

# GPU Implementation of JPEG2000 for Hyperspectral Image Compression

Milosz Ciznicki<sup>a</sup>, Krzysztof Kurowski<sup>a</sup> and <sup>b</sup>Antonio Plaza

<sup>a</sup>Poznan Supercomputing and Networking Center  
Noskowskiego 10, 61-704 Poznan, Poland.

<sup>b</sup>Hyperspectral Computing Laboratory,  
Department of Technology of Computers and Communications,  
University of Extremadura, Avda. de la Universidad s/n.  
10071 Cáceres, Spain

## ABSTRACT

Hyperspectral image compression has received considerable interest in recent years due to the enormous data volumes collected by imaging spectrometers for Earth Observation. JPEG2000 is an important technique for data compression which has been successfully used in the context of hyperspectral image compression, either in lossless and lossy fashion. Due to the increasing spatial, spectral and temporal resolution of remotely sensed hyperspectral data sets, fast (onboard) compression of hyperspectral data is becoming a very important and challenging objective, with the potential to reduce the limitations in the downlink connection between the Earth Observation platform and the receiving ground stations on Earth. For this purpose, implementation of hyperspectral image compression algorithms on specialized hardware devices are currently being investigated. In this paper, we develop an implementation of the JPEG2000 compression standard in commodity graphics processing units (GPUs). These hardware accelerators are characterized by their low cost and weight, and can bridge the gap towards on-board processing of remotely sensed hyperspectral data. Specifically, we develop GPU implementations of the lossless and lossy modes of JPEG2000. For the lossy mode, we investigate the utility of the compressed hyperspectral images for different compression ratios, using a standard technique for hyperspectral data exploitation such as spectral unmixing. In all cases, we investigate the speedups that can be gained by using the GPU implementations with regards to the serial implementations. Our study reveals that GPUs represent a source of computational power that is both accessible and applicable to obtaining compression results in valid response times in information extraction applications from remotely sensed hyperspectral imagery.

**Keywords:** Hyperspectral image compression, JPEG2000, commodity graphics processing units (GPUs).

## 1. INTRODUCTION

Hyperspectral imaging instruments are capable of collecting hundreds of images, corresponding to different wavelength channels, for the same area on the surface of the Earth.<sup>1</sup> For instance, NASA is continuously gathering imagery data with instruments such as the Jet Propulsion Laboratory's Airborne Visible-Infrared Imaging Spectrometer (AVIRIS), which is able to record the visible and near-infrared spectrum (wavelength region from 0.4 to 2.5 micrometers) of reflected light in an area 2 to 12 kilometers wide and several kilometers long, using 224 spectral bands.<sup>2</sup>

One of the main problems in hyperspectral data exploitation is that, during the collection of hyperspectral data, several GBs of multidimensional data volume is generated and set to the ground stations on Earth.<sup>3</sup> The downlink connection between the observation stations and receiving ground stations is limited, so different compression methods are employed.<sup>4</sup> Up to date, lossless compression techniques are the tool of choice, but the best lossless compression ratio reported in hyperspectral image analysis is around 3:1.<sup>4</sup> Due to the increasing spatial, spectral and temporal resolution of remotely sensed hyperspectral data sets, techniques with better

---

Send correspondence to Antonio J. Plaza:

E-mail: aplaza@unex.es; Telephone: +34 927 257000 (Ext. 51662); URL: <http://www.umbc.edu/rssipl/people/aplaza>

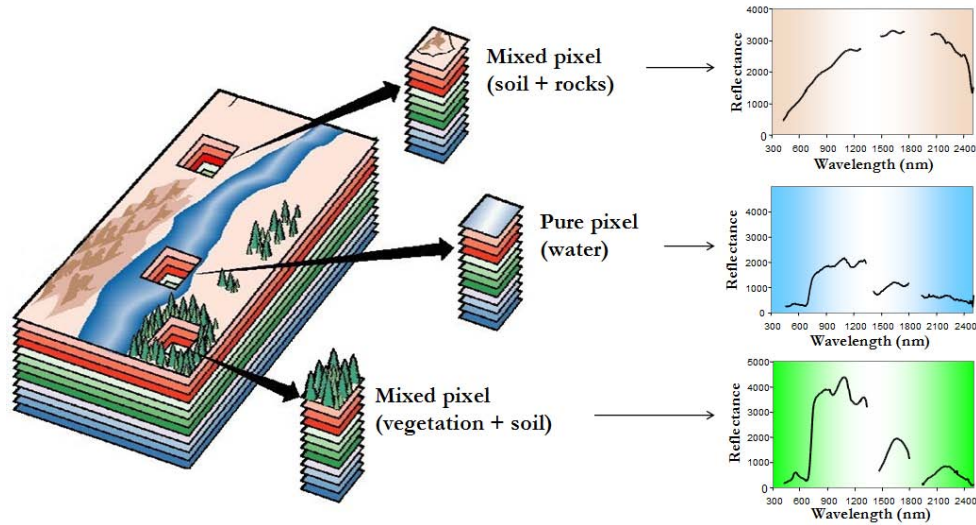


Figure 1. The mixture problem in hyperspectral data analysis.

compression ratios are needed and lossy compression becomes a reasonable alternative.<sup>5</sup> It turns out that JPEG2000<sup>6</sup> has been successfully used in the context of hyperspectral image compression, either in lossless and lossy fashion. Hence, it can be used to evaluate the impact of lossy compression on different techniques for hyperspectral data exploitation.

An important issue that has not been widely investigated in the past is the impact of lossy compression on spectral unmixing applications,<sup>7</sup> which are the tool of choice in order to deal with the phenomenon of mixed pixels,<sup>8</sup> i.e. pixels containing different macroscopically pure spectral substances, as illustrated in Fig. 1. In hyperspectral images, mixed spectral signatures may be collected due to several reasons. First, if the spatial resolution of the sensor is not fine enough to separate different pure signature classes at a macroscopic level, these can jointly occupy a single pixel, and the resulting spectral measurement will be a composite of the individual pure spectra, often called *endmembers* in hyperspectral analysis terminology.<sup>9</sup> Second, mixed pixels can also result when distinct materials are combined into a homogeneous or intimate mixture, and this circumstance occurs independently of the spatial resolution of the sensor.<sup>7</sup>

Although the unmixing chain maps nicely to high performance computing systems such as commodity clusters,<sup>10</sup> these systems are difficult to adapt to on-board processing requirements introduced by applications with real-time constraints such as wild land fire tracking, biological threat detection, monitoring of oil spills and other types of chemical contamination. In those cases, low-weight integrated components such as commodity graphics processing units (GPUs)<sup>11</sup> are essential to reduce mission payload. In this regard, the emergence of GPUs now offers a tremendous potential to bridge the gap towards real-time analysis of remotely sensed hyperspectral data.<sup>12-18</sup>

In this paper we develop an implementation of the JPEG2000 compression standard in commodity graphics processing units (GPUs) for hyperspectral data exploitation. Specifically, we develop GPU implementations of the lossless and lossy modes of JPEG2000. For the lossy mode, we investigate the utility of the compressed hyperspectral images for different compression ratios, using spectral unmixing as a case study. We also investigate the speedups that can be gained by using the GPU implementations with regards to the serial implementations in both the lossless and lossy modes. The remainder of the paper is organized as follows. Section 2 presents the JPEG2000 compression framework. Section 3 presents its GPU implementation. Section 4 first presents the hyperspectral data sets used for evaluation purposes, then briefly introduces the considered hyperspectral unmixing chain, and finally analyzes the proposed GPU implementation of JPEG2000 in terms of both unmixing accuracy (in the lossy mode) and computational performance (in both modes). Section 5 concludes with some remarks and hints at plausible future research.

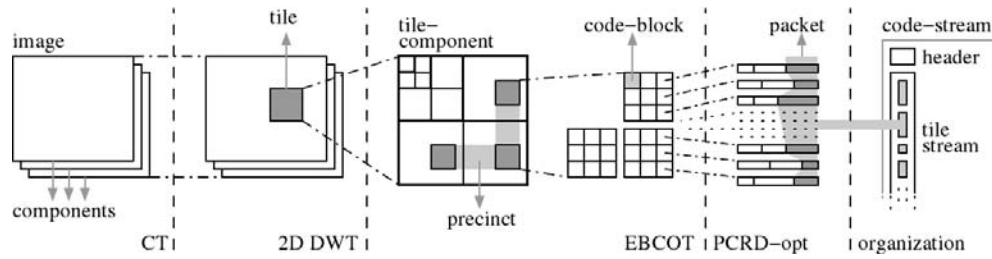


Figure 2. JPEG2000 data partitioning, coding and code-stream organization.

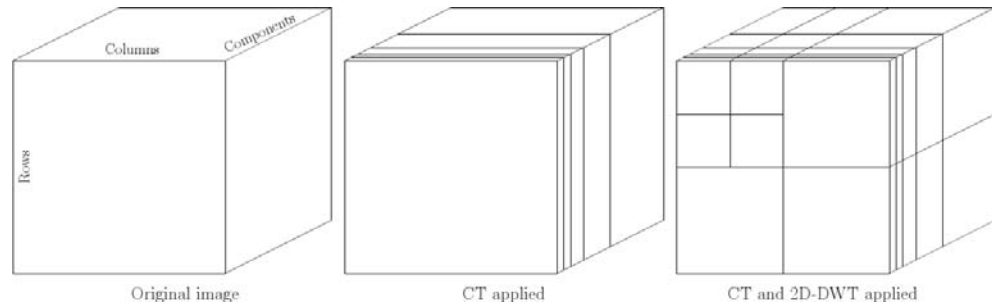


Figure 3. The hybrid scheme for 3D decorrelation. 4 levels for the CT and 2 levels for the ST.

## 2. OVERVIEW OF JPEG2000

The JPEG2000 standard is divided into several and incremental parts. Part 1<sup>19</sup> defines the core coding system and two basic image file formats. Part 2<sup>20</sup> specifies a set of extensions to the core coding system, such as spectral (inter-component) decorrelation and the use of different wavelet kernels, as well as a more extensible file format. These two first parts are the ones that are going to focus on along this paper. The rest of parts introduce some extensions for different applications. For example, Part 10 (named also JP3D)<sup>21</sup> is concerned with the coding of three-dimensional (3D) data.

Figure 2 shows an overview of the data partitioning made by the core coding system of JPEG2000, and how the elements of a three-component image (such as a color RGB image) are encoded and distributed. The first step in the JPEG2000 algorithm, not shown in the figure, is a level offset to guarantee that all the samples are signed. This is a requirement of the transform machinery that is going to be applied. After that, a CT (Component Transform, for example, a RGB to a YUV color transform) removes the inter-component redundancy that could be found in the image. The result of this stage is a new image in other domain, with the same number of components and samples per component. Next, as can be observed in the figure, each component of the image is divided into rectangular areas called tiles. Tiling is useful for compound images because the encoding parameters of each tile can be selected taking into account its characteristics. However, the tiling is rarely used in natural images due to the artifacts produced around the edges of the tiles, having commonly only one tile per component.

JPEG2000 allows the use of different decomposition patterns in the component domain, although the default one is the hybrid scheme (see Fig. 3). In the hybrid decomposition, a dyadic 1D-DWT (Discrete Wavelet Transform) is first applied to the component domain, and then a dyadic 2D-DWT, denoted by ST (Spatial Transform) in the rest of this paper, is applied to each tile-component. Currently, the dyadic DWT is widely used in the processing of scalable image contents because it facilitates the resolution scalability and improves the encoding efficiency, removing the intra-component (spatial) redundancy. In the example of Fig. 2 we can see how four different resolution levels are generated (remarked in blue) when a ST of three iterations is applied to a tile. These resolution levels are commonly referred by positive integer numbers starting from 0 for the highest one, the original image size.

JPEG2000 provides two working modes: lossy and lossless. The first one offers a better encoding efficiency at low bit-rates. When no loss of information is allowed, the lossless mode can be selected. A fixed point (integer) version of the DWT is used in this case. The resultant wavelet coefficients are grouped into rectangular areas (e.g.  $64 \times 64$  coefficients) called code-blocks, that are encoded independently by the EBCOT (Embedded Block

Coding with Optimal Truncation) algorithm.<sup>22</sup> In order to manage the image information more easily, the code-blocks related to the same rectangular location, within the same resolution level, are grouped into precincts. ROI (Region Of Interest) and the spatial scalabilities are achieved in JPEG2000 by means of the precincts. The compressed bit-streams of the code-blocks can be divided into a specific number of contiguous segments, or quality layers, by the PCRDOpt (Post-Compression Rate-Distortion Optimization) rate-allocation algorithm. The segments of all the code-blocks of a precinct associated to the same quality layer are stored in a packet. The packet is the storing unit in JPEG2000 and it is associated to a quality layer (L), a precinct (P), a resolution level (R), and a tile-component (C). The word formed by this four letters specifies the progression order used to store the image packets, existing five different possibilities: LRCP, RLCP, RPCL, PCRL and CPRL.

The distortion of the decoded image decreases as the amount of decoded data (packets) increases. The LRCP progression provides a fine-grain quality scalability mode. A sequentially decoded RLCP or RPCL image will produce a reconstruction of incremental resolution. The PCRL progression is useful in scan-based systems, like printers. Finally, an CPRL compressed image will be restored, component by component.

The most basic file format defined in the standard, in Part 1, contains only the code-stream of an image (see Fig. 2). This is composed by all the packets of the image and a set of markers with additional information. Markers can be located at any place in the code-stream, however the most important are included in the header. The image files with this format usually have the extension J2C or J2K. Part 1 also defines a more complex file format based on “boxes”. This format allows the coder to include additional information such as color palettes or meta-data. All the information is organized in boxes, contiguous segments of data, whose content is identified by a four-bytes code located at its header. It is possible to define a complex hierarchical structure since a box can contain many other boxes. The extension used to identify the image files with this format is JP2.

The box-based structure of the JP2 format is extensible. Just defining new four-bytes identifiers would allow to include new kind of boxes within an image file, maintaining the backward compatibility (an image viewer that does not understand certain box codes it just ignores them). Part 2 defines a new set of boxes with additional and powerful functionalities. For instance, multiple code-streams can be included within a file, as well as a complex composition scheme (animations, transparency masks, geometric transformations, user definable wavelet kernels, multi-component processing (CT), etc.), which will determine how the image decoding and displaying must be performed. The extension JPX is used to identify those files that contain boxes of Part 2. This is the file format used in our experiments.

### 3. GPU IMPLEMENTATION

GPUs can be abstracted by assuming a much larger availability of processing cores than in standard CPU processing, with smaller processing capability of the cores and small control units associated to each core (see Fig. 4). Hence, the GPU is appropriate for algorithms that need to execute many repetitive tasks with fine grain parallelism and few coordination between tasks. In the GPU, algorithms are constructed by chaining so-called *kernels*, which define the minimum units of computations performed in the cores. Thereby, data-level parallelism is exposed to hardware, and kernels can be concurrently applied without any sort of synchronization. The kernels can perform a kind of batch processing arranged in the form of a grid of blocks, as displayed in Fig. 5, where each block is composed by a group of threads which share data efficiently through the shared local memory and synchronize their execution for coordinating accesses to memory. There is a maximum number of threads that a block can contain but the number of threads that can be concurrently executed is much larger (several blocks executed by the same kernel can be managed concurrently, at the expense of reducing the cooperation between threads since the threads in different blocks of the same grid cannot synchronize with the other threads). Finally, Fig. 6 shows the architecture of the GPU, which can be seen as a set of multiprocessors. Each multiprocessor is characterized by a single instruction multiple data (SIMD) architecture, i.e., in each clock cycle each processor of the multiprocessor executes the same instruction but operating on multiple data streams. Each processor has access to a local shared memory and also to local cache memories in the multiprocessor, while the multiprocessors have access to the global GPU (device) memory.

As mentioned in Section 2 the JPEG2000 standard contains several encoding steps which are done in consecutive manner. The first one is the level offset which is performed on samples of components that are unsigned



only. This procedure involves only subtraction of the same quantity from all samples and as a result is bound to memory transfer on GPU. In order to obtain high efficiency every thread on GPU is responsible for calculations of several samples. It occurred that  $16 \times 16$  blocks of threads where each threads calculates values for four samples gives the best results. After that the component transform is applied to component domain using 1D-DWT. The 1D-DWT can be realized by iteration of filters with rescaling. This kind of implementation has high complexity, need a lot of memory and computational power. The better way is to use the lifting-based wavelet transform. Lifting-based filtering is done by using four lifting steps, which updates alternately odd or even sample values. During this process the spectral vector data is decomposed to the low pass (even) samples and the high pass (odd) samples. The low pass samples contains the most of the information and high pass samples account the residual information. As a result the high pass samples can be discarded during the compression processes and thus reduce file size. In the GPU implementation every thread is responsible for calculating and loading several samples to the shared memory. During the lifting process, all the samples at the edge of the shared memory array depend on samples which were not loaded to the shared memory. Around a data chunk within a thread block, there is a margin of samples that is required in order to calculate the component chunk. The margin of one block overlaps with adjacent blocks. The width of each margin depends on the side of the data chunk. In order to avoid idle threads, data from the margin is loaded by threads within block. Furthermore, the margins of the blocks on the edges of the image are symmetrically extended to avoid large errors at the boundaries. Before the lifting process samples are reordered to access array elements that are adjacent in the global memory. Next encoding step is tiling. Here the components from the hyperspectral dataset may be tiled, which means that they can be divided into several rectangular non-overlapping blocks, called tiles. The tiles at the edges are sometimes smaller, if tile size is not an integral multiple of the component size. A main advantage of tiling is less memory usage and different tiles can be processed parallel. It could be usefull for very large images, which different tiles could be processed independently on separate GPUs. A serious disadvantage of tiling is that artifacts can appear at the edges of the tiles. As hyperspectral data set easily fits into GPU memory no tiling is used during compression.

The subsequent step in compression process is the ST. The ST can be irreversible or reversible. The irreversible transform is implemented by means of the Daubechies filter (lossy transform). The reversible transformation is implemented by means of the Le Gall filter (lossless transform). This enables an intra-component decorrelation that concentrates the image information in a small and very localized area. By itself the ST does not compress image data. It restructures the image information so it is easier to compress it. During the ST the tile data is decomposed into the horizontal and vertical characteristics. This transform is similar to 1D-DWT in nature, but it is applied in the horizontal (rows) and the vertical (columns) directions which forms two-dimensional transform. Similar to the 1D-DWT, the component data block is loaded to the shared memory, however the processing is done on columns and rows during one kernel invocation including data reordering. As a result a number of kernel invocations and calls to the global memory is reduced. In order to get additional performance improvement every thread read and synchronize several samples from global to shared memory. Furthermore, all threads read adjacent sample values from the shared memory to registers, which are needed to correctly compute output samples. When registers are used, each thread from the block is able to calculate two sample values at one time. It gives more speedup, as memory transactions can be overlapped by arithmetic operations. After the lifting procedure samples are scaled and written to the global memory.

Quantization is only applied in the case of lossy compression. After the ST, all the resulting subbands are quantized, that means that floating point numbers are transformed into integers. Quantization is the main source of information loss, and therefore very important to obtain good compression rates. The quantizer maps several values that are in the range of some interval to one integer value. This results in a reduction of the bit-depth, thus compression. It involves only few computations, as a result each threads is responsible for quantization of 16 samples.

After quantization, the integer wavelet coefficients still contain a lot of redundancy and symmetries. This redundancy is removed by entropy coding and so the data is efficiently packed into a minimal size bit-stream. The problem with highly-compressed, entropy coded data is that few bit errors could completely corrupt the image information. This would be a big problem when JPEG2000 data is transmitted over a noisy communication channel, so each wavelet subband in subdivided into small code-blocks with typical sizes of  $32 \times 32$  or  $64 \times 64$ .

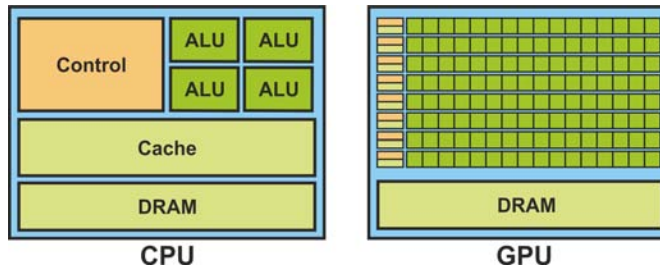


Figure 4. Comparison of CPU versus GPU architecture.

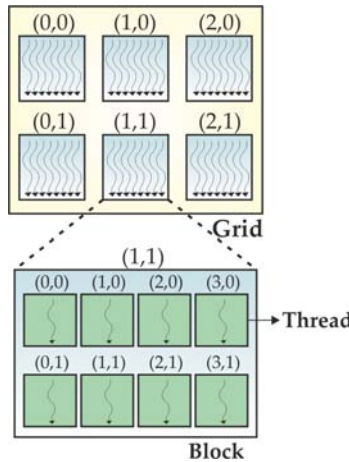


Figure 5. Processing in the GPU: grids made up of blocks with computing threads.

Each of these code-blocks is entropy coded separately which gives potential for parallelization. The process of entropy coding is highly sequential and difficult to efficiently parallelize to more threads, therefore each thread on GPU do entropy coding on whole code-block. However even for small input components it gives enough work to fill all multiprocessors with computations. For instance the input component with size  $512 \times 512$  and code-blocks size of  $32 \times 32$  will be spread to  $256/8 = 32$  blocks of threads.

The easiest way of rate control would be to change the precision of samples in the quantization process. Naturally, this is only possible for lossy compression. The disadvantage of this way is the high computation consumption, because after every change of precision of the samples, the whole entropy encoding has to be repeated. A much more elegant way is to use the Post-compression rate-distortion algorithm which generates optimal truncation points to minimize the distortion while still obtaining the target bit rate. PCR algorithm allows to compress hyperspectral data with target bitrate. The algorithm is based on calculation the distortion connected with including next bytes from code-blocks to the output stream. The main idea is to find for the given target output size the total sum of encoded bytes which minimizes the distortion. The calculation of distortion is based on bitplane which is actually encoded. For 16 bit precision hyperspectral component there are 16 bitplanes, from (highest) most significant to (lowest) less significant bitplane. If bit from higher bitplane is skipped during the encoding process, it introduces more distortion. Therefore bits from higher bitplanes have greater chance to be found in output file, because algorithm minimizes summary distortion. Similar to entropy coding each thread on GPU calculates distortions connected with one code-block, because it contains small number of computations.

The last step in compression process is creating and ordering the packets. This basically consists of writing the file and creating the progression order. At the end of the computations all the data have to be saved on the host memory, as a result this step is executed on CPU.

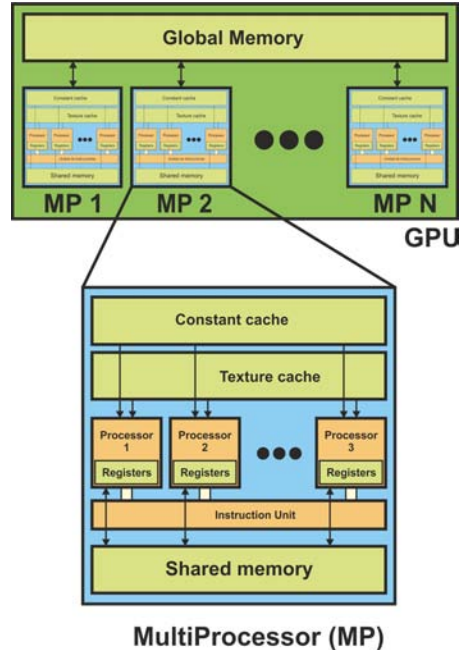


Figure 6. Hardware architecture of the GPU.

## 4. EXPERIMENTAL RESULTS

### 4.1 Hyperspectral data set

The hyperspectral data set used in this study is the well-known AVIRIS Cuprite data set, available online\* in reflectance units after atmospheric correction. This scene has been widely used to validate the performance of endmember extraction algorithms. The portion used in experiments corresponds to a  $350 \times 350$ -pixel subset of the sector labeled as `f970619t01p02.r02.sc03.a.rf1` in the online data. The scene comprises 224 spectral bands between  $0.4$  and  $2.5 \mu\text{m}$ , with full width at half maximum of  $10 \text{ nm}$ . Each reflectance value in the scene is represented using 16 bits, for a total image size of approximately 43.92 MB. Prior to the analysis, several bands (specifically, bands 1–2, 105–115 and 150–170) were removed due to water absorption and low SNR in those bands, leaving a total of 188 reflectance channels to be used in the experiments. The Cuprite site is well understood mineralogically,<sup>23,24</sup> and has several exposed minerals of interest including those used in the USGS library considered for the generation of simulated data sets. Five of these laboratory spectra (*alunite*, *buddingtonite*, *calcite*, *kaolinite* and *muscovite*) convolved in accordance with AVIRIS wavelength specifications, will be used to assess endmember signature purity in this work. For illustrative purposes, Fig. 7 shows the image data set considered in experiments and the USGS library signatures.

### 4.2 Hyperspectral unmixing chain

In order to evaluate the impact of lossy compression on hyperspectral data quality, we follow an exploitation-based approach which consists of applying a spectral unmixing chain to the decompressed hyperspectral data at different quality levels. Let  $\mathbf{x}$  be a hyperspectral pixel vector given by a collection of values at different wavelengths. In the context of linear spectral unmixing,<sup>25</sup> such vector can be modeled as:

$$\mathbf{x} \approx \mathbf{E}\mathbf{a} + \mathbf{n} = \sum_{i=1}^p \mathbf{e}_i a_i + \mathbf{n}, \quad (1)$$

where  $\mathbf{E} = \{\mathbf{e}_i\}_{i=1}^p$  is a matrix containing  $p$  pure spectral signatures (endmembers),  $\mathbf{a} = [a_1, a_2, \dots, a_p]$  is a  $p$ -dimensional vector containing the abundance fractions for each of the  $p$  endmembers in  $\mathbf{x}$ , and  $\mathbf{n}$  is a noise

\*<http://aviris.jpl.nasa.gov/html/aviris.freedata.html>

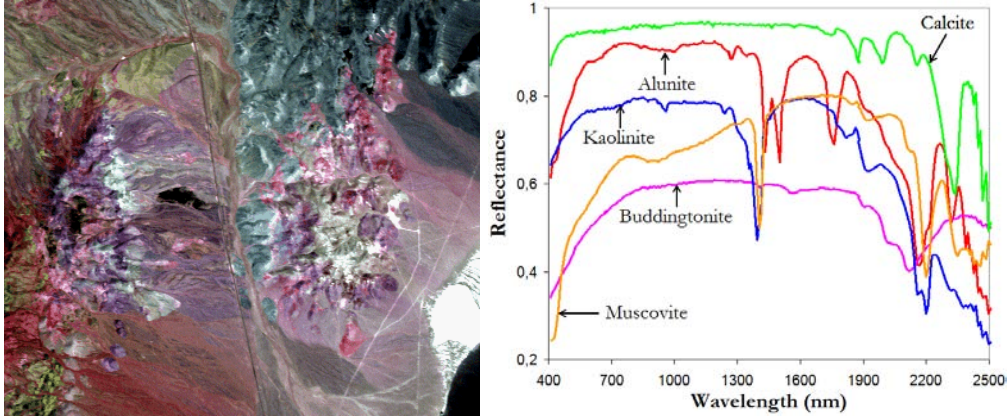


Figure 7. (a) Hyperspectral image over the AVIRIS Cuprite mining district in Nevada. (b) USGS reference signatures used to assess endmember signature purity.

term. Solving the linear mixture model involves: 1) estimating the right subspace for the hyperspectral data, 2) identifying a collection of  $\{\mathbf{e}_i\}_{i=1}^p$  endmembers in the image, and 3) estimating their abundance in each pixel of the scene. In this work, the dimensional reduction step is performed using principal component analysis (PCA),<sup>26,27</sup> a popular tool for feature extraction in different areas including remote sensing. For the endmember selection part, we rely on the well-known N-FINDR algorithm,<sup>28</sup> which is a standard in the hyperspectral imaging community. Finally, abundance estimation is carried out using unconstrained least-squares estimation<sup>25</sup> due to its simplicity. This unmixing chain has been shown to perform in real-time in different GPU platforms.<sup>11</sup> It should be noted that evaluating the accuracy of abundance estimation in real analysis scenarios is very difficult due to the lack of ground-truth information, hence in this work we only evaluate the endmember extraction accuracy of the considered spectral unmixing chain.

### 4.3 Impact of lossy compression on unmixing accuracy

In this experiment we evaluate the impact of applying lossy JPEG2000-based compression to the considered hyperspectral image in terms of the degradation of unmixing accuracy as we increase the compression ratio. In our experiments, the full scene has been compressed (using different compression ratios) from 0.2 to 2.0 bits per pixel per band (bpppb). This ranges from 80:1 compression ratio (0.2 bpppb) to 2 bpppb (8:1 compression ratio). Our original idea was to compress the hyperspectral data in two different manners: spectral domain and spatial domain. In the spectral domain every pixel (vector) is compressed separately, so the spectral information is utilized. On the other hand compressing pixels separately does not take advantage of redundancy between spatially adjacent pixels. Spatial domain-based compression is done on separate hyperspectral bands. This strategy utilizes spatial data redundancy, but is not specifically intended to preserve spectral properties. Subsequently, in this work we focus mainly on spatial domain decomposition as it exploits spatial information (an important source of correlation in hyperspectral images) and can be applied to the full hyperspectral image without truncation.

The quantitative and comparative algorithm assessment that is performed in this work is intended to measure the impact of compression on the quality of the endmembers extracted by N-FINDR algorithm in the considered unmixing chain. Specifically, the quality of the endmembers extracted from the decompressed scene (i.e., the one obtained after coding/decoding the scene using JPEG2000) is measured in terms of their spectral similarity with regards to the five reference USGS spectral signatures in Fig. 7(b). The spectral similarity between an endmember extracted from the original scene,  $\mathbf{e}_i$ , and a USGS reference signature  $\mathbf{s}_j$ , is measured by the spectral angle (SA), a well known metric for hyperspectral data processing<sup>7</sup> which is defined as follows:

$$SA(\mathbf{e}_i, \mathbf{s}_j) = \cos^{-1} \frac{\mathbf{e}_i \cdot \mathbf{s}_j}{\|\mathbf{e}_i\| \|\mathbf{s}_j\|}. \quad (2)$$

It should be noted that the SA is given by the arc cosine of the spectral angle formed by  $n$ -dimensional vectors. As a result, this metric is invariant in the multiplication of  $\mathbf{e}_i$  and  $\mathbf{s}_j$  by constants and, consequently,

Table 1. Spectral angle values (in degrees) between the pixels extracted by N-FINDR algorithm from the AVIRIS Cuprite scene (using different compression ratios) and the USGS reference signatures.

Compression	Alunite	Buddingtonite	Calcite	Kaolinite	Muscovite
No compression	4,81°	4,16°	9,52°	10,76°	5,29°
2.0 bpppb	7,44°	7,03°	12,31°	13,45°	8,16°
1.8 bpppb	7,93°	7,58°	12,84°	14,01°	8,57°
1.6 bpppb	8,40°	8,55°	13,36°	14,59°	9,12°
1.4 bpppb	9,22°	9,34°	13,92°	15,26°	9,61°
1.2 bpppb	9,95°	10,06°	14,44°	15,80°	10,23°
1.0 bpppb	10,76°	10,59°	15,20°	16,43°	10,99°
0.8 bpppb	11,17°	11,06°	15,79°	17,12°	11,88°
0.6 bpppb	12,03°	11,94°	16,50°	17,91°	12,60°
0.4 bpppb	13,12°	13,02°	17,61°	19,04°	13,72°
0.2 bpppb	14,31°	14,23°	18,88°	20,11°	13,90°

Table 2. Real-time compression results obtained on the NVidia GeForce GTX 480 GPU for the AVIRIS Cuprite image using both lossless and lossy compression modes.

Input image size	Lossless compression			Lossy compression		
	Output image size	Ratio	bpppb	Output image size	Ratio	bpppb
43.92 MB	25.00 MB	0.57	9.11	21.97 MB	0.5	8.00

is invariant before unknown multiplicative scalings that may arise due to differences in illumination and angular orientation.<sup>7</sup> For illustrative purposes, Table 1 shows the SA values (in degrees) measured as the compression ratio was increased (the lower the SA, the better the obtained results, with results in the order of 15 degrees considered sufficiently similar in spectral terms<sup>9</sup>). As we can observe in our preliminary assessment, the quality of endmember extraction decreases for higher compression ratios as expected, with values of 1.2 bpppb (around 13:1 compression ratio) still providing acceptable results in terms of spectral similarity for all considered USGS minerals. Further experiments should be conducted in order to assess the impact of additional endmember extraction and abundance estimation algorithms for different compression ratios.

#### 4.4 Analysis of parallel performance

Our GPU implementation was evaluated using the same AVIRIS Cuprite scene addressed in previous subsections. In our experiments, we used a CPU with Intel Core 2 Duo E8400 with 3,00GHz processor and 6GB RAM. The GPU platform used for evaluation purposes was the NVidia GeForce GTX 480, which features 480 processor cores. For the GPU implementation, we used the the compute unified device architecture (CUDA) as the development environment.

Our main goal in experiments was to assess the possibility of obtaining compression results in real-time, as the main interest of onboard compression using specialized hardware devices is to reduce the limitations imposed by the downlink connection when sending the data to a control station on Earth. As a result, real-time compression (as the data is collected by the sensor) is highly desirable. It should be noted that the cross-track line scan time in AVIRIS, a push-broom instrument,<sup>2</sup> is quite fast (8.3 milliseconds to collect 512 full pixel vectors). This introduces the need to compress the considered AVIRIS Cuprite scene (350 × 350 pixels and 188 spectral bands) in less than 1985 milliseconds to fully achieve real-time performance. In our experiments, we observed that the proposed GPU implementation of JPEG2000 was able to compress hyperspectral data sets in valid response time for both lossless and lossy compression. The time measured for lossless compression in the GPU was 1580 milliseconds, while the time measured for lossy compression (with 8 bpppb) in the GPU was 1490 milliseconds. The time difference between the lossless and lossy mode results from compression technique used in JPEG2000. During the lossy mode the quantization is used which considerably reduces the accuracy of the pixels and thus speedups the encoder. Whereas in lossless mode quantization is not used, so all information included in pixels are compressed by the encoder. Therefore duration of lossless compression is an upper bound for compression times with different ratios.

On the other hand, Table 2 shows the compression ratios and number of bpppb's obtained in real-time compression of the AVIRIS Cuprite data set. As it can be seen in Table 2, lossless compression provides approximately 2:1 compression ratio. Due to increasing spatial and spectral resolution of remotely sensed hyperspectral data sets, this ratio is not sufficient for onboard compression. On the other hand, in our experiments we observed that lossy compression (with 8 bpppb) is also able to provide 2:1 compression ratio in real-time. Obviously, further developments are required in order to increase the lossy compression ratios that can be achieved in real-time to fully benefit from lossy compression in order to optimize downlink data transmission.

## 5. CONCLUSIONS AND FUTURE LINES

In this paper we have developed a GPU implementation of JPEG2000 intended for real-time compression of remotely sensed hyperspectral images, with the ultimate goal of designing a specialized hardware implementation that can be used onboard hyperspectral imaging instruments. Although the power consumption rates of GPUs are still much higher than those provided by other hardware devices such as field programmable gate arrays (FPGAs), we believe in the future potential of GPU hardware accelerators for onboard hyperspectral image compression. Specifically, we have developed GPU implementations for the lossless and lossy compression modes of JPEG2000. For the lossy mode, we investigated the utility of the compressed hyperspectral images for different compression using a standard techniques for hyperspectral data exploitation such as linear spectral unmixing. In both cases we investigated the processing times that can be obtained using the GPU implementations, concluding that real-time performance can be obtained around 2:1 compression ratios in both cases. Further developments are required in order to increase the lossy compression ratios that can be achieved in real-time. In order to achieve this goal, we are planning on using the post-compression rate-distortion (PCRD) algorithm.<sup>29</sup> It allows to optimally truncate codestream depending on given target size. PCRD algorithm calculates distortion associated with truncation points. The process of calculating distortion is based on weights which are connected to the image characteristics. Therefore it is very important to choose suitable weights which reflect the image properties.

## ACKNOWLEDGEMENT

This work has been supported by the European Community's Marie Curie Research Training Networks Programme under reference MRTN-CT-2006-035927 (HYPER-I-NET). Funding from the Spanish Ministry of Science and Innovation (HYPERCOMP/EODIX project, reference AYA2008-05965-C04-02) is gratefully acknowledged.

## REFERENCES

1. A. F. H. Goetz, G. Vane, J. E. Solomon, and B. N. Rock, "Imaging spectrometry for Earth remote sensing," *Science* **228**, pp. 1147–1153, 1985.
2. R. O. Green, M. L. Eastwood, C. M. Sarture, T. G. Chrien, M. Aronsson, B. J. Chippendale, J. A. Faust, B. E. Pavri, C. J. Chovit, M. Solis, *et al.*, "Imaging spectroscopy and the airborne visible/infrared imaging spectrometer (AVIRIS)," *Remote Sensing of Environment* **65**(3), pp. 227–248, 1998.
3. A. Plaza and C.-I. Chang, *High Performance Computing in Remote Sensing*, Taylor & Francis: Boca Raton, FL, 2007.
4. G. Motta, F. Rizzo, and J. A. Storer, *Hyperspectral data compression*, Springer-Verlag, New York, 2006.
5. Q. Du and C.-I. Chang, "Linear mixture analysis-based compression for hyperspectral image analysis," *IEEE Trans. Geosci. Rem. Sens.* **42**, pp. 875–891, 2004. [doi:10.1109/IGARSS.2000.861638].
6. D. S. Taubman and M. W. Marcellin, *JPEG2000: Image compression fundamentals, standard and practice*, Kluwer, Boston, 2002.
7. N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Processing Magazine* **19**(1), pp. 44–57, 2002.
8. J. B. Adams, M. O. Smith, and P. E. Johnson, "Spectral mixture modeling: a new analysis of rock and soil types at the Viking Lander 1 site," *Journal of Geophysical Research* **91**, pp. 8098–8112, 1986.

9. A. Plaza, P. Martinez, R. Perez, and J. Plaza, "A quantitative and comparative analysis of endmember extraction algorithms from hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing* **42**(3), pp. 650–663, 2004.
10. A. Plaza, J. Plaza, A. Paz, and S. Sanchez, "Parallel hyperspectral image and signal processing," *IEEE Signal Processing Magazine* **28**(3), pp. 119–126, 2011.
11. S. Sanchez, A. Paz, G. Martin, and A. Plaza, "Parallel unmixing of remotely sensed hyperspectral images on commodity graphics processing units," *Concurrency and Computation: Practice and Experience* **23**(12), 2011.
12. C. A. Lee, S. D. Gasster, A. Plaza, C.-I. Chang, and B. Huang, "Recent developments in high performance computing for remote sensing: A review," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **4**(3), 2011.
13. S.-C. Wei and B. Huang, "GPU acceleration of predictive partitioned vector quantization for ultraspectral sounder data compression," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **4**(3), 2011.
14. H. Yang, Q. Du, and G. Chen, "Unsupervised hyperspectral band selection using graphics processing units," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **4**(3), 2011.
15. J. A. Goodman, D. Kaeli, and D. Schaa, "Accelerating an imaging spectroscopy algorithm for submerged marine environments using graphics processing units," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **4**(3), 2011.
16. E. Christophe, J. Michel, and J. Inglada, "Remote sensing processing: From multicore to GPU," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **4**(3), 2011.
17. J. Mielikainen, B. Huang, and A. Huang, "GPU-accelerated multi-profile radiative transfer model for the infrared atmospheric sounding interferometer," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **4**(3), 2011.
18. C.-C. Chang, Y.-L. Chang, M.-Y. Huang, and B. Huang, "Accelerating regular LDPC code decoders on GPUs," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **4**(3), 2011.
19. International Organization for Standardization, "Information Technology - JPEG2000 Image Coding System - Part 1: Core Coding System." ISO/IEC 15444-1:2004, 2004.
20. International Organization for Standardization, "Information Technology - JPEG2000 Image Coding System - Part 2: Extensions." ISO/IEC 15444-2:2004, 2004.
21. International Organization for Standardization, "Information technology - JPEG2000 image coding system: Extensions for three-dimensional data." ISO/IEC 15444-10:2008, 2008.
22. D. Taubman, "High performance scalable image compression with EBCOT," *IEEE Trans. Image Process.* **9**, pp. 1158–1170, 2000. [doi:10.1109/83.847830].
23. R. N. Clark, G. A. Swayze, K. E. Livo, R. F. Kokaly, S. J. Sutley, J. B. Dalton, R. R. McDougal, and C. A. Gent, "Imaging spectroscopy: Earth and planetary remote sensing with the usgs tetracorder and expert systems," *Journal of Geophysical Research* **108**, pp. 1–44, 2003.
24. G. Swayze, R. N. Clark, F. Kruse, S. Sutley, and A. Gallagher, "Ground-truthing AVIRIS mineral mapping at Cuprite, Nevada," *Proc. JPL Airborne Earth Sci. Workshop* , pp. 47–49, 1992.
25. D. Heinz and C.-I. Chang, "Fully constrained least squares linear mixture analysis for material quantification in hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing* **39**, pp. 529–545, 2001.
26. D. A. Landgrebe, *Signal Theory Methods in Multispectral Remote Sensing*, John Wiley & Sons: New York, 2003.
27. J. A. Richards and X. Jia, *Remote Sensing Digital Image Analysis: An Introduction*, Springer, 2006.
28. M. E. Winter, "N-FINDR: An algorithm for fast autonomous spectral endmember determination in hyperspectral data," *Proceedings of SPIE* **3753**, pp. 266–277, 1999.
29. D. Taubman and M. Marcellin, *JPEG2000 Image Compression Fundamentals, Standards and Practice*, Springer, Berlin, 2002.





# Keeping it simple: the Alabama Digital Preservation Network (ADPNet)

Keeping it simple

Aaron Trehub

*Auburn University Libraries, Auburn University, Auburn, Alabama, USA, and*

Thomas C. Wilson

*University of Alabama Libraries, University of Alabama, Tuscaloosa, Alabama, USA*

245

Received 8 February 2010  
Revised 9 February 2010  
Accepted 15 February 2010

## Abstract

**Purpose** – The purpose of this paper is to present a brief overview of the current state of distributed digital preservation (DDP) networks in North America and to provide a detailed technical, administrative, and financial description of a working, self-supporting DDP network: the Alabama Digital Preservation Network (ADPNet).

**Design/methodology/approach** – The paper reviews current regional and national initiatives in the field of digital preservation using a variety of sources and considers ADPNet in the context of generally accepted requirements for a robust DDP network. The authors view ADPNet in a comparative perspective with other Private LOCKSS Networks (PLNs) and argue that the Alabama model represents a promising approach to DDP for other states and consortia.

**Findings** – The paper finds that cultural memory organizations in a number of countries have identified digital preservation as a critical issue and are crafting strategies to address it, with DDP-based solutions gaining in popularity in North America. It also identifies an array of technical, administrative, and financial challenges that DDP networks must resolve in order to be viable in the long term.

**Practical implications** – The paper describes a working model for building a low-cost but robust DDP network.

**Originality/value** – The paper is one of the first comprehensive descriptions of a working, self-sustaining DDP network.

**Keywords** Digital storage, Information networks, United States of America, Archiving

**Paper type** Case study

If the first decade of the twenty-first century was the decade of mass digitization, the second decade looks likely to become the decade of digital preservation. Although precise figures are hard to come by, it is generally recognized that most of the world's information is currently being produced in digital form, not as print documents or analog artifacts. This poses a serious challenge to libraries, archives, museums, and other cultural memory organizations, as well as government departments and public agencies. Unlike their tangible counterparts, digital files are inherently susceptible to decay, destruction, and disappearance. Given the vulnerability of digital content to fires, floods, hurricanes, power blackouts, cyber attacks, and a variety of hardware and software failures, cultural memory organizations need to begin incorporating long-term digital preservation services for locally created digital content into their routine operations, or risk losing that content irrevocably. The advent of a "digital dark age" is not just a clever conceit; it is a real danger.





---

How do we care for these new digital resources, which run the gamut from government web sites and institutional records to digital archives, scanned images, and born-digital audio and video recordings? A number of countries have recognized the challenge and embarked on ambitious digital preservation programs at the national level. In the United States, the Library of Congress initiated the National Digital Information Infrastructure and Preservation Program (NDIIPP) almost ten years ago, and recently launched the National Digital Stewardship Alliance (NDSA) (Library of Congress, 2002; Library of Congress, 2010). In the United Kingdom, the Digital Curation Centre of the Joint Information Systems Committee (JISC) provides a national focus for digital preservation issues (DCC, 2010). Similar initiatives are underway in Canada, Australia, New Zealand, France, Germany, and other European countries (IIPC, 2010).

Several lessons have already emerged from these initiatives. One of them is the importance of collaboration among institutions, states, and even countries. In digital preservation as in many other endeavors, there is strength in numbers. With numbers comes complexity, however, and comprehensive digital preservation programs inevitably raise difficult technical, administrative, financial, and even legal questions. That said, these questions are not unresolvable. Indeed, they are being resolved, or at least addressed, by a number of preservation programs in the United States, Canada, and other countries. There is a growing body of experience that shows that it is possible to build technically and administratively robust digital preservation networks across institutional and geographical borders without compromising those networks' long-term viability through excessive complexity and cost.

### **Distributed Digital Preservation (DDP) and LOCKSS**

One especially promising approach combines Distributed Digital Preservation (DDP) with LOCKSS ("Lots of copies keep stuff safe") software in Private LOCKSS Networks (PLNs).

As its name implies, DDP is based on the idea of distributing copies of digital files to server computers at geographically dispersed locations in order to maximize their chances of surviving a natural or man-made disaster, power failure, or other disruption. DDP networks consist of multiple preservation sites, selected with the following principles in mind:

- sites preserving the same content should not be within a 75-125-mile radius of one another;
- preservation sites should be distributed beyond the typical pathways of natural disasters, such as hurricanes, typhoons, and tornadoes;
- preservation sites should be distributed across different power grids;
- preservation sites should be under the control of different systems administrators;
- content preserved in disparate sites should be on live media and should be checked on a regular basis for bit-rot and other issues; and
- content should be replicated at least three times in accordance with the principles detailed above (Skinner and Mevenkamp, 2010, pp. 12-13).

---

LOCKSS was developed and is currently maintained at the Stanford University Libraries (LOCKSS, 2010a). It is ideally suited for use in DDP networks. Originally designed to harvest, cache, and preserve digital copies of journals for academic libraries, LOCKSS is also effective at harvesting, caching, and preserving locally created digital content for cultural memory organizations in general. LOCKSS servers (also called LOCKSS boxes or LOCKSS caches) typically perform the following functions:

- they collect content from target web sites using a web crawler similar to those used by search engines;
- they continually compare the content they have collected with the same content collected by other LOCKSS Boxes, and repair any differences;
- they act as a web proxy or cache, providing browsers in the library's community with access to the publisher's content or the preserved content as appropriate; and
- they provide a web-based administrative interface that allows the library staff to target new content for preservation, monitor the state of the content being preserved, and control access to the preserved content. (LOCKSS, 2010b)

Although LOCKSS is open source software and therefore available for further development by the open source community, its maintenance and development have been concentrated in the LOCKSS development team at Stanford.

Private LOCKSS Networks (PLNs) represent a specialized application of the same software that runs the public LOCKSS network. Simply defined, a PLN "is a closed group of geographically distributed servers (known as 'caches' in LOCKSS terminology) that are configured to run the open source LOCKSS software package" (Skinner and Mevenkamp, 2010). PLNs are secure peer-to-peer networks. All the LOCKSS caches in a PLN have the same rights and responsibilities; once they have been set up, they can run indefinitely and independently of a central server. If the digital content on a PLN cache is destroyed or lost, it can be restored from the other caches in the network; content can also be restored to the server from which it was originally harvested. Like the public LOCKSS network, PLNs rely on LOCKSS' built-in polling and voting mechanism to monitor the condition of the content in the network and to repair damaged or degraded files. Since LOCKSS is format-agnostic, PLNs may collect and archive any type of digital content. Prior to ingest, the content may require reorganization in order to be effectively harvested into the private network – a process often referred to as "data wrangling". In order to ensure robustness and operate effectively, a PLN should consist of at least six geographically dispersed LOCKSS caches.

LOCKSS has a well-deserved reputation for being easy to install and configure, and installing the LOCKSS software on a single server in the public network is indeed a fairly simple procedure (a YouTube video suggests that it can be done in about five minutes: Dresselhaus, 2007). Setting up and running a PLN is a considerably more complex undertaking, however. There are a number of technical questions that must be addressed at the outset, including the choice of operating system and LOCKSS installation (OpenBSD or Linux/CentOS, Live CD or package installation), the type of hardware to be used (identical or different server makes and configurations), network

---

capacity (typically in terabytes – more is better), the number of members in the network (preferably six or more), and the responsibility for creating and hosting such essential LOCKSS components as the manifest pages, plugins, plugin repository, and title database. In addition to these and other technical issues, there are also questions related to network governance, administration, and sustainability. In some ways, governance questions are the most difficult to resolve.

---

Fortunately, there is a growing body of practical experience and governance models that prospective PLNs can draw on. The first PLN in North America was the MetaArchive Cooperative, which was established in 2004 with funding from the Library of Congress's National Digital Information Infrastructure and Preservation Program (NDIIPP). The Cooperative was the first successful attempt to apply LOCKSS to the preservation of locally created digital content – in this case, content pertaining to the history and cultures of the American South. The Cooperative began with six universities in the southeastern United States: Auburn University, Emory University, Florida State University, the Georgia Institute of Technology, the University of Louisville, and the Virginia Polytechnic Institute and State University. In the years since its inception, the Cooperative has more than doubled its original membership and now serves fourteen member institutions in the United States, the United Kingdom, and Brazil. In 2008, the Cooperative formed a 501(c)(3) organization – the Atlanta, Georgia-based Educopia Institute – to manage the network (MetaArchive, 2010).

As of January 2010, there were at least five other PLNs in operation in the United States and Canada, including the Alabama Digital Preservation Network (ADPNet), the Persistent Digital Archives and Library System (PeDALS), the Data Preservation Alliance for the Social Sciences (Data-PASS)/Inter-University Consortium for Political and Social Research (ICPSR), the US Government Documents PLN, and the Council of Prairie and Pacific University Libraries (COPPUL) PLN (ADPNet, 2010a; PeDALS, 2010); Data-PASS, 2010; USDocs, 2010; COPPUL, 2010). Most of these networks are so-called “dark archives”, meaning that access to the cached content is restricted to the network participants.

### **The Alabama Digital Preservation Network (ADPNet)**

The Alabama Digital Preservation Network (ADPNet) is a geographically distributed digital preservation network for the state of Alabama – the first working single-state PLN in the United States. Its origins go back to a series of conversations in late 2005 at the Network of Alabama Academic Libraries (NAAL), a department of the Alabama Commission on Higher Education in Montgomery. Inspired by Auburn University's experience as a founding member of the NDIIPP-supported MetaArchive Cooperative, the directors of six other Alabama libraries and archives agreed to pool resources to build a Private LOCKSS Network for the state if external start-up funding could also be obtained. With this commitment in hand, NAAL Director Sue Medina and Auburn University Director of Library Technology Aaron Trehub submitted a funding proposal to the US Institute of Museum and Library Services (IMLS) at the beginning of 2006. The proposal was funded, and work began on the network at the end of 2006 under a two-year IMLS National Leadership Grant. The grant provided support for equipment and associated expenses to the seven participating institutions; crucially, it also covered those institutions' annual membership fees in the LOCKSS Alliance for the same period. For their part, the participating institutions split the equipment costs

---

with IMLS and contributed staff time and other in-house resources to the project. A LOCKSS staff member was assigned to the project to provide technical support and guidance.

Alabama was an attractive candidate for a DDP network for several reasons. The first is the frequency of hurricanes, tornadoes, flooding, and other natural disasters, especially on and around Alabama's Gulf coast. In the past decade, Alabama has been hit by at least four major hurricanes and many more tropical storms. In 2005, Hurricane Katrina devastated the coastal communities of Bayou la Batre and Coden and flooded downtown Mobile. The coastal communities are not the only parts of the state that have suffered from natural disasters, however. The interior of the state is vulnerable to tornadoes: in March 2007 a tornado swept through Enterprise, Alabama, destroying a high school and causing ten deaths.

The second factor is Alabama's financial situation. Alabama is a relatively poor state, ranking 47th out of 51 states and territories in median household income in 2008 (US Census Bureau, 2008). The worldwide financial crisis of 2008-2009 exacerbated the state's economic difficulties and led to cutbacks in the state budget for higher education. There isn't much state money available for new initiatives, which means that technical solutions have to be simple, robust, and above all inexpensive to implement and maintain.

Finally, Alabama is home to a growing number of digital collections at libraries, archives, and museums. Many of these collections trace their creation back to 2001, when NAAL received a three-year National Leadership Grant from the IMLS to develop a shared statewide digital repository on all aspects of Alabama's history, geography, and culture. This repository eventually became AlabamaMosaic (AlabamaMosaic, 2010). AlabamaMosaic currently contains over 25,000 digital objects from eighteen institutions around the state, and the number continues to grow.

The growing number of digital collections in Alabama threw the problem of digital preservation into relief. Although academic librarians took the lead in AlabamaMosaic, an important part of the project was encouraging all types of cultural memory organizations in Alabama to digitize their unique materials and contribute them to the online repository. Organizations participating in the project were expected to maintain their own archival files and to follow current best practices for preserving their digital masters in archival conditions. It soon became clear, however, that archival storage and digital preservation were simply not in the lexicon of the smaller institutions, not to mention quite a few of the larger ones. The smaller institutions in particular were often staffed by volunteers and lacked the funding for off-site storage solutions or the technical resources and expertise for server-based backup routines. For their part, faculty members and students at colleges and universities throughout the state stored their unique digital files on desktop PCs, with rudimentary or no backup and no provision for long-term preservation.

In short, AlabamaMosaic revealed that the ability to create new digital content had far outstripped awareness of the need to preserve it. Alabama is not unique in this respect. A 2005 survey conducted by Liz Bishoff and Tom Clareson for the Northeast Document Conservation Center (NEDCC) revealed that 66 percent of the institutions surveyed did not have a staff member specifically assigned to digital preservation and 30 percent of the institutions had digital collections that had been backed up only once or not at all (Bishoff, 2007).

---

This combination of circumstances – extreme weather conditions, meagre state financial resources, and a growing number of rich but vulnerable digital collections – has made Alabama an ideal test-case for a LOCKSS-based DDP solution. The IMLS grant ended in September 2008, and ADPNet is now a self-sustaining, member-managed DDP network operating under the auspices of NAAL. The network currently has seven member institutions: the Alabama Department of Archives and History (ADAH), Auburn University, Spring Hill College, Troy University, the University of Alabama, the University of Alabama at Birmingham, and the University of North Alabama. With seven locally managed caches in different parts of the state and on different power grids, the Alabama network meets most of the functional criteria for distributed digital preservation networks listed at the beginning of this article.

### **ADPNet: technical aspects**

Like other PLNs, ADPNet had to resolve a number of technical questions at the beginning of the project. Briefly, these involved picking hardware for the LOCKSS caches, choosing an operating system and agreeing on a LOCKSS installation method, and deciding whether to rely on the LOCKSS team at Stanford to host the network's LOCKSS plugin repository and title database or to manage those crucial network components locally.

In each case, the ADPNet members opted for the simplest and least expensive solution. In the case of hardware, they followed the recommendation of the LOCKSS liaison person and purchased low-cost hardware – one-terabyte servers with two hard drives – from a small manufacturer in San Jose, California. In LOCKSS' experience, smaller vendors are slower to move their inventory and thus tend to have slightly older chips installed in their machines – a benefit where LOCKSS is concerned, because the open source community is more likely to have developed drivers for those chips. Furthermore, the members agreed to purchase identical servers, in order to normalize costs across institutions, minimize time spent on configuration and testing, and make trouble-shooting easier.

Having chosen to purchase identical, low-cost servers from the same manufacturer, the ADPNet institutions chose to use the OpenBSD operating system and the Live CD installation of the LOCKSS platform. Under this scenario, the OpenBSD OS runs from a CD-ROM, while the configuration parameters are stored on write-protected media. In ADPNet's case, the write-protected media was a flash drive connected to a USB port on the cache computers. The flash drive included a write-protect option so that the configuration data could not be accidentally overwritten and to provide an appropriate level of security. This installation method was attractive to the Alabama network members because it did not require high-level IT expertise to implement. Moreover, this method is inherently secure because the operating system loads and runs from non-writable media. Finally, this type of installation has the advantage of ensuring automatic updates to new versions of the LOCKSS daemon.

Like most of the other PLNs in North America, ADPNet decided to have the LOCKSS team host the network's plugin repository and title database. Plugins provide specific instructions to the LOCKSS software for harvesting digital content into the network. They are usually written by, or in cooperation with, the content contributor (a content contributor in a PLN is a member site that prepares collections for harvesting). The title database connects the LOCKSS plugins to their corresponding archival units (AUs), making the AUs selectable for harvest.

---

ADPNet began with a network capacity of one terabyte at each of the seven LOCKSS caches. This seemed adequate for the early stages of the project, when the focus was on getting the network up and running and testing its basic viability. As the network moved from being a grant-supported project to an ongoing, self-supporting program, the member institutions identified a need for more network storage space. In 2009, the network members purchased and installed servers with eight terabytes of storage, expandable to 16 terabytes. They also transitioned from the OpenBSD operating system and the LOCKSS Live CD installation to the CentOS operating system and the LOCKSS package installation. In addition to more network capacity, it also became clear that ADPNet needed a better tool for keeping track of the archival units being harvested into the network. Accordingly, ADPNet is now looking at deploying the conspectus database developed by the MetaArchive Cooperative to organize and manage collection-level metadata in DDP networks – just one example of how PLNs learn from each other.

Although the transition to larger caches and a different operating system reflects the network's growing maturity and technical sophistication, it has also highlighted different levels of technical expertise and financial resources among the network members. Managing these differences while preserving the integrity of the network is one of several challenges facing ADPNet as it moves forward (more on this below).

#### **ADPNet: governance, membership, and costs**

ADPNet's mission is fourfold: to provide a reliable, low-cost, low-maintenance preservation network for the long-term preservation of publicly available digital resources created by Alabama libraries, archives, and other cultural memory organizations; to promote awareness of the importance of digital preservation throughout the state; to create a geographically dispersed "dark archive" of digital content that can be drawn upon to restore collections at participating institutions in the event of loss or damage; and to serve as a model and resource to other states and consortia that are interested in setting up digital preservation networks of their own.

ADPNet was designed from the outset primarily to be a simple, inexpensive digital preservation solution for libraries, archives, museums, and other cultural memory organizations in Alabama. This emphasis on simplicity and cost-effectiveness is reflected in the ADPNet governance documents. Devising a governance structure that meets the needs of all the network members while ensuring long-term sustainability is perhaps the most difficult part of setting up a PLN. Here ADPNet benefited from the fact that one of its members had participated in drafting the original MetaArchive Cooperative Charter and was able to draw on that experience to craft a similar document for the Alabama network. The ADPNet Governance Policy was formally adopted by the NAAL Advisory Council at its annual business meeting in October 2008. The policy can be found on the ADPNet web site, along with the ADPNet Application for Membership and the ADPNet Technical Specifications (ADPNet, 2010b).

ADPNet has a lightweight governance structure consisting of two committees: the ADPNet Steering Committee and the ADPNet Technical Policy Committee. The Steering Committee consists of an elected chair and an appointed representative from each of the member institutions. The chair has a term of one year and the position rotates among the member institutions. In addition to spreading the experience of



---

coordinating a PLN among the member institutions, this arrangement ensures that no one institution can dominate the network. The Technical Policy Committee consists of IT specialists appointed by the member institutions; it oversees the network's technical needs and makes recommendations to the Steering Committee. Together, the two committees are responsible for the day-to-day management of ADPNet. The NAAL Advisory Council exercises general oversight of ADPNet and considers and votes on policy recommendations submitted to it by the Steering Committee.

In keeping with the network's guiding principles, the requirements for membership are as simple and inexpensive as the members could make them. Membership in ADPNet is open to all libraries, archives, and cultural memory organizations in Alabama. In addition to that basic condition, there are three other requirements for membership: participating institutions must agree to install and run a LOCKSS server in the network, contribute content to the network, and join the LOCKSS Alliance for an annual fee. There is a provision in the governance policy that allows the Steering Committee to waive even these minimal requirements in the case of small, poorly resourced organizations that have digital content in urgent need of harvest and preservation. Finally and perhaps most importantly, there is no ADPNet membership fee.

ADPNet has tried to keep financial and in-kind expenses to an absolute minimum. Communication costs are negligible. There is an ADPNet listserv for e-mail correspondence and a monthly conference call for voice-to-voice communication. There is also an annual face-to-face meeting in Alabama, usually held just before the yearly business meeting of the NAAL Advisory Council. Other meetings may be called by the chair of the ADPNet Steering Committee, at his or her discretion. The amount of staff time required to keep the network running is also extremely modest, typically amounting to a couple of hours of IT staff time per institution per month. This can increase during equipment upgrades or when preparing new digital content for harvest into the network, but so far none of the ADPNet members have indicated that the staff time required is beyond their capacities. The major expenses for the network to date have been hardware upgrades and the annual LOCKSS Alliance fee, which is based on the institutions' Carnegie classifications and can range from \$1,080 per year for community colleges to \$10,800 per year for large research universities (LOCKSS, 2010c).

### **ADPNet in comparative perspective**

Work is underway to describe and classify the digital preservation landscape. Tyler Walters of the Georgia Institute of Technology has posited three governance models for distributed digital preservation networks:

- (1) The Participant Model (e.g. ADPNet).
- (2) The Network Administrative Model (e.g. the MetaArchive Cooperative and the CLOCKSS Archive).
- (3) The Lead Organization Model (e.g. the PeDALS PLN led by the Arizona State Library and Archives; and the SRB-based Chronopolis Network led by the University of California at San Diego) (Walters, 2010, pp. 38-40).

Although ADPNet was originally inspired by and has some similarities with the MetaArchive Cooperative, there are important differences between the two networks.

---

First and most importantly, the Alabama network is a single-state PLN. This has simplified governance and allowed the network to be absorbed into an existing legal and administrative entity (NAAL), one with bylaws and a committee structure already in place.

Second, the Alabama network was designed to be a practical solution to a pressing statewide problem, not a research-and-development project. In order to attract participants, ADPNet had to be simple, robust, and above all inexpensive. This, and the fact that only one or two institutions in Alabama had had any prior experience with LOCKSS, meant that the members opted for the simplest, least expensive hardware and software solutions available, in the hope that these would be easier to deploy and manage and more attractive to other institutions in the state.

That said, ADPNet has made some technical contributions to the larger PLN community. One example is the collaboration between the Auburn University Libraries and the LOCKSS team to develop a plugin for harvesting content hosted in collections that use CONTENTdm, a popular digital content management software package marketed by OCLC. The procedure requires exporting the metadata in the standard Dublin Core (RDF) format, then saving the resulting file at the same web location as the LOCKSS manifest page. Each CONTENTdm collection is defined using a volume parameter, allowing one plugin to harvest many collections. As a result, the display objects (images, PDFs, etc.) for the collection can be harvested and cached in LOCKSS together with their associated metadata. CONTENTdm compound documents cause special concerns due to the way CONTENTdm stores the metadata for individual “pages” within a document. Also, full-resolution collections may contain pointers to the archival images in the metadata. These full-resolution images may or may not be hosted in the same location as the presentation images. Further developments are planned to support harvesting collections of compound documents in CONTENTdm, as well as harvesting archival images for full-resolution collections. If JPEG2000 is accepted as a standard preservation format, CONTENTdm’s ability to use this format for presentation will simplify the process of archiving and caching high-quality archival images in a PLN.

Another example of technical innovation comes from work done at the University of Alabama on scalable, straightforward ways to restructure the normally chaotic file-naming conventions and storage practices that make up digital repositories. The idea is to harness the file and directory structure within an operating system to mirror holding organizations, collections, items, and delivery sequences. Digital objects, associated metadata, and documentation are stored at whatever level is appropriate. The resulting hierarchy is both human- and machine-parsable (DeRidder, 2009). This type of structure supports running scripts for verifying correct file names, validating and confirming content files (e.g. the JSTOR/Harvard Object Validation Environment: JHOVE, 2010), and generating LOCKSS manifests for archival units that are ready to be harvested.

Finally, ADPNet is not a fee-based service organization. Rather, the network is intended to complement AlabamaMosaic, the other statewide initiative that was originally started with an IMLS grant but that has been kept going by in-kind contributions from its participating institutions. In other words, ADPNet was designed to run on relatively small expenditures and on sweat equity, not on recurring infusions of grant money or annual membership fees. To some degree these differences reflect



---

Alabama's institutional culture, which is extremely expense-averse. They also reflect a preference for simplicity and informality where administrative arrangements are concerned.

### Challenges

As was mentioned above, ADPNet's transition from a grant-supported project to a self-sustaining program has highlighted a number of management issues within the network. Chief among these are ensuring equitable network use, promoting technical diversity and geographic dispersion, attracting new members, and achieving long-term financial sustainability.

One of the ADPNet's strengths is that it serves a range of institutions, from large research universities to small liberal arts colleges. The network also includes the state archives, a government agency. This diverse membership suggests that the ADPNet model is suitable for similarly heterogeneous groups of institutions. It has, however, thrown the question of equitable network use into relief. This problem is common to any shared network that serves different categories of institutions. Large institutions typically have large amounts of digital content that they want to preserve in the network; small institutions have smaller amounts of content. Since the network has to be large enough to accommodate the preservation needs of its largest members, smaller institutions may find themselves subsidizing storage capacity that they do not need and cannot use. The challenge is to apportion costs fairly among the network members according to network use.

Thomas C. Wilson, the Associate Dean for Library Technology at the University of Alabama and the current chair of the ADPNet Steering Committee, has opened a discussion on how to ensure equitable use of the network. One idea is to allow the member institutions to make their own hardware purchases, or to repurpose or cannibalize hardware they already have. LOCKSS does not require that PLN members use identical equipment. The ADPNet members, however, agreed at the beginning of the IMLS grant to use the same hardware in order to minimize compatibility problems and make troubleshooting easier, especially since only one ADPNet institution had had substantial experience with LOCKSS at that time. Now that all the network members are familiar with LOCKSS and its technical requirements, having identical hardware is less critical. The smallest member of the network, Spring Hill College in Mobile, has already started building its own ADPNet server from surplus equipment. Other ADPNet members may follow suit. In addition to saving money that would otherwise be spent on new hardware, instituting a policy of hardware diversity will also test LOCKSS' scalability in heterogeneous networks.

The members have also discussed moving to a two-tiered network: a network of large caches for the large institutions, paid for and administered by those institutions, and a parallel network of smaller caches for the small institutions. This idea is attractive because most of the original one-terabyte servers that ADPNet started with are still available and could be used to build the small network. Choosing this alternative, however, would add an element of complexity to the network administration; it might also undermine the sense of institutional solidarity that ADPNet has succeeded in building over the past four years. A variation on this theme would be to create one or more caches on a single-tiered network that would be shared among two or more smaller institutions each. In this way, the burden for

---

administration of a server would be lessened and the overall cost of maintaining the cache and covering the LOCKSS fees would be shared as well. In this scenario, it would also be possible to host the separate caches at one or more of the larger member institutions, completely eliminating the need for smaller organizations to provide system support.

Other ideas include devising a mechanism to compensate sites that do not contribute a lot of content to the network, but still have to support capacity needed for the total amount of content that is being preserved; instituting a graduated fee system for storing content above a certain amount (i.e. a tax on content); or setting up a contingency fund, perhaps through NAAL, that could be drawn upon when the network needs to expand. These are all financial and possibly political issues, and they promise to occupy the network in the coming year. The one point of agreement is that ADPNet needs to adopt a phased schedule for expansion that will manage the members' concerns about size and cost.

In most ecological systems, diversity promotes survivability. In PLNs, diversity implies technical diversity and geographic dispersion (Rosenthal *et al.*, 2005). The ADPNet members are moving toward one type of technical diversity: using different hardware within the network. Thanks to Auburn University's participation in the MetaArchive Cooperative, ADPNet is also poised to benefit from that PLN's collaboration with the Chronopolis Digital Preservation Demonstration Project at the University of California at San Diego (Chronopolis, 2010). The goal is to build a bridge between LOCKSS-based PLNs and Chronopolis' large-scale data-grid approach to distributed digital preservation, thereby enhancing the survivability of both types of networks.

In order to increase the Alabama network's geographic dispersion, ADPNet and the Edmonton, Alberta-based Council of Prairie and Pacific University Libraries (COPPUL) PLN have discussed hosting servers in their respective locations, an initiative that has tentatively been dubbed "The Alberta Project". Similar discussions have taken place with representatives of the Arizona-based Persistent Digital Archives and Library System (PeDALS) PLN.

ADPNet is committed to expanding the network by attracting new members in Alabama. This will add to the store of preserved digital content and enhance the network's robustness; it will also increase awareness of and experience with digital preservation throughout the state. The network is especially interested in attracting museums, public libraries, and local historical societies – institutions that are not currently represented in ADPNet. One obstacle to achieving this goal has been the annual LOCKSS Alliance fee, which many smaller non-academic institutions regard as prohibitive. For this reason, the ADPNet Steering Committee is exploring alternatives with the LOCKSS management, including alternate fee schedules or a consortial payment system.

An overarching issue is the question of long-term financial sustainability. As indicated above, several options are being pursued to ensure the financial attractiveness of ADPNet to other cultural memory organizations in the state of Alabama. In the long term, having a larger membership should reduce the financial burden on all participants, but it may take a while before ADPNet achieves that level of saturation. In addition to raising the awareness of Alabamians with regard to the need for digital preservation, ADPNet must also focus attention on the real costs of such activities. However small we may be able to make these costs, they are not zero.

---

Some options for remaking the financial model do exist. For example, NAAL could establish an ADPNet fund that would be seeded by the organizations that surpass a well-defined quota of preservation space in the network. This approach, in conjunction with negotiating a quasi-consortial fee schedule with LOCKSS, could move ADPNet to a model that not only sustains at current levels of preservation, but also supports ongoing increases in the quantity and size of the archival units being preserved.

These are somewhat formidable challenges indeed, but the same spirit and pragmatism that led to the creation of ADPNet will rise to address them.

### **The future**

At its inception, ADPNet identified four specific tasks. First, to highlight the importance of preserving digital content among libraries, academic institutions, state agencies, and other cultural heritage institutions in Alabama. Second, to demonstrate the feasibility of state-based, low-cost models for digital preservation by creating a working example of such a network in Alabama. Third, to create an administrative structure to manage the network and assure its long-term sustainability. And fourth, to demonstrate that the network can support different types of digital content from different types of institutions, from public libraries and small colleges to large state agencies.

The network has achieved all four tasks. On the technical side, ADPNet has been up and running since 2007. All seven member institutions have contributed content to the network, and almost fifty digital collections have been harvested to date. These consist primarily of archival audio, video, and still image files and include the Alabama Department of Archives and History World War I Gold Star Database; the Auburn University Historical Maps Collection and Sesquicentennial Lecture Series; the Troy University Postcard Collection; the University of Alabama 1968 Student Government Association Emphasis Symposium, with audio files of historic speeches by Robert F. Kennedy, Ferenc Nagy, and John Kenneth Galbraith; the University of Alabama at Birmingham Oral History Collection; and digital collections from the University of North Alabama on local historian William McDonald and the US Nitrate Plant in Muscle Shoals. ADPNet plans to harvest several terabytes of new content into the network in 2010. The network also hopes to recruit new member institutions in the coming year.

Online surveys have shown that ADPNet has succeeded in raising awareness of the importance of digital preservation among Alabama libraries, archives, and state agencies. The task now is to translate this increased awareness into broader participation in ADPNet.

Although ADPNet's main mission is to build and sustain a robust, inexpensive distributed digital preservation network for Alabama, the members also hope that it will serve as a model for similar networks in other states and countries and as a low-cost alternative to more-expensive digital preservation solutions. In the past few years, the ADPNet team has shared its experience with academic librarians in Virginia, the state libraries of Montana and Nevada, museum directors in Colorado and Oklahoma, a consortium of Canadian research libraries, and staff members from a regional library consortium in the southeast. In 2008-2009 alone, ADPNet representatives gave presentations about the network at the NDIIPP partners meeting in Washington, DC; the LITA national forum in Cincinnati, Ohio; the DigCCurr

conference in Chapel Hill, North Carolina; the annual meeting of the Society of American Archivists in Austin, Texas; the Best Practice Exchange (BPE) meeting in Albany, New York; and the iPres conference in San Francisco, California. LOCKSS team members have also promoted ADPNet as an exemplary model for state-based or regional DDP networks.

If ADPNet had an official motto, it would be “keep it simple and keep it cheap”. This approach appears to have worked well so far in Alabama. It remains to be seen whether it will work for other states and consortia, but the signs are encouraging.

## References

- Alabama Digital Preservation Network (ADPNet) (2010a), “About ADPNet”, available at: <http://adpn.org/index.html> (accessed 10 February 2010).
- Alabama Digital Preservation Network (ADPNet) (2010b), “Resources”, available at: <http://adpn.org/resources.html> (accessed 10 February 2010).
- AlabamaMosaic (2010), “About AlabamaMosaic”, available at: [www.alabamamosaic.org/about.php](http://www.alabamamosaic.org/about.php) (accessed 10 February 2010).
- Bishoff, L. (2007), “Digital preservation assessment: readying cultural heritage institutions for digital preservation”, available at: [www.ils.unc.edu/digccurr2007/slides/bishoff\\_slides\\_8-3.pdf](http://www.ils.unc.edu/digccurr2007/slides/bishoff_slides_8-3.pdf) (accessed 10 February 2010).
- Chronopolis (2010), “Federated digital preservation across space and time”, available at: <http://chronopolis.sdsc.edu/index.html> (accessed 10 February 2010).
- Council of Prairie and Pacific University Libraries (COPPUL) (2010), “Who are we?”, available at: [www.coppul.ca/](http://www.coppul.ca/) (accessed 11 February 2010).
- Data Preservation Alliance for the Social Sciences (Data-PASS) (2010), “About Data-PASS”, available at: [www.icpsr.umich.edu/icpsrweb/DATAPASS/](http://www.icpsr.umich.edu/icpsrweb/DATAPASS/) (accessed 11 February 2010).
- DeRidder, J.L. (2009), “Preparing for the future as we build collections”, *Archiving 2009: Final Program and Proceedings, May 2009, Bern, Switzerland*, Vol. 6, Society for Imaging Science and Technology, Washington, DC, pp. 53-7.
- Digital Curation Centre (DCC) (2010), *Digital Curation Centre*, available at: [www.dcc.ac.uk](http://www.dcc.ac.uk) (accessed 9 February 2010).
- Dresselhaus, A.S. (2010), *Configuring a LOCKSS Box*, available at: [www.youtube.com/watch?v=0wdcnXrQkaI](http://www.youtube.com/watch?v=0wdcnXrQkaI) (accessed 9 February 2007).
- International Internet Preservation Consortium (IIPC) (2010), “About the consortium”, available at: <http://netpreserve.org/about/index.php> (accessed 9 February 2010).
- JSTOR/Harvard Object Validation Environment (JHOVE) (2010) (2010), “JSTOR/Harvard object validation environment”, available at: <http://hul.harvard.edu/jhove/> (accessed 10 February 2010).
- Library of Congress (2002), “Preserving our digital heritage: plan for the national digital information infrastructure and preservation program”, available at: [www.digitalpreservation.gov/library/resources/pubs/docs/ndiipp\\_plan.pdf](http://www.digitalpreservation.gov/library/resources/pubs/docs/ndiipp_plan.pdf) (accessed 9 February 2010).
- Library of Congress (2010), “Digital preservation”, available at: [www.digitalpreservation.gov](http://www.digitalpreservation.gov) (accessed 9 February 2010).
- LOCKSS (2010a), “Home”, available at: [www.lockss.org/lockss/Home](http://www.lockss.org/lockss/Home) (accessed 9 February 2010).
- LOCKSS (2010b), “How it works”, available at: [www.lockss.org/lockss/How\\_It\\_Works](http://www.lockss.org/lockss/How_It_Works) (accessed 9 February 2010).

- LOCKSS (2010c), "LOCKSS alliance participant invoice", available at: [www.lockss.org/locksswiki/files/LOCKSS\\_ALLIANCE\\_MEMBERSHIP\\_FORM.pdf](http://www.lockss.org/locksswiki/files/LOCKSS_ALLIANCE_MEMBERSHIP_FORM.pdf) (accessed 10 February 2010).
- MetaArchive Cooperative (MetaArchive) (2010), *The MetaArchive Cooperative*, available at: [www.metaarchive.org/](http://www.metaarchive.org/) (accessed 10 February 2010).
- Persistent Digital Archives and Library System (PeDALS) (2010), "About PeDALS", available at: [www.pedalspreservation.org/](http://www.pedalspreservation.org/) (accessed 11 February 2010).
- Rosenthal, D.S.H., Robertson, T.S., Lipkis, T., Reich, V. and Morabito, S. (2005), "Requirements for digital preservation systems: a bottom-up approach", *D-Lib Magazine*, Vol. 11 No. 11, available at: [www.dlib.org/dlib/november05/rosenthal/11rosenthal.html](http://www.dlib.org/dlib/november05/rosenthal/11rosenthal.html) (accessed 10 February 2010).
- Skinner, K. and Mevenkamp, M. (2010), "Chapter 2: DDP architecture", in Skinner, K. and Schultz, M. (Eds), *A Guide to Distributed Digital Preservation*, The Educopia Institute, Atlanta, GA, pp. 11-25.
- US Census Bureau (2008), "Table R1901: median household income (in 2008 inflation-adjusted dollars)", available at [http://factfinder.census.gov/servlet/GRTSelectServlet?ds\\_name=ACS\\_2008\\_1YR\\_G00\\_](http://factfinder.census.gov/servlet/GRTSelectServlet?ds_name=ACS_2008_1YR_G00_) (accessed 10 February 2010).
- US Government Documents PLN (USDocs) (2010), "Government documents PLN", available at: [www.lockss.org/lockss/Government\\_Documents\\_PLN](http://www.lockss.org/lockss/Government_Documents_PLN) (accessed 11 February 2010).
- Walters, T.O. (2010), "Chapter 4: Organizational considerations", in Skinner, K. and Schultz, M. (Eds), *A Guide to Distributed Digital Preservation*, The Educopia Institute, Atlanta, GA, pp. 37-48.

#### **About the authors**

Aaron Trehub is the Assistant Dean for Technology and Technical Services at the Auburn University Libraries in Auburn, Alabama, USA. Aaron Trehub is the corresponding author and can be contacted at: [trehuaj@auburn.edu](mailto:trehuaj@auburn.edu)

Thomas C. Wilson is the Associate Dean for Library Technology at the University of Alabama in Tuscaloosa, Alabama, USA.

# Fractal Analysis of Periapical Bone from Lossy Compressed Radiographs: A Comparison of Two Lossy Compression Methods

B. Güniz Baksi · Aleš Fidler

Published online: 5 April 2011  
© Society for Imaging Informatics in Medicine 2011

**Abstract** The aim of the study was to evaluate the effect of two lossy image compression methods on fractal dimension (FD) calculation. Ten periapical images of the posterior teeth with no restorations or previous root canal therapy were obtained using storage phosphor plates and were saved in TIF format. Then, all images were compressed with lossy JPEG and JPEG2000 compression methods at five compression levels, i.e., 90, 70, 50, 30, and 10. Compressed file sizes from all images and compression ratios were calculated. On each image, two regions of interest (ROIs) containing healthy trabecular bone in the posterior periapical area were selected. The FD of each ROI on the original and compressed images was calculated using differential box counting method. Both image compression and analysis were performed by a public domain software. Altogether, the FD of 220 ROIs was calculated. FDs were compared using ANOVA and Dunnett tests. The FD decreased gradually with compression level. A statistically significant decrease of the FD values was found for JPEG 10, JPEG2000 10, and JPEG2000 30 compression levels ( $p < 0.05$ ). At comparable file sizes, the JPEG induced a smaller FD difference. In conclusion, lossy compressed images with appropriate compression level may be used for FD calculation.

**Keywords** Compression · Computer analysis · Computer-assisted detection

## Introduction

A fractal analysis is a method for quantitative evaluation of complex geometric structures that exhibit patterns throughout the image. The complexity of the structure is represented by a single number, the fractal dimension (FD), which is calculated with a computer algorithm.[1] In medical radiology, the FD calculation is used to enhance the diagnosis of osteoporosis[2] or breast cancer.[3] In dental radiology, the FD calculation was used to evaluate and quantify a trabecular bone structure for the detection of bone changes associated with periapical periodontitis,[4,5] periodontal disease,[6] bone surgery,[7] and systemic diseases.[8,9] Several methods for FD calculation were proposed, with box counting method[10] being the most often used in dental radiology.[4]

Due to the benefits of digital radiography,[11] its use in dentistry is increasing, further facilitating the application of fractal analysis as images are readily available in digital format. However, storage and communication of digital images still remain a challenge.[12] Hardware requirements for picture archival and communication systems can be efficiently reduced by utilization of lossy image compression.[13] Two standardized lossy compression methods, namely JPEG[14] and JPEG2000[15] are widely accepted in dental radiography.[16] They offer considerably higher compression ratios compared to lossless compression, but on the cost of image information loss, adjusted by compression level. It is of utmost importance, that diagnostic accuracy of image is preserved. Therefore, to maximize file size reduction, the highest amount of image

---

B. G. Baksi  
School of Dentistry, Department of Oral Diagnosis and Radiology,  
Ege University,  
Izmir, Turkey  
e-mail: bgunb@yahoo.com

A. Fidler (✉)  
Department of Restorative Dentistry and Endodontics,  
Faculty of Medicine, University of Ljubljana,  
Hrvatski trg 6,  
1000 Ljubljana, Slovenia  
e-mail: ales.fidler@mf.uni-lj.si



information loss that is still preserving diagnostic accuracy should be determined and applied. Unfortunately, the amount of acceptable image information loss cannot be universally recommended as it is rather task specific.[13] In dental radiology, the compression ratio (CR) between 1:6.5 [17] and 1:28[18] was reported acceptable for visual interpretation.

Due to the concerns regarding diagnostic accuracy, the use of lossy compression is generally discouraged for computer-aided image evaluation methods, as they are more sensitive and consequently supposed to be more susceptible to compression-induced information loss. In contrast to this general belief, the accuracy of one computer-aided method, namely digital subtraction radiography (DSR) was not affected by lossy compression with CR of 1:7.[19] This was explained by the fact that a slight lossy image compression performs as a noise reduction filter. Fractal analysis, in comparison to DSR, was reported to be a more robust computer-aided image analysis method, insensible to variations in film exposure, limited image geometry variations and sizes and positions of region of interest (ROI).[1,20] However, the effect of lossy image compression on FD calculation has not been evaluated yet.

Therefore, the aim of the study was to evaluate the effect of two standard lossy image compression methods on FD calculation and to determine the highest acceptable degree of information loss, still preserving the diagnostic accuracy of FD calculation.

## Materials and Methods

### Radiographic Technique

Dry human mandibles, containing premolars and molars at least on one side, with no restorations or previous root canal therapy were selected. Specimens were radiographed with storage phosphor plates (SPP) of Digora® Optime (Soredex Corporation, Helsinki, Finland) system to ensure the absence of periapical pathology. Ten mandibles meeting the criteria were used in the study. An optical bench was used to standardize the projection geometry. Size 2 (31 × 41 mm) blue SPPs were exposed at a focus receptor distance of 25 cm with a Gendex Oralix DC (Gendex Dental Systems, Milan, Italy) dental X-ray unit operating at 60 kVp, 7 mA, and 1.5 mm Al equivalent filtration. The image plates were exposed for 0.12 s and scanned immediately after exposure in the Digora® Optime scanner with a matrix size of 620 × 476 pixels and resolution of 400 dpi. The acquired images were saved uncompressed in TIF format with Digora for Windows software (Soredex Corporation, Helsinki, Finland).

### Image Compression

Images were compressed with a public domain IrfanView software[21] with two lossy compression methods. The first, JPEG (JP) compression method is based on discrete cosine transformation of image tiles and discarding frequency information,[14] while the second, the JPEG2000 (J2) compression method is utilizing the discrete wavelet transformation and converts an image into series of wavelets.[19] Images were compressed for both compression methods at five different compression levels (CL) of 90, 70, 50, 30, and 10. A CL, sometimes referred to as quality factor, is a value from a scale from 100 to 1, where a higher number means a lower amount of image information loss. The average file sizes and compression ratio of compressed images were calculated for each CL and compression method.

### Fractal Dimension Calculation

On each original image, two nonoverlapping rectangular ROIs were selected in periapical trabecular bone not including roots or periodontal space. The positions and sizes of ROIs were determined according to the size and shape of the periapical region[22] resulting in sizes ranging between 3.77 and 118.15 mm<sup>2</sup>. Position and size of each ROI in original and corresponding compressed images was identical. In total, FD was calculated on 220 ROIs (20 ROIs × 11 image types—original + 2 × 5 compressed) with public domain Image J software[23] and FracLac plug-in,[24] implementing a differential box counting method.[10] The maximum box size was 45% of each ROI and ranged from 5 to 57 pixels, depending on the ROI size; the minimum box size was always two and the box series was linear. These parameters were independent from compression level and method. The FD of each ROI was determined as the mean of four calculations inside the ROI. For every combination of compression method and CL the mean FD was calculated. For comparison of the two compression methods, a plot depicting the relationship between the compressed file size and induced FD difference was created, as compression scales of different compression methods does not represent the same amount of information loss.[16]

### Statistical Analysis

The fractal dimensions of ROIs from the original and compressed images were compared using ANOVA ( $p < 0.05$ ). Post hoc pairwise comparisons between FDs from the original and compressed images were made with the Dunnett test ( $p < 0.05$ ).

## Results

### Image Compression

With decreasing CL from 90 to 10, the amount of image information loss increases. This results in image alteration, which is ranges from noise reduction and blurring to introduction of artifacts and finally image degradation (Fig. 1). Concurrently, a file size is reduced (Fig. 2) with smaller file sizes for J2 at all compression levels ( $p < 0.01$ ).

### Fractal Dimension at Different Compression Levels

In general, FD decreased with decreasing CL from 90 to 10 for both compression methods. A decrease in FD was more pronounced for the J2 compression method (Fig. 3) at all compression levels. There was no statistically significant difference in the FDs of the original images and images compressed with CL 90 to CL 30 for JP ( $p > 0.05$ ), while for J2, there was no statistically significant difference in the FDs for CL 90 to CL 50 ( $p > 0.05$ ) (Fig. 3). This results in a CR of 1:31 and 1:35 for JP and J2, respectively. At CL 10, the mean FD for JP and J2 was nearly the same, i.e., 2.40 and 2.39, respectively. At comparable file sizes down to 10 kB, JP induced a slightly less FD difference than J2 compression method (Fig. 4). Below this file size, an opposite relationship was found. The same FD difference of  $-0.036$  was achieved at 9.7 kB with JP 30 and 13.8 kB with J2 70 compression method (Fig. 4). For both compression methods, the standard deviation increased with the reduction of CL (Fig. 4).

## Discussion

The results of this study indicate that fractal analysis seems to be insensible to lossy image compression, namely to JPEG and JPEG 2000 at approximate compression ratio of 1:30. This result confirms the robustness of fractal analysis, as previously reported to be insensible to variations in film exposure, image geometry, and size and position of ROI. [1,20] Certainly, there is a limit in the acceptable amount of information loss, as found to be the CL 30 and CL 50 for JPEG and JPEG2000, respectively. With the use of lossy compression, high-frequency image content is lost first and as the compression level decreases, lower frequencies in image content are progressively reduced. Visually, this was represented as noise reduction at the beginning, then the image becomes progressively blurred, and finally compression artifacts become apparent, as it is clearly depicted in Fig. 3. Concurrently, the image complexity is progressively reduced resulting in the reduction of FD. Together with the loss of information, the file size reduces, which is the

primary aim of lossy compression. At the abovementioned limits of information loss for FD calculation, a considerable file size reduction was achieved, i.e., a CR of 1:31 and 1:35 for JP and J2, respectively.

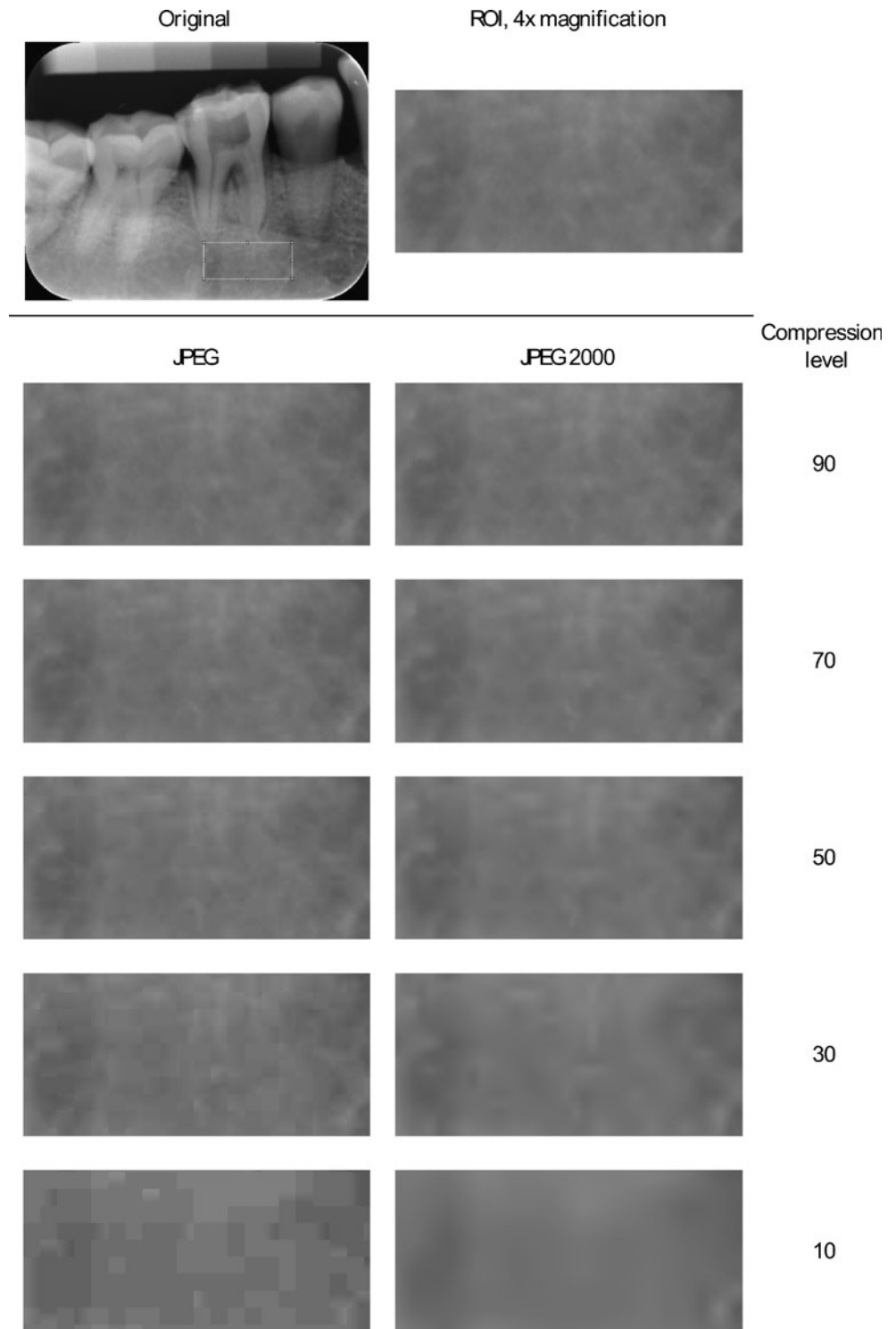
A comparison of our results with other studies is not possible as this is the first study evaluating the effect of lossy image compression on FD calculation. In general, due to the absence of normative data, fractal dimension at various conditions/pathologies and for various image types could only be evaluated as relative measurements. The limit of detection with fractal analysis method was reported only by Southard et al. It was stated that at optimal beam angulation a 5.7% decalcification of maxillary alveolar bone was the limit of detection with fractal analysis.[25] According to the results obtained, a significant difference in FD values as compared to the originals was found at CL 30 and CL 10 for JPEG2000 and only at CL 10 for JPEG. The FD of the compressed images for JPEG and JPEG2000 at CL 10 demonstrated respectively 0.10 and 0.11 lower values than the FD of their originals and therefore approximately 4.4% difference. On the other hand, the FD of images compressed with JPEG2000 at CL 30 was 0.07 lower than the original FD resulting in a 3% difference as calculated by the differential box counting method.

Originally, a box counting method for FD calculation was developed for the analysis of binary images. As radiographs are grayscale images, they should be converted to binary images before fractal analysis was performed. The process precisely described by White et al.[26] has several steps and is time consuming. To facilitate the fractal analysis in various application fields employing grayscale images, a modification of box counting method, namely differential box counting method was proposed.[27] It was proven that the differential box counting method not only has a more precise estimated value of fractal dimension, but also consumes less computational time than the so-called traditional box counting method.[28] In biomedicine, it has been used in ultrasonography for the characterization of salivary gland tumors.[29] In dental radiology, this is the first time this method has been used.

The efficient reduction of file size with lossy image compression requires applying the highest degree of information loss yet still preserving the diagnostic value of the image, resulting in the smallest possible file size. The determination of a more efficient compression method could be simply done with the comparison of compressed file sizes obtained at the same compression level. In our study, file sizes were smaller for JPEG2000 as compared to JPEG compression method at the same compression level, indicating that JPEG2000 is a more efficient compression method. It should be emphasized that this would be an erroneous approach as compression scales are different and even same compression methods do not have a standard-

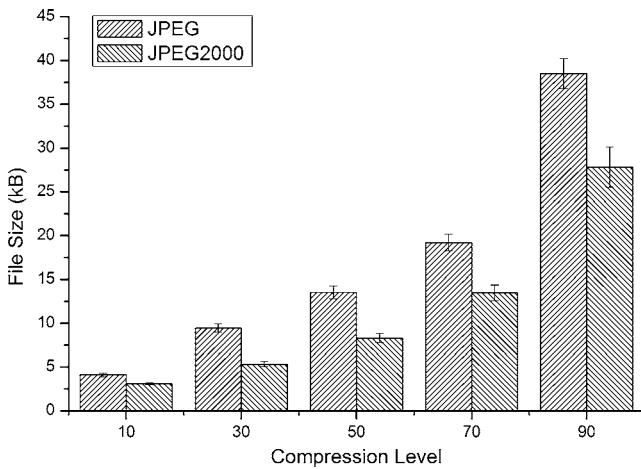


**Fig. 1** Example of original image with marked ROI and 4× magnified ROIs of original image and compressed images, which were compressed with JPEG and JPEG2000 compression method at compression level 90, 70, 50, 30, and 10



ized compression scale.[16] In this study, at compression levels above 30, the JPEG2000 compression method obviously induced more image information loss at the same compression level, resulting in smaller file size and bigger FD difference. A truly more efficient compression method would need to exhibit either the same FD difference at a smaller file size or a smaller FD difference at the same

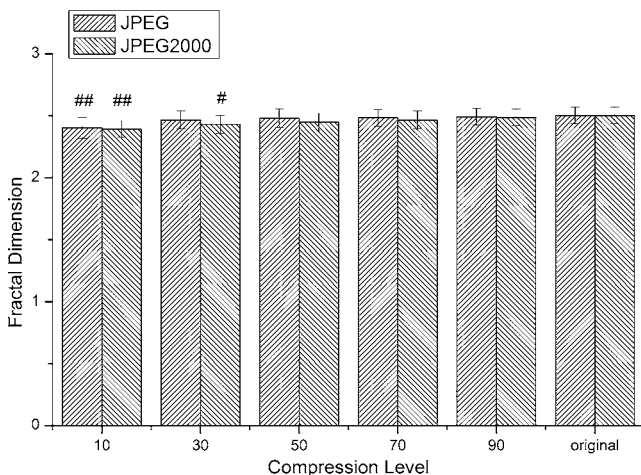
file size. For the correct comparison of the efficiency of compression methods, a plot was generated to reveal the relationship between a compressed file size and induced FD difference. This comparison demonstrated that JPEG performed slightly better than JPEG2000, i.e., induced less FD difference at the same file size, although JPEG2000 is a newer method. However, this difference would be negligi-



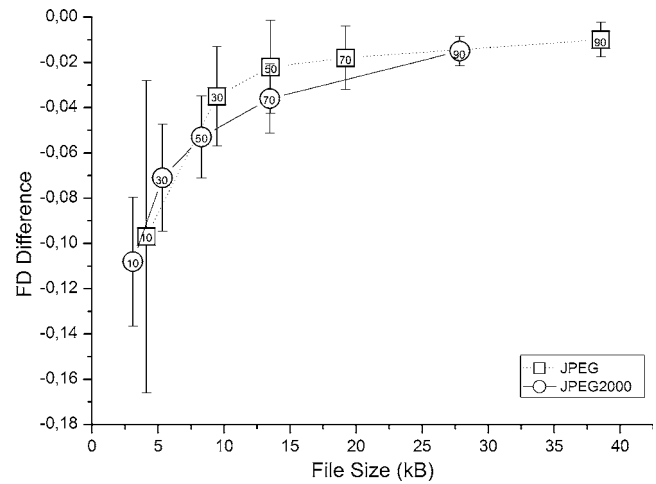
**Fig. 2** Mean file sizes at different compression levels for both compression methods

ble in the clinical setting. Similar results were reported in a study evaluating the effect of both lossy compression methods in DSR.[19]

The most common application of irreversible compression in radiology is teleradiology, while another application is to reduce the storage and bandwidth requirements required to deliver images to clinicians.[30] Teleradiology has a particular benefit from irreversible compression due to the low bandwidth connections most homes have. Although technologies like cable modems and digital subscriber lines have increased bandwidth substantially, the need for compression seems to remain particularly due to the massive amounts of data generated by cone beam computed tomography scanners. Lossy compression methods were not recommended and may not be needed for primary image storage because of the present day availability of very large sized mass storages. However, it



**Fig. 3** Mean fractal dimension at different compression levels compared to FD on original images. #,  $p < 0.05$ ; ##,  $p < 0.01$



**Fig. 4** Fractal dimension difference and file sizes at different compression levels

becomes an obligation because of the critical issue of dental/medical imaging applications to transmit and display the archived images promptly when requested.

The greatest concern in using lossy compression for dental/medical images is that subtle findings would be lost in the compressed image, which may not be always true. Subtle findings may be difficult for the human eye to discern due to the low contrast of the image, but if the image has a significant spatial extent, they are characterized by low frequencies in the spectral domain, which are well preserved by many compression methods.[31] Information belonging to subtle pathologies such as a thin fracture line or faint periapical radiolucency that may not be perceivable by the naked eye in the compressed image may be uncovered by image analysis techniques. In other words, the hidden diagnostic information in the compressed image may be revealed. At this point, the importance of testing the vulnerability of various image analysis techniques to different compression methods becomes evident. It is necessary for radiologists to be equally familiar with image compression techniques and effects of various image analyses techniques on compressed images. Such an evaluation using dental images was previously done to test the effect of JPEG and JPEG2000 compression methods on subtraction radiography.[19,32]

The lack of medicolegal standards is a significant difficulty for the widespread use of irreversible compression for diagnosis. Yet, it was stated that compression was not essentially different from any other step in the imaging chain (creation and presentation).[13] There is increasing evidence that some forms of irreversible compression can be used with no measurable degradation in diagnostic value.[13] This issue is of particular importance for clinical setting.

## Conclusions

This study confirms that FD calculation is a robust method, which can be readily performed on lossy compressed images. The JPEG compression method performed only slightly better than JPEG2000 since it showed less FD difference at the same compressed file size down to JPEG 30 CL. However, the difference between the two methods was small and it may be negligible in a clinical setting. Nevertheless, the question of the acceptable loss of information for detecting changes in bone structure using fractal analysis requires further studies, including studies on artificially generated test fractals, in which the fractal dimension may be computed analytically.

## References

- Jolley L, Majumdar S, Kapila S: Technical factors in fractal analysis of periapical radiographs. *Dento-Maxillo-Facial Radiol* 35:393–397, 2006
- Lee RL, Dacre JE, James MF: Image processing assessment of femoral osteopenia. *J Digit Imaging* 10:218–221, 1997
- Sankar D, Thomas T: A new fast fractal modeling approach for the detection of microcalcifications in mammograms. *J Digit Imaging* 23:538–546, 2010
- Chen SK, Oviir T, Lin CH, Leu LJ, Cho BH, Hollender L: Digital imaging analysis with mathematical morphology and fractal dimension for evaluation of periapical lesions following endodontic treatment. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 100:467–472, 2005
- Yu YY, et al: Fractal dimension analysis of periapical reactive bone in response to root canal treatment. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 107:283–288, 2009
- Updike SX, Nowzari H: Fractal analysis of dental radiographs to detect periodontitis-induced trabecular changes. *J Periodontol Res* 43:658–664, 2008
- Heo MS, et al: Fractal analysis of mandibular bony healing after orthognathic surgery. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 94:763–767, 2002
- Bollen AM, Taguchi A, Hujoel PP, Hollender LG: Fractal dimension on dental radiographs. *Dento-Maxillo-Facial Radiol* 30:270–275, 2001
- Ergun S, Saracoglu A, Guneri P, Ozpinar B: Application of fractal analysis in hyperparathyroidism. *Dento-Maxillo-Facial Radiol* 38:281–288, 2009
- Majumdar S, Weinstein RS, Prasad RR: Application of fractal geometry techniques to the study of trabecular bone. *Med Phys* 20:1611–1619, 1993
- Miles DA, Razzano MR: The future of digital imaging in dentistry. *Dent Clin North Am* 44:427–438, 2000
- Wenzel A, Moystad A: Work flow with digital intraoral radiography: a systematic review. *Acta Odontol Scand* 68:106–114, 2010
- Erickson BJ: Irreversible compression of medical images. *J Digit Imaging* 15:5–14, 2002
- Wallace GK: The JPEG still picture compression picture standard. *Commun Assoc Computing Machinery* 34:30–44, 1991
- ISO/IEC JTC 1/SC 29/WG 1 IIF: Information technology—JPEG 2000 image coding system: Core coding system [WG 1 N 1890], September 2000
- Fidler A, Likar B, Skaleric U: Lossy JPEG compression: easy to compress, hard to compare. *Dento-Maxillo-Facial Radiol* 35:67–73, 2006
- Siragusa M, McDonnell DJ: Indirect digital images: limit of image compression for diagnosis in endodontics. *Int Endod J* 35:991–995, 2002
- Koenig L, Parks E, Analoui M, Eckert G: The impact of image compression on diagnostic quality of digital images for detection of chemically-induced periapical lesions. *Dento-Maxillo-Facial Radiol* 33:37–43, 2004
- Fidler A, Likar B, Pernus F, Skaleric U: Comparative evaluation of JPEG and JPEG2000 compression in quantitative digital subtraction radiography. *Dento-Maxillo-Facial Radiol* 31:379–384, 2002
- Shrout MK, Potter BJ, Hildebolt CF: The effect of image variations on fractal dimension calculations. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 84:96–100, 1997
- Irfanview (2010) <http://www.irfanview.com/>. Accessed 18 Dec 2010
- Shrout MK, Roberson B, Potter BJ, Mailhot JM, Hildebolt CF: A comparison of 2 patient populations using fractal analysis. *J Periodontol* 69:9–13, 1998
- ImageJ (2010) <http://rsbweb.nih.gov/ij/>. Accessed 18 Dec 2010
- FracLac (2010) <http://rsbweb.nih.gov/ij/plugins/fraclac/fraclac.html>. Accessed 18 Dec 2010
- Southard TE, Southard KA: Detection of simulated osteoporosis in maxillae using radiographic texture analysis. *IEEE Trans Biomed Eng* 43:123–132, 1996
- White SC, Rudolph DJ: Alterations of the trabecular pattern of the jaws in patients with osteoporosis. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 88:628–635, 1999
- Sarkar N, Chaudhuri BB: An efficient differential box-counting approach to compute fractal dimension of image. *IEEE Trans Syst Man Cybern* 24:115–120, 1994
- Liu S: An improved differential box-counting approach to compute fractal dimension of gray-level image: In: Yu F, Yue G, Chen Z, Zhang J Eds. *Proceedings of the International Symposium on Information Science and Engineering*; 2008 Dec 20–22; Shanghai, China. IEEE, Los Alamitos, pp 303–306
- Chikui T, Tokumori K, Yoshiura K, Oobu K, Nakamura S, Nakamura K: Sonographic texture characterization of salivary gland tumors by fractal analyses. *Ultrasound Med Biol* 31:1297–1304, 2005
- Erickson BJ, Ryan WJ, Gehring DG, Beebe C: Image display for clinicians on medical record workstations. *J Digit Imaging* 10:38–40, 1997
- Persons K, Palisson P, Manduca A, Erickson BJ, Savcenko V: An analytical look at the effects of compression on medical images. *J Digit Imaging* 10:60–66, 1997
- Gegler A, Mahl C, Fontanella V: Reproducibility of and file format effect on digital subtraction radiography of simulated external root resorptions. *Dento-Maxillo-Facial Radiol* 35:10–13, 2006

# Preservation Plan for Microsoft - Update

---

Digital Preservation Team

---

## Document history

Version	Date	Author	Status/Change
0.1	30/04/2007	Rory McLeod	Draft
0.2	04/05/07	Rory McLeod	Reviewed and approved by DPT members Paul Wheatley and Peter Bright.
0.3	24/5/07	Paul Wheatley	Minor changes following discussions between DPT and DOM Programme Manager.
0.4	19/6/07	Paul Wheatley	Update

## DOM responsibility- Document expiry (document valid for three years)

Owner	Status	Start Date	Document Expiry Date <sup>1</sup>	Reviewed	Reviewed by (sign)
DOM Programme Manager	Checkpoint with DPT at 12 month intervals	19/06/2007	19/06/2010	2008 2009 2010	

## DPT responsibility- Preservation plan (this plan is reviewed annually)

Owner	Format	Status	Start Date	Review Date <sup>2</sup>	Reviewed by (sign)
DPT	JPEG2000	Monitor	19/06/2007	19/06/2008	
DPT	PDF 1.6	Monitor	19/06/2007	19/06/2008	
DPT	METS/ALTO	Monitor	19/06/2007	19/06/2008	

---

<sup>1</sup> At document expiry date (36 months), DOM programme requests an updated three-year document from DPT. During the life of this document, it will be reviewed annually by the DOM team.

<sup>2</sup> At preservation review date (12 months) the preservation plan is reviewed, updated and reissued to DOM by DPT. Changes are then incorporated into the DOM three year plan.

	Microsoft Update	Version 0.3	Date 24/05/07	
	Digital Preservation Team			

1	FOREWORD.....	3
2	INTRODUCTION .....	3
2.1	PURPOSE.....	3
2.2	DOCUMENT REVIEW .....	3
3	CONSTRAINTS .....	3
4	PRESERVATION PLAN TIMEFRAME AND OPERATIONAL HANDOVER..	3
5	ANALYSIS OF CONTENT .....	4
6	INGEST PROCEDURE .....	4
7	FORMAT ANALYSIS .....	4
8	PRESERVATION PLAN .....	5
8.1	CONTENT TO BE PRESERVED.....	5
8.2	FUTURE USE- ACCESS COPY .....	5
8.3	FUTURE USE- PRESERVATION COPY .....	5
8.4	FUTURE USE- METADATA .....	5
8.5	PRESERVATION ACTIVITIES.....	5
8.5.1	Preservation action .....	5
8.5.2	Preservation watch .....	5
9	FUTURE PRESERVATION SUPPORT.....	6

	Microsoft Update	Version 0.3	Date 24/05/07	
	Digital Preservation Team			

## 1 Foreword

British Library Preservation Plans are living documents that will continue to evolve over time.

Key reasons for making revisions will include:

- Changes in the content profile necessitating update or expansion
- Better content characterisation facilities enabling more detailed description and analysis of the content
- New preservation technology enabling more detailed planning for preservation actions

The role and scope of Preservation Plans will change over time, as developments in preservation metadata, particularly representation information, are progressed.

In light of these issues, a detailed and frequent review schedule has been established to ensure the published Preservation Plans remain up to date and relevant.

## 2 Introduction

This document defines the updated preservation plan for the Microsoft Live Book data ingested into the DOM system at The British Library (BL). It is updated here to reflect the vendor change from Internet Archive to CCS.

This document will refer repeatedly to the document MLB\_v2.doc that contains the detail of the project, and has already been signed off by the Microsoft Project Board. This document will serve only to update the sections that have altered under this supplier change.

### 2.1 Purpose

The purpose of this document is to approve the formats to be retained by the project, and identify the tools and methods used to preserve for the long-term the associated files for the Microsoft project. It will:

- Approve the formats based upon the previous work done
- Provide a framework for future preservation decisions for material of this type
- Provide a practical plan for long-term preservation of the data

### 2.2 Document review

This document, and the principles herein, will be reviewed yearly and re-assessed where necessary.

## 3 Constraints

Where project constraints are identified, they will be recorded here to produce a decision audit trail.

## 4 Preservation plan timeframe and operational handover

This preservation plan is due for completion in April 2007. Operational aspects will be determined by the project.

	Microsoft Update	Version 0.3	Date 24/05/07
	Digital Preservation Team		

## 5 Analysis of content

As noted above, this is a revision to reflect any changes from the original documentation.

The content remains as stated in MLB\_v2.doc, namely 19<sup>th</sup> Century British Library out of copyright printed books. These books will be scanned by the new supplier CCS.

## 6 Ingest Procedure

Ingest procedures will follow that outlined in MLB\_v2.doc. Files must be checked at an object level by JHOVE for well-formedness and validity. The following files will be output by the Microsoft Digitisation Project:

- one METS file per book
- one PDF file per book
- one ALTO file per page (containing the OCR text)
- one JPEG2000 file per page

The proposal by the DOM Programme to use WARC containers to collect large numbers of related files (e.g. the JPEG2000 and ALTO files from a single book) for storage in the DOM System is endorsed as being a suitable mechanism. Whether one or more containers is used per book is an operational decision that is outside the scope of this document. It would be inappropriate for DPT to make further recommendations in this area due to the analysis operational areas will need to undertake.

SHA-1 hashes should be provided for the images and PDF to enable the detection of data loss during transit to DOM. These hashes should be recorded in the METS metadata. An additional SHA-1 hash of the METS file should be provided as a separate file. This file will consist of 40 hexadecimal digits representing the hash, two spaces, and the filename of the METS file (this is the de facto standard mechanism for recording SHA-1 hashes in standalone files).

## 7 Format analysis

The DPT have taken the view that since the budget for hard drive storage for this project has already been allocated, it would be impractical to recommend a change in the specifics as far as file format is concerned for this project. As such, we recommend retaining the formats originally agreed in MLB\_v2.doc. These are:

Linearized PDF 1.6 files for access, with the “first page” being either the table of contents, or the first page of chapter one, depending on the specifics of the book being scanned.

JPEG 2000 files compressed to 70 dB PSNR for the preservation copy.

METS/ALTO<sup>3</sup> XML for metadata. ALTO is an extension schema to METS that describes the layout and content of text pages. ALTO will be used to encode the output of the OCR process; it will describe the text and its position of the text on

---

<sup>3</sup> METS / ALTO XML Object Model, <<http://www.ccs-gmbh.com/alto>>. I.



	Microsoft Update	Version 0.3	Date 24/05/07	
	Digital Preservation Team			

the page. MODS will be embedded into the METS document to record descriptive metadata.

The use of MODS and ALTO is new; MLB\_v2.doc did not specify either. The use of MODS is consistent with current bibliographic and preservation standards. The use of ALTO provides richer resource discovery options. These changes do not change the previous decision to accept the METS file and its content as an acceptable preservation format.

## 8 Preservation plan

### 8.1 Content to be preserved

All content received will be ingested into the archival store.

### 8.2 Future Use- Access copy

PDF 1.6 files will be retained for access as per the original project specifications. Each PDF will represent an entire book, and will be linearised. The fast load page will be either the contents page or Chapter 1 (i.e. the first "content" page).

### 8.3 Future Use- Preservation copy

JPEG2000 files will be retained for preservation as per the original project specifications.

### 8.4 Future Use- Metadata

METS/ALTO files will be created to include both logical structural data (METS) and physical layout data (ALTO) as per the standards definition.

### 8.5 Preservation activities

#### 8.5.1 Preservation action

METS/ALTO, JPEG 2000, and PDF are not considered to be at risk at the current time and no action will be taken to migrate or otherwise perform preservation actions on them. Over time, it is expected that risks will be identified by Preservation Watch activities (see below) and recorded in revisions of this document under the section "Format analysis". When a risk is determined to be sufficiently serious, this section will be expanded to define appropriate preservation actions that will ensure accessibility to the material in question.

#### 8.5.2 Preservation watch

DPT will monitor the following formats annually for risks to their accessibility:

- JPEG 2000
- PDF 1.6
- METS
- Alto

No immediate risks are identified with the formats used within this project.

-The PDF is for access purposes but is a well-defined and widely used standard.



	Microsoft Update	Version 0.3	Date 24/05/07	
	Digital Preservation Team			

-The JP2 files fulfil the role of master file but a lack of industry take-up is a slight concern from a preservation viewpoint. However, the format is well defined and documented and poses no immediate risk.

-The JP2 format has yet to be added to the BL technical standards document however, a summer workshop of industry experts has been organised at the BL in London to discuss this matter. Any relevant findings will be added to this document at this time.

-Both METS and Alto are existing metadata schemas that are approved by DPT as preservation standards.

Peter Bright of the Digital Preservation Team will conduct the monitoring of these formats annually. The DPT will perform a full review of the preservation plan and an update to any sections where concerns or changes have been identified that would affect the long-term stability of the data. This may also result in preservation actions where appropriate.

## 9 Future preservation support

It is expected that the procedures and technology for supporting preservation planning and execution within DOM will be significantly enhanced over the next few years. It is expected that facilities for characterisation will be enhanced, the systems for storing information about formats, tools that render formats and the environments tools run within will become available (possibly based on PRONOM), and tools for executing preservation actions will be made available from the Planets Project. As these developments become available, this preservation plan will be updated and expanded.



## British Newspaper Archive: Digitising the nation's memory

A look at the technology behind the British Library's project to put 300 years of newspapers online

By Sophie Curtis | [Techworld](#) | Published: 10:35 GMT, 06 February 12

Consuming content in digital form has become the norm for many of us. We watch videos on smartphones, we skim the news on tablets, we share photos on social networks and we read books on e-readers. But in a world where digital rules, where does a traditional organisation like The British Library fit in?



The first thing most people find out about The British Library is that it holds at least one copy of every book produced in the United Kingdom and the Republic of Ireland. The Library adds some three million volumes every year, occupying roughly 11 kilometres of new shelf space.

It also owns an almost complete collection of British and Irish newspapers since 1840. Housed in its own building in Colindale, North London, the collection consists of more than 660,000 bound volumes and 370,000 reels of microfilm, containing tens of millions of newspapers.

It may have come as a surprise, therefore, when The British Library – an organisation that places such high value on paper objects – announced in May 2010 that it was teaming up with online publisher brightsolid to digitise a large portion of the British Newspaper Archive and make it available via a dedicated website.

### The British Newspaper Archive

By the time The British Newspaper Archive website went live in November 2011, it offered access to up to 4 million fully searchable pages, featuring more than 200 newspaper titles from every part of the UK and Ireland. Since then, the Library has been scanning between 5,000 and 10,000 pages every day, and the digital archive now contains around 197TB of data.

The newspapers – which mainly date from the 19th century, but which include runs dating back to the first half of the 18th century – cover every aspect of local, regional and national news. The archive also offers a wealth of material for people researching family history, including family notices, announcements and obituaries.

According to Nick Townend, head of digital operations at The British Library, the idea of the project is to ensure the stability of the collection and make it available to as many people as possible.

"The library has traditionally had quite an academic research focus, but the definition of research has maybe broadened to mean everybody who's interested in doing research, and I think the library's trying to respond to that and make the collections more accessible," said Townend.

The British Library and brightsolid have set themselves a minimum target of scanning 40 million pages over ten years. "That's actually a relatively small percentage of the total collection," said Townend. The entire collection consists of 750 million pages.

"The digitisation project gives us a really good audit of the physical condition of the collection items," he added. "Some of the earlier collections were made on very thin paper and it's just naturally degraded over time, so they've effectively become 'at risk' collection items. Making a digital surrogate is part of the longer term preservation of the collection."

### Eight thousand pages a day

The fragility of some items in the collection is the reason why the scanning process has to take place on-site at Colindale, according to Malcolm Dobson, chief technology officer at brightsolid. He explained that the company set up a scanning facility there at the start of the project, with five very high-spec scanners from Zeutschel.

"We do fairly high resolution scanning – 400 DPI, 24-bit colour. The full-res image sizes vary from anything from 100MB up to 600MB per page," said Dobson. "At 400DPI these can be 12,000 pixels by 10,000 pixels – very large bitmaps. So even compressed, they are massive."

The pages are scanned in TIF format, and then converted into JPEG 2000 files. According to Dobson, JPEG 2000 provides a good quality of compression and retains a much better representation of the image than standard JPEG.



"We throw away the TIF files because they're just too big to keep," said Dobson. "To put it into perspective, we've probably got something like 250TB of JPEG 2000, and we have 3 copies of each file, so it's a lot of data. If we'd just been going with the uncompressed TIF, that would probably be something in excess of a petabyte and a half."

Once scanned, the images are transported over a Gigabit Ethernet connection to brightsolid's data centre in Dundee. The transfer happens over night, and usually takes around five to six hours.

The scanned images are entered into an OCR workflow system, where they are cropped, de-skewed, and turned into searchable text using optical character recognition. They are also "zoned" using an offshore arrangement in Cambodia. This means that areas of the page are manually catalogued by content – such as births, marriages, adverts or photographs – and referenced to coordinates.

"We end up with quite a comprehensive metadata package that accompanies the image, and it's that metadata package along with the OCR information that forms the basis of the material that's then searchable," said Townened.

### Searchable content

Having gone through this process, one copy of the file is uploaded to The British Newspaper Archive website, and another is sent to the British Library, to be ingested into its digital library system.

Dobson explained that, while JPEG 2000 is a perfect file format for storing and transferring high resolution images, most browsers are not able to render it. The images are therefore converted into a set of JPEG tiles.

"We decided to take a tiling image server, and the format we use is something called Deep Zoom, or Seadragon. There's various resolution layers stored or created, so that when someone initially goes to view an image they're looking at a sampled smaller image, and it's served up as a number of tiles," said Dobson.

"As you zoom in, only the tiles relating to that area you're looking at are delivered. You can zoom in further and further, so you get quite a good experience in terms of looking at a very high resolution image without the obvious latencies."



*Illustration of Queen Victoria - Supplement to the Bucks Herald, June 25 1887*

The website is delivered using a virtualised blade solution from IBM, consisting of a virtualised HS22 blade environment in an IBM BladeCenter H Chassis with an IBM x3755 rack server and associated SAN fabric.

According to Michael Mauchlin, IBM's systems sales manager for Scotland, the IBM blade platform is highly energy efficient, and consumes 10 percent less energy than the nearest rival platform.

"The other key attribute is that it has no single point of failure," said Mauchlin. "This is not the case with all blade designs, and system availability was a key reason why brightsolid chose IBM for this platform."

The availability aspect was of particular importance when The British Newspaper Archive website first launched in 2011, and received extensive coverage in the mainstream press. Dobson said that this prompted a surge of traffic, as people visited the site to try out the free search service.

Brightsolid was keen to provide a good experience for all these people, so it used sizing and load testing to model what it thought the peak demand was likely to be, and then create a suite of load testing scripts that represented various user journeys and activities and run that on the hardware.

"So we were able to prove that we could deliver the kind of numbers of searches per second, the number of images and tiles associated with those things per second to meet that peak demand," said Dobson. "All the evidence from our customers on that day was that the experience was good."

## Digital library system

Meanwhile, the copy that is sent to the British Library enters into its digital library system.

"We have a four node digital library system, where we have a base in Yorkshire, one in London, one in Aberystwith and one in Edinburgh, and effectively we replicate content across all of the nodes," explained Townend. "Each of the nodes has a slightly different technical architecture and hardware setup, so that even if one node were to develop a fault, technically it shouldn't be the same fault replicated across all four sites."

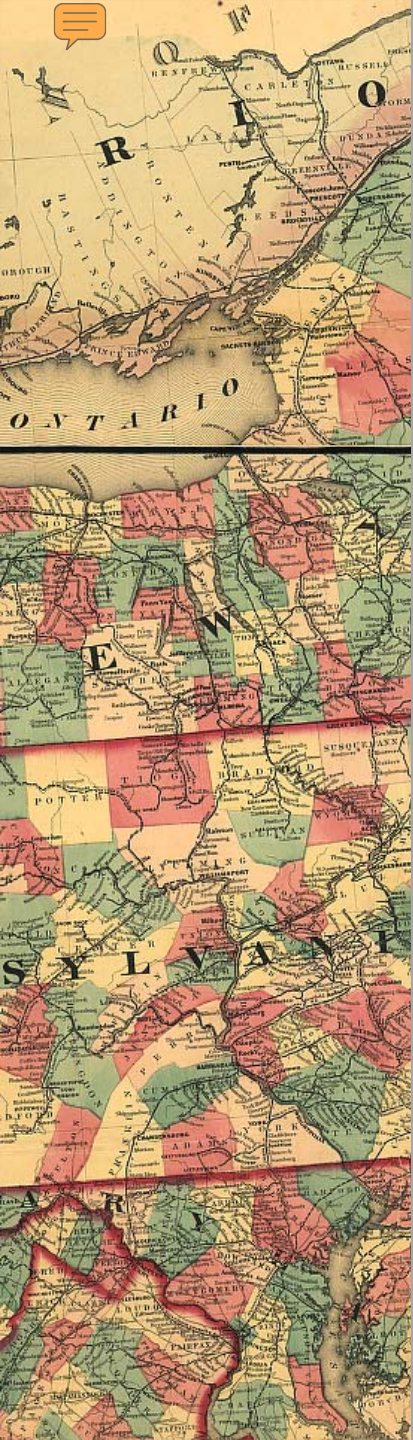
The British Library's ambitions don't stop there, however. Townend explained that the organisation is always looking to make more content available in more interesting ways. For example, it has massive ambitions in terms of "born digital" material.

"The government is currently in the process of reviewing new legislation that will allow us to collect the digital items under law, but that's not in place at the moment," explained Townend. "We're currently working on a voluntary deposit base. The legislation will allow us to capture the UK web domain, in terms of every .co.uk website, which is a fairly frightening prospect."

It also hopes to enter into a licensing agreement with copyright owners, so that more up-to-date newspaper content can be published on the site and accessed in digital format. However, this could be a long process, according to Townend. "This is a huge challenge in terms of copyright and content management," he said.

<http://features.techworld.com/applications/3333256/british-newspaper-archive-digitising-the-nations-memory/>





# The Benefits of JPEG 2000 for Image Access and Preservation

Robert Buckley  
University of Rochester /  
NewMarket Imaging

JPEG 2000 Summit  
Washington, DC  
May 12, 2011





# What's changing

2

- Image files are getting bigger and there are more of them
- Demand for access to online image collections
- Several mass digitization projects underway
- The economics and sustainability of preservation

# Why JPEG 2000?

3

- Cost
- User Experience
- New Opportunities

# JPEG 2000 Features

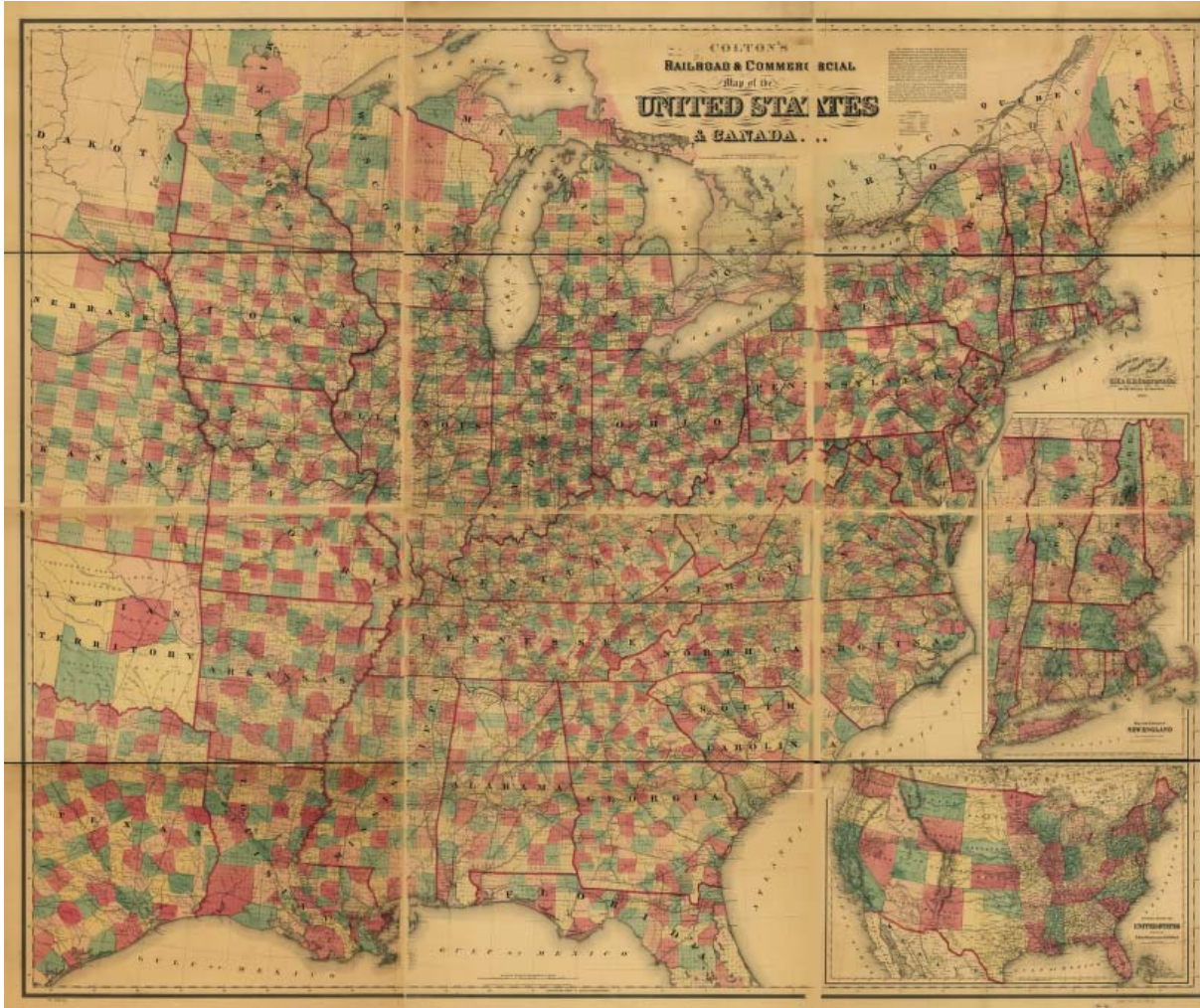
4

- ❑ Open Standard without license fees or royalties in Part 1
- ❑ A single approach to lossless and lossy compression
- ❑ Progressive display with scalable resolution and quality
- ❑ Region-of-Interest (ROI) on coding and access
- ❑ Create compressed image with specified size or quality
- ❑ Support for domain-specific metadata in file format
- ❑ Low loss across multiple decompress-compress cycles
- ❑ Error resilience



# On-Demand / Just-in-Time Imaging

5



Give me  
**x,y** coordinates  
of **c** component  
at **z** resolution  
and **p** quality

# Error Resilience

6



TIFF



JP2



# Benefits of JPEG 2000

7

- Reduced costs for storage and maintenance
  - ▣ Smaller file size compared to uncompressed TIFF
  - ▣ One master replaces multiple derivatives
- Enhanced handling of large images
  - ▣ Fast access to image subsets
  - ▣ Intelligent metadata support
- Enables new opportunities



# Cost savings with JPEG 2000

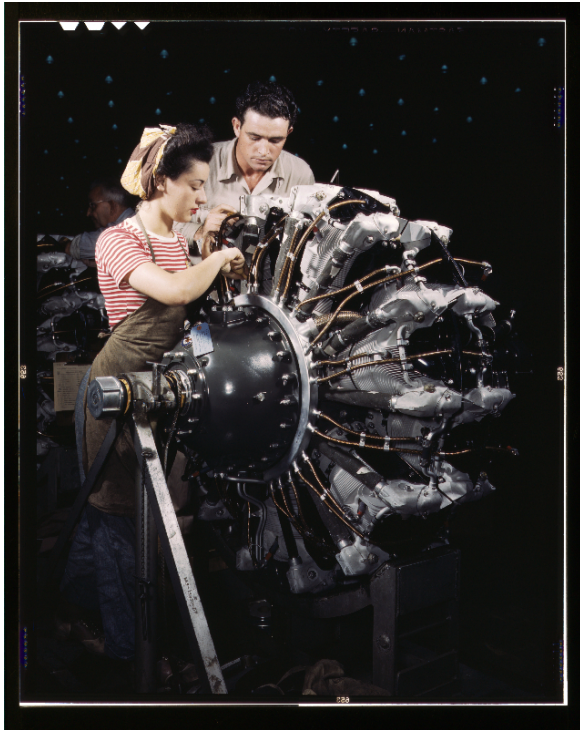
8

- Produces smaller files
  - Storage costs \$2000-\$4000/TB/year
  - Lossless 2:1 compression of 100 TB of uncompressed TIFF could save \$100,000 to \$200,000 per year
  - Lossy compression would save even more
- Eliminates need for multiple derivatives
  - Smart decoding with just-in-time imaging
  - Fewer files to create, manage and keep track of
- Does not require license fees or royalties



# User Experience

9





# JPEG 2000 Profiles

10

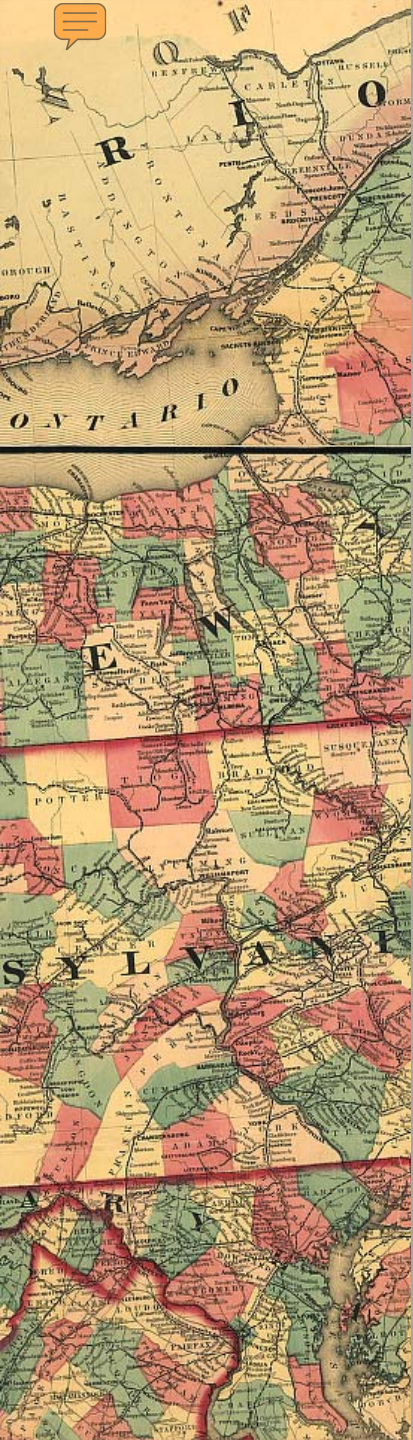
- Realizing the benefits of JPEG 2000
  - ▣ Matching JPEG 2000 capabilities with your application
- Managing JPEG 2000 use
  - ▣ Defined set of JPEG 2000 parameters
  - ▣ Establish reference point
- Performance tuning
  - ▣ Meet requirements
  - ▣ Allow optimizations



# Looking ahead

11

- JPEG 2000 = Compression + Services
  - ▣ Lossy compression: access
  - ▣ Lossless → Lossy compression: preservation
- What questions remain?
  - ▣ Sustainability, color support, implementation status
- Economics of preservation



# The Benefits of JPEG 2000 for Image Access and Preservation

Robert Buckley  
robert.buckley@rochester.edu  
rob@NewMarketImaging.com

JPEG 2000 Summit  
Washington, DC  
May 12, 2011



# JPEG2000

## Specifications for The National Library of the Czech Republic

**Lecture:** Wellcome Library, NOV 16, 2010  
**Lecturer:** Bedřich Vychodil  
**Contact:** [bedrich@gmail.com](mailto:bedrich@gmail.com)  
[bedrich.vychodil@nkp.cz](mailto:bedrich.vychodil@nkp.cz)



# Compression Ratio TEST

Format	BMP	TIFF	TIFF LZW	PNG (6)	JPEG (12)	JPEG (11)	JP2 (0)	JP2 (1:1)	JP2 (1:10)	JP2 (1:25)	JPM photo	JPM standard/good	JPM standard/low	DJV photo MAX	DJV photo preset	DJV manuscript	
Comparison (MB)	I. 8bit (B/W)	8,70	8,71	3,81	2,70	2,50	1,69	2,44	3,00	0,87	0,35	0,31	0,16	0,13	1,62	0,46	0,12
	I. 24bit (color)	26,01	26,13*	9,80	6,43	4,13	2,74	5,57	5,19	2,61	1,05	0,92	0,30	0,21	2,39	0,57	0,15
	II. 8bit (B/W)	8,70	8,71	0,61	0,48	2,19	1,74	1,61	2,34	0,87	0,35	0,31	0,12	0,11	1,94	0,87	0,02
	II. 24bit (color)	26,01	26,13*	0,96	0,44	2,56	2,05	1,59	2,38	2,38	1,04	0,92	0,19	0,17	1,94	0,87	0,02
Info	windows	interleaved	interleaved	none	baseline	baseline	quality-axis	quality-axis	quality-axis	quality-axis	no licence (watermark)	no licence (watermark)	no licence (watermark)				
Compression	none	none	LZW (lossless)	Adaptive DEFLATE (3 to 7)	DCT 12 (max)	DCT 11	RCT	ICT 1/1	ICT 1/10	ICT 1/25	ICT/JPM Profile: Photo, save thumbnail, quality Good	ICT/JPM Profile: Standard, save thumbnail, quality Good	ICT/JPM Profile: Standard, save thumbnail, quality low	IW44 photo (max setting) BGsubsample6, BG100	IW44/JBIG2 photo (preset setting) BG sample3, FG sample12, FG/BG75, GB quality100, Text lossless	IW44/JBIG2 manuscript (preset setting) BG sample3, FG sample12, FG/BG60, GB quality75, Text medium-loss	
SW	Photoshop	Photoshop	Photoshop	Photoshop	Photoshop	Photoshop	LEAD plugin Photoshop	LEAD plugin Photoshop	LEAD plugin Photoshop	LEAD plugin Photoshop	IrfanView Plugins 425 (Lura Document)	IrfanView Plugins 425 (Lura Document)	IrfanView Plugins 425 (Lura Document)	Enterprise edition, LizardTech	Enterprise edition, LizardTech	Enterprise edition, LizardTech	

# Parameters chart

MC	PMC	PMC
Books, periodicals, maps, manuscripts (masters)	Books, periodicals (production masters)	Maps, manuscripts (production masters)
Kakadu	Kakadu	Kakadu
Part 1 (.jp2)	Part 1 (.jp2)	Part 1 (.jp2)
Lossless	Lossy	Lossy
2:1 to 3:1	20:1 to 30:1 it is high compress ratio, it may be changed according our needs	8:1 to 10:1
4096x4096	1024x1024	1024x1024
RPCL	RPCL	RPCL
5 or 6	5	5 or 6 /6 layers for over-sized material/
1	12 /logarithmic/	12 /logarithmic/
6	6	6
5-3 reversible filter	9-7 irreversible filter	9-7 irreversible filter
256x256 for first two decomp. levels, 128 by 128 for lower levels	256x256 for first two decomp. levels, 128 by 128 for lower levels	256x256 for first two decomp. levels, 128 by 128 for lower levels
No	No	No
64x64	64x64	64x64
Yes "R"	Yes "R"	Yes "R"
YES	YES	YES

# Kakadu Command-lines



## MC (Master Copy):

```
kdu_compress -i example.tif -o example.jp2 "Cblk={64,64}" Corder=RPCL "Stiles={4096,4096}"  
"Cprecincts={256,256},{128,128}" ORGtparts=R Creversible=yes Clayers=1 Clevels=5 "Cmodes={BYPASS}"
```

## PMC (Production Master Copy):

### Compress Ratio 1:8

```
kdu_compress -i example.tif -o example.jp2 "Cblk={64,64}" Corder=RPCL "Stiles={1024,1024}"  
"Cprecincts={256,256},{128,128}" ORGtparts=R -rate 3 Clayers=12 Clevels=5 "Cmodes={BYPASS}"
```

### Compress Ratio 1:10

```
kdu_compress -i example.tif -o example.jp2 "Cblk={64,64}" Corder=RPCL "Stiles={1024,1024}"  
"Cprecincts={256,256},{128,128}" ORGtparts=R -rate 2.4 Clayers=12 Clevels=5 "Cmodes={BYPASS}"
```

### Compress Ratio 1:20

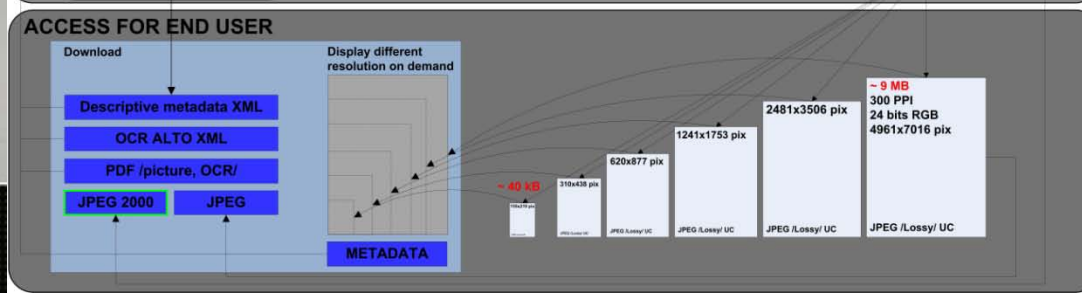
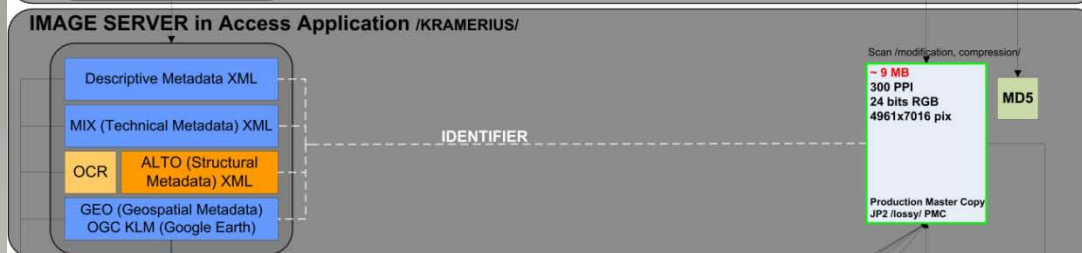
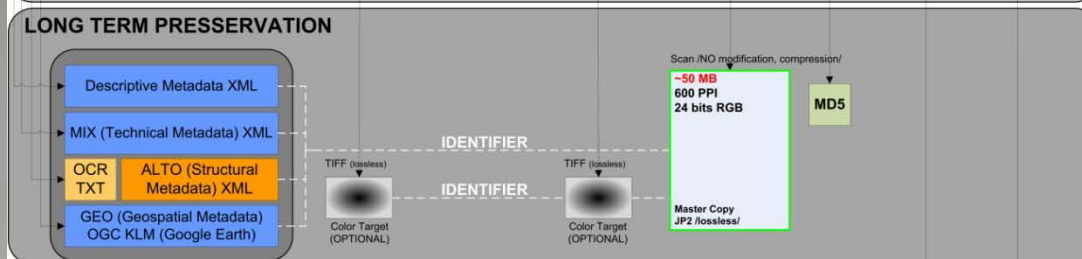
```
kdu_compress -i example.tif -o example.jp2 "Cblk={64,64}" Corder=RPCL "Stiles={1024,1024}"  
"Cprecincts={256,256},{128,128}" ORGtparts=R -rate 1.2 Clayers=12 Clevels=5 "Cmodes={BYPASS}"
```

### Compress Ratio 1:30

```
kdu_compress -i example.tif -o example.jp2 "Cblk={64,64}" Corder=RPCL "Stiles={1024,1024}"  
"Cprecincts={256,256},{128,128}" ORGtparts=R -rate 0.8 Clayers=12 Clevels=5 "Cmodes={BYPASS}"
```



# Workf



# NEW DIGITIZATION

OUTSOURCING /STORAGE, DIGITIZATION/

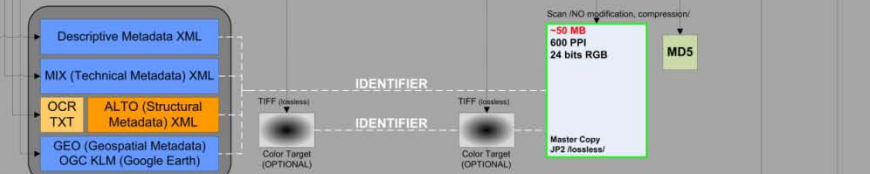
## DIGITIZATION



## PROCESSING



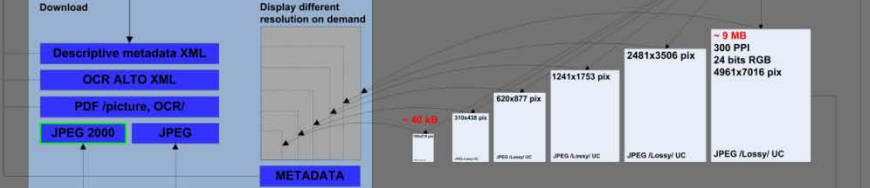
## LONG TERM PRESERVATION



## IMAGE SERVER in Access Application /KRAMERIUS/



## ACCESS FOR END USER

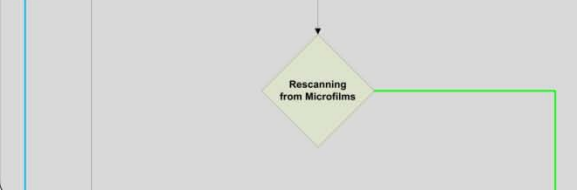


# DATA MIGRATION FROM K3 TO K4

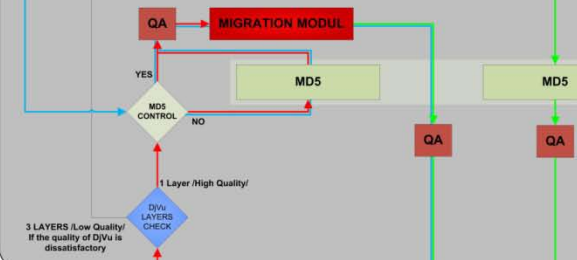
EXTERNAL STORAGE /UNSECURED/



## DIGITIZATION



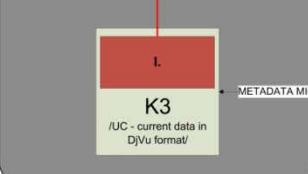
## PROCESSING



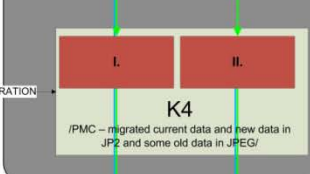
## LONG TERM PRESERVATION



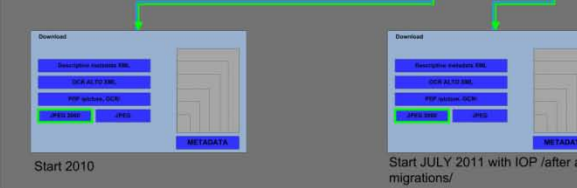
## ACCESS APL. /KRAMERIUS/



## ACCESS APL. /KRAMERIUS-Image server/



## ACCESS FOR END USER





# Tests

123 MB TIFF /No compression/, MB JP2 /No compression/, RGB, 24bitů, 300 PPI  
PhotoShop CS5



IrfanView 4.27

11,5 MB JP2, RGB, 24bitů, 1:8



4,6 MB JP2, RGB, 24bitů, 1:20

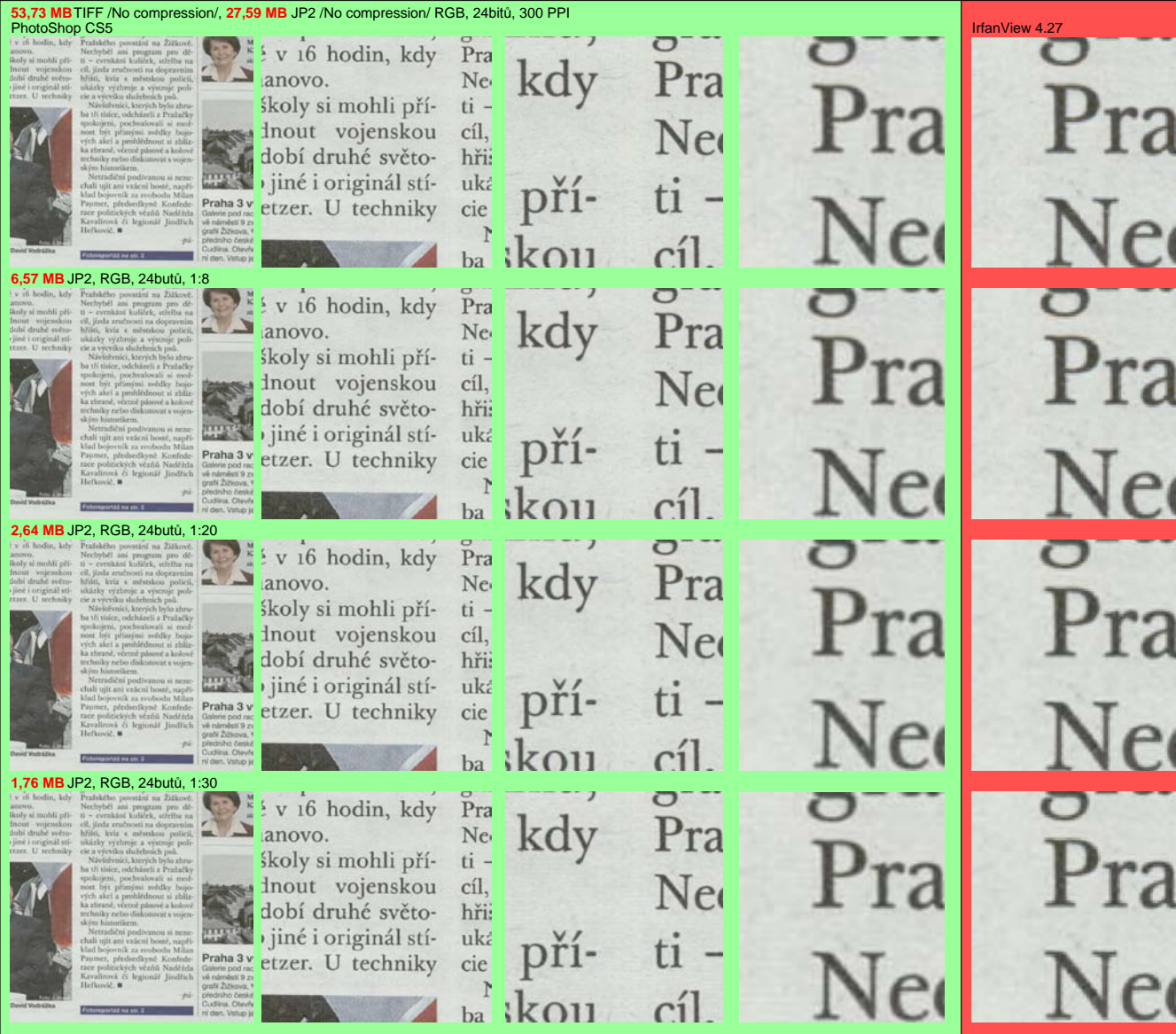


3,0 MB JP2, RGB, 24bitů, 1:30



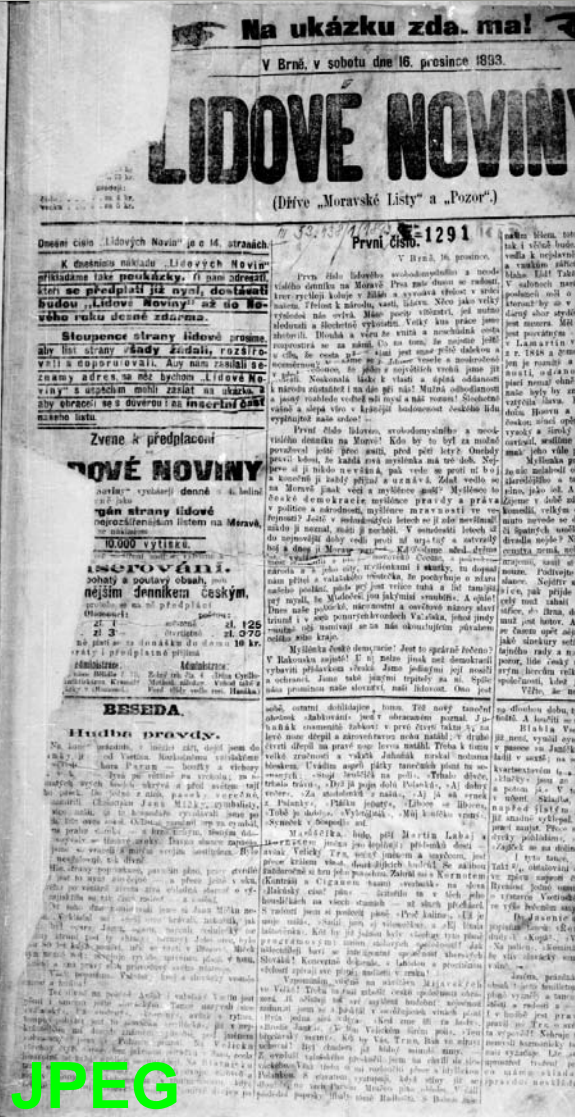


# Tests





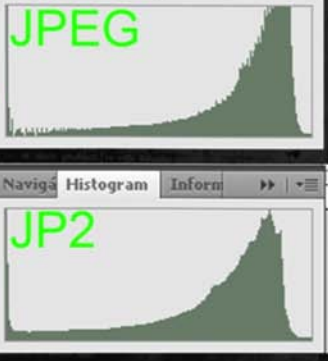
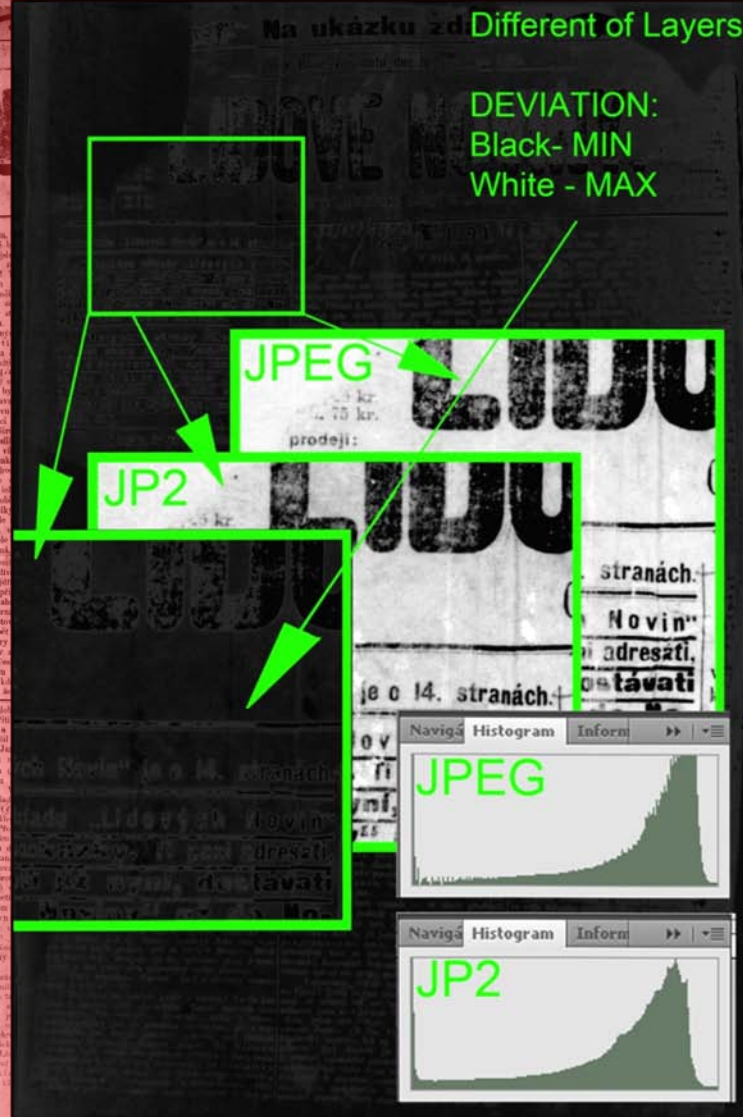
# Migration from JPEG to JP2



JPEG



JPEG2000 / P2/



Different of Layers

DEVIATION:  
Black- MIN  
White - MAX

JPEG

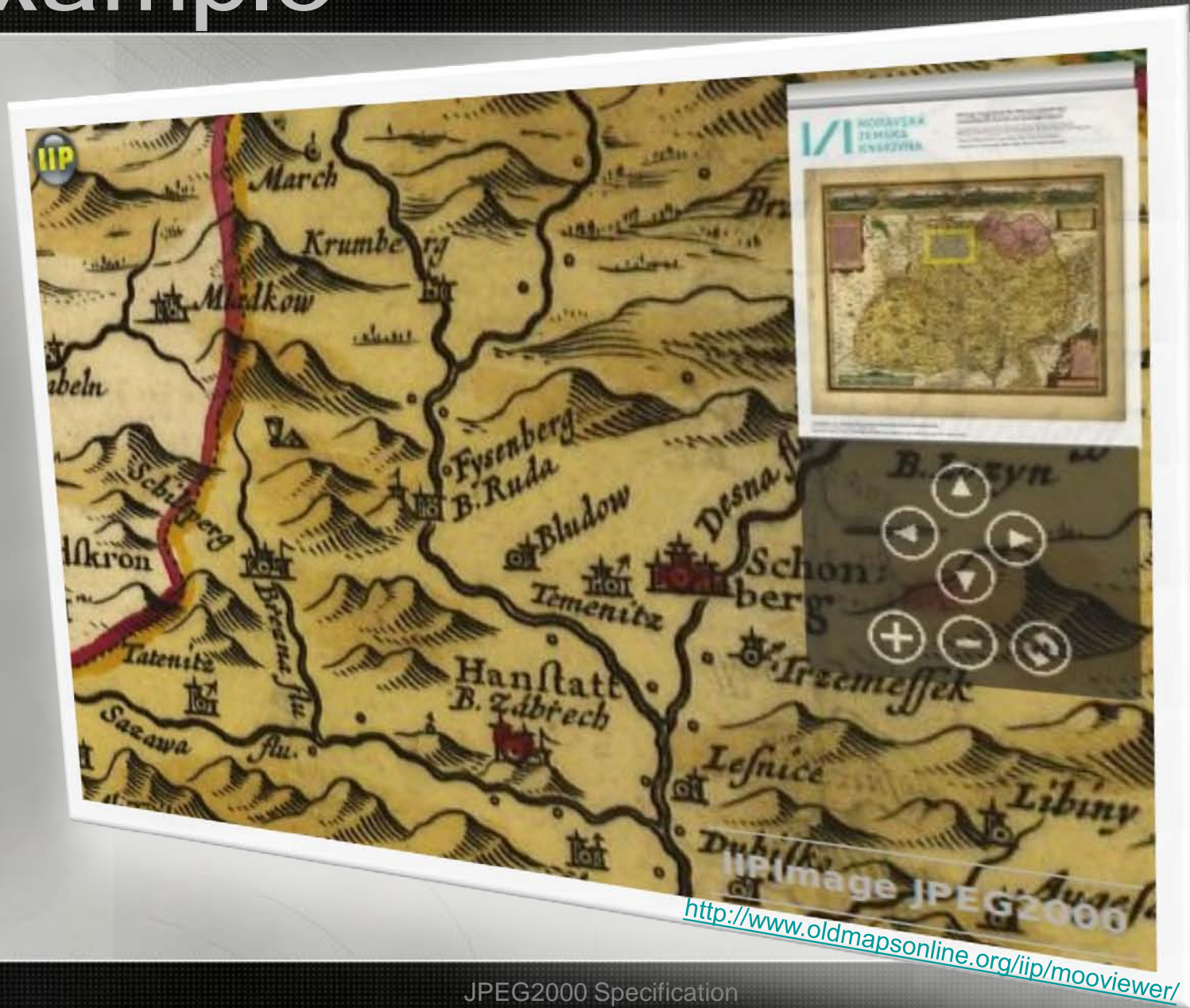
JP2

JPEG

JP2



# Example



# The End...

## Questions...?

Lecture: Wellcome Library, NOV 16, 2010  
Lecturer: Bedřich Vychodil  
Contact: [bedrich@gmail.com](mailto:bedrich@gmail.com)  
[bedrich.vychodil@nkp.cz](mailto:bedrich.vychodil@nkp.cz)

RE TIFF vs. JPEG 2000

From: Kibble, Matt [Matt.Kibble@proquest.co.uk]  
Sent: Wednesday, February 29, 2012 11:46 AM  
To: Buchner, Mary  
Subject: RE: TIFF vs. JPEG 2000

Dear Mary,

You're welcome! Please feel free to cite this email, and I have no objection to my name being included. But just to clarify, JP2 is our preferred archival standard for Early European Books, not for EEBO (Early English Books Online), where we use TIFFs partly for legacy reasons, and because the file sizes for these black and white images are not such a problem.

Best wishes,  
Matt

> -----Original Message-----

> From: Buchner, Mary [mailto:mbuc@loc.gov]  
> Sent: 29 February 2012 16:31  
> To: Kibble, Matt  
> Subject: RE: TIFF vs. JPEG 2000

>  
> Dear Matt,

>  
> Thank you very much for your reply! Your email is very helpful for my  
> research.

>  
> I'm currently updating a list of institutions and their decisions  
> regarding JPEG 2000 as an archival format, and this may effect the  
> federal guidelines as made by FADGI (federal agencies digitization  
> guidelines initiative). You can find the website at  
> www.digitizationguidelines.gov/  
>

> Would you mind if I cited this email as a part of my list? I can keep  
> your name anonymous if you prefer, but the fact that you prefer to use  
> JP2 for EEBO's archival standard is an interesting fact that I would  
> like to note.

>  
> Thank you very much,

>  
> Mary Buchner  
> Intern, Office of Strategic Initiatives U.S. Library of Congress  
> mbuc@loc.gov  
>

> -----Original Message-----

> From: Kibble, Matt [mailto:Matt.Kibble@proquest.co.uk]  
> Sent: Wednesday, February 29, 2012 11:21 AM  
> To: Buchner, Mary; INTL-eebo-webmaster  
> Subject: RE: TIFF vs. JPEG 2000

>  
> Dear Mary,

>  
> Thank you for your mail. This is an interesting question, and is one  
> which has come up for us recently for other projects.

>  
> For EEBO, the decision to use TIFFs was taken in the late 1990s, and  
> so predates the advent of the JP2 format. Until recently, this was the  
> standard preservation format which we used for images for a number of  
> our databases. For EEBO, we create the TIFFs ourselves by scanning  
> from microfilm. These are converted into gifs for web delivery and  
> display, and we offer users the option to download the full resolution  
> TIFF files.

RE TIFF vs. JPEG 2000

>  
> We have recently begun a different project, Early European Books,  
> where we are scanning rare books in full colour on site in partner  
> libraries throughout Europe. For this project, we capture either TIFFs  
> or JP2 files, depending on the preference of the library (since we are  
> also providing digital files back to the library). We then convert  
> these into jpegs for web delivery, and offer users the option of  
> downloading a PDF which is driven by the JPEG files. We don't deliver  
> the full-resolution TIFFs to web users because of the huge file sizes  
> involved for these rich colour images. For the archived files, our  
> preference is to use JP2 files because they require much less storage  
> space with very little loss of visual information, but we still create  
> TIFF files for those institutions who require that format as their archival  
standard.

>  
> I hope that information is helpful - please let me know if you have  
> any further queries.

>  
> Best wishes,  
> Matt

>  
>  
> Matt Kibble  
> Senior Product Manager, Arts and Humanities ProQuest | Cambridge, UK  
> matt.kibble@proquest.co.uk www.proquest.co.uk

>  
> This email and any files transmitted with it are confidential and  
> intended solely for the use of the individual or entity to whom they  
> are addressed. If you have received this message in error, please  
> notify the sender immediately by reply and immediately delete this  
> message and any attachments.

>  
>  
>  
> > -----Original Message-----  
> > From: mbuc@loc.gov [mailto:mbuc@loc.gov]  
> > Sent: 28 February 2012 15:46  
> > To: INTL-eebo-webmaster  
> > Subject: TIFF vs. JPEG 2000

> >  
> > WEBMASTER QUERY : Early English Books Online (  
> > <http://eebo.chadwyck.com/> )

> >  
> > Name: Mary Buchner  
> > Institution: United States Library of Congress  
> > Email: mbuc@loc.gov  
> > Status: researcher  
> > First Time: No  
> > Subject: TIFF vs. JPEG 2000

> >  
> > Message: Hello!

> >  
> > I'm currently researching various digitization efforts and their  
> > decision to use .TIFF or .jpg2 (jpeg 2000). I've used EEBO  
> > previously for other researched and noticed that you only offer  
> > files for download with the .TIFF extension. Do you have any  
> > documentation regarding this decision? What institutions provide  
> > the digitized content for your database?

> >  
> > Thank you for your help.

> >



RE TIFF vs. JPEG 2000

> > -Mary Buchner  
> >  
> > Cc:  
> >  
> > IP: 140.147.236.195  
> > UID: libcong  
> > SUB OPTIONS: Z39.50 access: NO,  
> > Early English Books Online: Basic Package: YES,  
> > Activate Library Branding in EEBO: NO,  
> > Access to Full Text in EEBO (for TCP members only): NO,  
> > Access to works in all collections: YES,  
> > Access to works in the STC1 collection: NO,  
> > Access to works in the STC2 collection: NO,  
> > Access to works in the Thomason collection: NO,  
> > Access to works in the Biblical Studies collection: NO,  
> > Donor Branding: NO,  
> > Access to Lion Author Pages (on US server): NO,  
> > Access to Lion Author Pages (on UK server): NO,  
> > Access to Variant Spelling functionality (CIC  
> > consortium  
> > only): NO,  
> > Enable ECCO cross-searching (Gale ECCO subscribers  
> only):  
> > NO,  
> > Access to additional TCP2 Full Text in EEBO: NO  
> > BROWSER: Mozilla/5.0 (Windows NT 5.1; rv:2.0.1) Gecko/20100101  
> > Firefox/4.0.1  
>  
>

# A survey on JPEG2000 encryption

Dominik Engel · Thomas Stütz · Andreas Uhl

Received: 11 June 2008 / Accepted: 16 December 2008 / Published online: 16 January 2009  
© Springer-Verlag 2009

**Abstract** Image and video encryption has become a widely discussed topic; especially for the fully featured JPEG2000 compression standard numerous approaches have been proposed. A comprehensive survey of state-of-the-art JPEG2000 encryption is given. JPEG2000 encryption schemes are assessed in terms of security, runtime and compression performance and their suitability for a wide range of application scenarios.

## 1 Introduction

A clear trend toward the increased employment of JPEG2000 for specialized applications has been observable recently, especially where a high degree of quality or scalability is desired. For example, the Digital Cinema Initiative (DCI), an entity created by seven major motion picture studios, has adopted JPEG2000 as the (video!) compression standard in their specification for a unified Digital Cinema System [8]. With increasing usage comes the increasing need for practical security methods for JPEG2000. Over the last years, a significant number of different encryption schemes for visual data types have been proposed (see [23, 60] for extensive overviews). Recently, the awareness for JPEG2000 secu-

rity has grown with the finalization of part 8 of the JPEG2000 standard, JPSEC [32].

The most secure method for the encryption of visual data, sometimes referred to as the naive method, is to encrypt the whole multimedia stream (e.g., a JPEG2000 code-stream) with the aid of a cryptographically strong cipher like AES [7]. The most prominent reasons not to stick to classical full encryption of this type for multimedia applications are

- to maintain format compliance and/or associated functionalities like scalability (which is usually achieved by parsing operations and marker avoidance strategies),
- to achieve higher robustness against channel and storage errors, and
- to reduce the computational effort (which is usually achieved by trading off security, as is the case in partial or soft encryption schemes).

These issues immediately make clear that encryption methods for visual data types need to be specifically tailored to fulfill the requirements of a particular multimedia application with respect to security on the one hand and other functionalities on the other hand.

A number of proposals for JPEG2000 encryption have been put forward to date. The approaches differ significantly in their fields of applications, their levels of security, the functionalities they provide and their computational demands.

In this paper our aim is to give a comprehensive survey of the existing approaches. For this purpose, we first present different categories for the classification of JPEG2000 encryption schemes. We systematically describe, discuss, evaluate, and compare the various techniques, especially with respect to their impact on JPEG2000 compression performance,

---

Communicated by A.U. Mauthe.

D. Engel · T. Stütz · A. Uhl (✉)  
Department of Computer Sciences, Salzburg University,  
Jakob-Haringerstr. 2, Salzburg, Austria  
e-mail: uhl@cosy.sbg.ac.at

D. Engel  
e-mail: dengel@cosy.sbg.ac.at

T. Stütz  
e-mail: tstuetz@cosy.sbg.ac.at

concerning their security, and regarding their computational performance.

In Sect. 2 we give an introduction to media encryption. Section 3 provides an overview of the JPEG2000 standard suite, focusing on the parts relevant to our survey, most importantly Part 8, JPSEC. In Sect. 4 we discuss evaluation criteria for JPEG2000 encryption schemes.

In Sect. 5 we cover methods for bitstream-oriented encryption techniques, Sect. 6 is devoted to compression-integrated methods. In Sect. 7 we discuss the findings of this survey and we give recommendations which techniques should preferably be used in specific application scenarios. Section 8 concludes the paper.

## 2 Media encryption

In the following, we discuss a number of useful categories for the classification of media encryption schemes, which of course are also relevant for JPEG2000 encryption.

### 2.1 Security and quality constraints

Encryption may have an entirely different aim as opposed to maximal confidentiality or privacy in the context of certain multimedia applications. “Transparent encryption” [41] has been introduced mainly in the context of digital TV broadcasting (also called “perceptual encryption” predominantly in the area of audio encryption): a pay TV broadcaster does not always intend to prevent unauthorized viewers from receiving and watching his program, but rather intends to promote a contract with non-paying watchers. This can be facilitated by providing a low quality version of the broadcast program for everyone, only legitimate (paying) users get access to the full quality visual data (which has been already broadcast together with the low quality version in encrypted form). Also, the degree of confidentiality varies from application to application. Whereas a high degree is required for applications like video conferencing, telemedicine, or surveillance, in some scenarios it might be sufficient for digital rights management schemes to degrading the visual quality to an extent where a pleasant viewing experience is no longer possible (“sufficient encryption”). Only transparent encryption guarantees a minimum quality of the preview image (the encrypted image transparently decoded).

We can summarize the following distinct application scenarios and requirements as follows:

- **Highest Level Security/Cryptographic Security**  
Applications that require a very high level of security, no information about the plaintext (image and compressed file) shall be deducible from the ciphertext.
- **Content Security/Confidentiality**  
Information of the plaintext may leak, but the image content must not be discernible.
- **Sufficient encryption/Commercial application of encryption**  
The content must not be consumable due to the high distortion (DRM systems).
- **Transparent/Perceptual encryption**  
A preview image has to be decodable, but the high quality version has to be hidden. Another application is privacy protection.

### 2.2 Selective/partial and lightweight encryption

In order to serve the purpose of reducing computational effort in the encryption process, more efficient methods as opposed to full encryption with cryptographically strong ciphers have been designed. Such systems—often denoted as “selective/partial” or “soft” encryption systems—usually trade off security for runtime performance, and are therefore—in terms of security—somewhat weaker than the naive method. Whereas selective or partial encryption approaches restrict the encryption process (employing classical ciphers like AES) to certain parts of the visual data by exploiting application-specific data structures or by encrypting only perceptually relevant information (e.g., encryption of I-macroblocks in MPEG, packet data of leading layers in JPEG2000), the soft encryption approach employs weaker encryption systems (like permutations) to accelerate the processing speed. Often, selective/partial encryption or soft encryption are termed “lightweight encryption”.

### 2.3 Bitstream-oriented versus compression-integrated encryption

Bitstream-oriented techniques only operate on the final compressed stream, i.e., the codestream. Although they may parse the codestream and for example use meta-information from the codestream, they do not access the encoding (or decoding) pipeline. Classical methods for encryption fall into this category, and also many selective/partial encryption schemes that only encrypt parts of the codestream.

Compression-integrated techniques apply encryption as part of the compression step, sometimes going so far that part of the compression actually is the encryption. One possibility is to apply classical encryption after the transform step (which in most cases inevitably destroys compression performance). For other approaches the transform step is also the encryption step at the same time. Another possibility to achieve compression-integrated encryption is by selecting the transform domain to be used for encoding based on a key.

## 2.4 On-line/off-line scenario

Two application scenarios exist for the employment of encryption technology in multimedia environments [46] if we distinguish whether the data is given as plain image data (i.e., not compressed) or in form of a codestream resulting from prior compression. In applications where the data is acquired before being further processed, the plain image data may be accessed directly for encryption after being captured by a digitizer. We denote such applications as “on-line”. Examples for this scenario are video conferencing and on-line surveillance. On the other hand, as soon as visual data has been stored or transmitted once, it has usually been compressed in some way. Applications where codestreams are handled or encrypted are denoted “off-line”. Examples are video on demand and retrieval of medical images from a database.

Note that while this distinction is related to the distinction between bitstream-oriented and compression-integrated encryption, it is a distinction by application, not by procedure. In principle, both bitstream-oriented and compression-integrated methods may be suited for either of the two scenarios. However, the application of compression-integrated methods in an off-line scenario will in general not be very efficient, for obvious reasons.

## 2.5 Format-compliance

The aim of format-compliant encryption is to preserve—carefully—selected parts of the (meta-)information in the codestream so that the encrypted data is compliant to the format of the unencrypted data. If format compliance is desired, the classical cryptographic approach (the naive method) cannot be employed as no (meta-)information is preserved. In many cases, header information is left in plaintext and the actual visual information is encrypted avoiding the emulation of marker and header sequences in the ciphertext parts. In this manner, the properties of the original codestream carry over to the encrypted stream. For example, rate adaptation may be done in the encrypted domain easily, provided the original codestream facilitates this functionality as well (which is true for scalable or embedded codestreams, for example). While the headers are not encrypted in most approaches proposed to date, they may be encrypted in a format-compliant way as well.

The requirement of format compliance can safely be assumed to be of great importance. Format-compliance enables the transparent application of encryption, leading to numerous benefits such as signal processing in the encrypted domain, rate adaptation, or reduction of deployment costs.

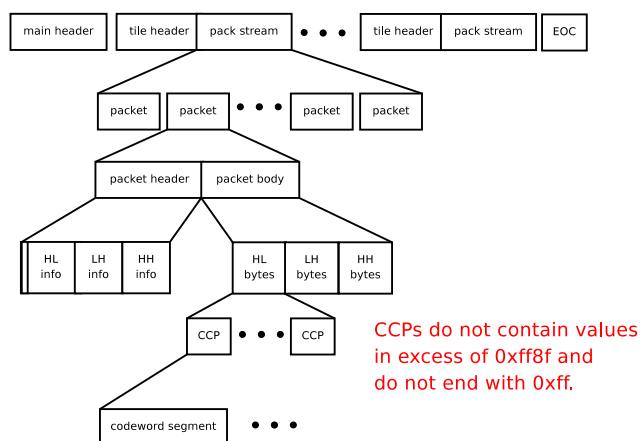
## 3 The JPEG2000 standard suite

JPEG2000 has 13 parts (part 7 has been abandoned). For the focus of this survey our interest is in Part 1 (the core coding system), Part 2 (extensions), Part 4 (conformance testing) and Part 8 (JPSEC).

### 3.1 Part 1: the core coding system

JPEG2000 [56] employs a wavelet transform; Part 1 of the standard specifies an irreversible 9/7 and a reversible integer 5/3 wavelet transform and requires the application of classical pyramidal wavelet decomposition. The components of the image (after an optional multi-component transform) are subdivided into tiles, each of these tiles is independently wavelet-transformed. For a detailed description of the data partitioning refer to [56, p. 449] or to [34, p. 42]. After the wavelet transform the coefficients are quantized and encoded using the EBCOT scheme, which renders quality scalability possible. Thereby the coefficients are grouped into codeblocks and these are encoded bitplane by bitplane, each with three coding passes (except the first bitplane). The coding passes may contribute to a certain quality layer. A packet body contains CCPs (codeblock contribution to packet) of codeblocks of a certain resolution, quality layer and precinct (a spatial inter-subband partitioning structure that contains one to several codeblocks) of a tile of a certain component. A CCP may consist of a single or multiple codeword segments. Multiple codeword segments arise when a coding pass (in the CCP) is terminated. This will happen if all coding passes are terminated (JJ2000 option: `-Cterm all`).

The JPEG2000 codestream—the standard’s term for the JPEG2000 stream (cf. Sect. 3.3)—consists of headers (main header, tile headers, tile part headers) and packets that consist of packet headers and packet bodies (cf. Fig. 1). The compressed coefficient data is contained in the packet bodies.



**Fig. 1** Restrictions within the CCPs

The CCPs must not contain any two byte sequence in excess of  $0xff8f$  nor end with a  $0xff$  byte (bitstream compliance) [34, p. 56]. The arithmetic coding of the bitplanes is referred to as tier 1 encoding, while the partitioning of the coding passes into quality layers and the generation of the packet headers is referred to as tier 2 encoding.

### 3.1.1 JPEG2000 headers

The main header and tile-part header contain information about the specific compression parameters (e.g., image size, tile size, number of components, codeblock size, wavelet filters, ...). The packet header contains the following data items: inclusion information for each codeblock (does the codeblock contribute to this packet?), the lengths of the CCPs, the number of contributed coding passes for each codeblock, and the number of leading zero bitplanes for each codeblock (LZB).

### 3.2 Part 2: extensions

Part 2 of JPEG2000 specifies extended decoding processes, an extended codestream syntax containing information for interpreting the compressed image data, an extended file format, a container to store image meta-data and a standard set of image meta-data. The extensions of Part 2 allow employing custom wavelet transforms and arbitrary decomposition structures.

The extended coding processes are beneficial for certain applications, such as fingerprint compression and medical image compression [4,40,57]. As fingerprints and medical images contain sensitive information, security concerns naturally arise.

### 3.3 Part 1 and 4: bitstream, format and JPEG2000 compliance

The term “bitstream” in its common meaning refers to an arbitrary stream of bits. In the MPEG-4 standards bitstreams denote the compressed video stream [33]. The term “bitstream-oriented encryption” refers to the encryption of the compressed stream, i.e., the JPEG2000 codestream. However, in the JPEG2000 standard the term “bitstream” has a precisely defined alternate meaning. According to the JPEG2000 standard [30], “bitstream” is defined in the following way: “The actual sequence of bits resulting from the coding of a sequence of symbols. It does not include the markers or marker segments in the main and tile-part headers or the EOC marker. It does include any packet headers and in stream markers and marker segments not found within the main or tile-part headers” [30, p. 2].

Sequences in excess of  $0xff8f$  are used to signal in-bitstream markers and marker segments and therefore must

not be generated in the encryption process (schemes fulfilling this requirement and avoiding  $0xff$  bytes at the end of an encryption unit, i.e., codeword segment, CCP, or packet body, and preserving the length of the encryption unit are denoted bitstream-compliant). An encryption scheme delivering a valid JPEG2000 codestream (in the sense that it is decodable by the reference software) is denoted as format-compliant. Part 4 of the JPEG2000 standard suite (conformance testing) [31] defines the term “compliance” for JPEG2000 decoders and encoders. While JPEG2000 decoders have to decode certain test sets within given error bounds in order to be compliant, the only requirement for encoder compliance is to produce compliant codestreams (decodable by the reference software); any other requirements using quality criteria are not part of the standard [31, p. 30]. JPEG2000 compression with a compliant encoder, which is followed by encryption that results in a decodable JPEG2000 codestream, is therefore JPEG2000 compliant in the sense of [31].

### 3.4 Part 8: JPSEC

JPEG2000 Part 8 (JPSEC) has only recently become an official ISO standard (ISO/IEC 15444-8 [32]). The standardization process started with a call for proposals in March 2003 and since then quite a number of contributions have been made [1,2,5,12,13,62,63]. JPSEC is an open security framework for JPEG2000 that offers solutions for

- Encryption
- Conditional access
- Secure scalable streaming
- Verification of data integrity
- Authentication

Encryption, conditional access and secure scalable streaming overlap with the topic of this survey.

#### 3.4.1 JPSEC architecture

The JPSEC framework offers a syntax for the definition of JPEG2000 security services. This syntax specifies the JPSEC codestream. A JPSEC codestream is created from either an image, a JPEG2000 codestream or an existing JPSEC codestream. The last case applies if several security tools are applied subsequently.

Currently security tools are grouped into three types of tools, namely template, registration authority, and user-defined tools. Template tools are defined by the normative part of the standard, registration authority tools are registered with and defined by a JPSEC registration authority, and user-defined tools can be freely defined by users or applications. The standard defines a normative process for the registration



of registration authority tools. The registration authority and the user-defined tools enable the application of custom and proprietary encryption methods, which leads to a flexible framework.

In the following section a more detailed summary of the JPSEC codestream syntax and semantics is given.

### 3.4.2 The JPSEC syntax and semantics

JPSEC defines a new marker segment for the JPEG2000 main header (SEC marker segment), which is preceded only by the SIZ marker segment [32, p. 9]. Therefore the information of the SIZ marker segment is always preserved by JPSEC encryption. The SIZ marker contains information about the number of components of the source image, their resolutions (subsampling factors), their precision, as well as the chosen tile size. Note that this information is always accessible even if the most secure settings are chosen for JPSEC encryption (e.g., AES encryption of the entire remaining codestream).

The first SEC marker segment in a JPSEC codestream defines if INSEC marker segments are employed, the number of applied tools and the TRLCP format specification (the number of necessary bits to specify tile, resolution, layer, component, and precinct uniquely; in conjunction these indices uniquely identify a packet). The INSEC marker segment is used in conjunction with a non-normative tool and it may be present in the bitstream. The INSEC marker segment makes use of the fact that the JPEG2000 decoder stops decoding if a termination marker (a sequence in excess of  $0\text{x}\text{f}\text{f}\text{8}\text{f}$ ) is encountered. Thus encryption specific information can be placed directly in the JPEG2000 bitstream. The application of INSEC markers, though not without merits, also leads to certain drawbacks. First, the preservation of JPEG2000 format compliance, as defined in Sect. 3.3, requires the packet header to be changed (cf. to the approach of [25] in Sect. 5.2 for details). Second, if no specifically tailored encryption routines are employed (bitstream-compliant), the INSEC marker segment may not be parsed correctly. Therefore a useful application of INSEC markers is together with bitstream-compliant encryption algorithms (see Sect. 5.3).

The SEC marker segment also contains a list of tool specifications (one for each tool). The JPSEC tool specification follows a normative syntax and defines which type of tool is applied (either normative or non-normative), which specific tool is used, where it is applied (ZOI:= zone of influence) and its parameters (e.g., keys, initialization vectors, ...).

The ZOI can be specified via image or non-image related parameters. A ZOI specification consists of one or multiple zone descriptions, the ZOI is the union of all the zones. Each zone is described by several parameters of a description class (image related or non-image related). For image related parameters a zone is the region where all the parameters are met (intersection). If multiple non-image related parameters

are given, the specified regions should correspond to each other in a one to one manner, e.g., if packets and byte ranges are employed, the byte ranges specify the packet borders. In this manner the ZOI can be used to store meta-data of the codestream, e.g., where certain parts of the image are located in the codestream.

The image related description class allows to specify a zone via image regions, tiles, resolution levels, layers, components, precincts, TRLCP tags, packets, subbands, codeblocks, ROIs (regions of interest), and bitrates. The non-image related description class allows to specify packets, byte ranges (padded and unpadded ranges if padding bytes are added), TRLCP tags, distortion values, and relative importance. The distortion value and the relative importance may be set to signal to a decoder or adaptation element the importance of the specified ZOI. While the distortion value gives the total squared error if the corresponding ZOI is not available for decoding, the relative importance field is not tied to a specific quality metric. By employing these fields efficiently and in an informed way, transcoding can be conducted even if the JPSEC codestream consists of fully encrypted segments (see Sect. 3.4.3). The parameters of a tool also have to be specified; for normative tools the parameter description follows a distinct syntax, while non-normative tools may define their own syntax and semantics.

The parameter description for JPSEC normative tools consists of a template identifier and the corresponding template parameters for the tool, the processing domain, the granularity, and a list of the actual parameter values VL (initialization vectors, MAC values, digital signatures, ...).

There are three basic templates for JPSEC normative tools, namely the decryption template, the authentication template and the hash template. These are further subdivided. For the decryption template a block cipher template, a stream cipher template, and an asymmetric cipher template are defined. Several block ciphers are available (AES, TDEA, MISTY1, Camellia, Cast-128, and Seed), one stream cipher (SNOW 2), and one asymmetric cipher (RSA-OAEP).

The processing domain is used to indicate in which domain the JPSEC tool is applied. The possible domains are: pixel domain, wavelet domain, quantized wavelet domain, and codestream domain.

The granularity defines the processing order (independently of the actual progression order of the JPEG2000 codestream) and the granularity level. The granularity level may be component, resolution, layer, precinct, packet, subband, codeblock, or the entire ZOI. Thus the ZOI specifies a subset of the image data (either in the image domain or in the compressed domain), while the processing order specifies in which order these data are processed (which may differ from the progression order of the protected JPEG2000 codestream). The granularity level specifies the units in which the data are processed (which can be a further subset of



the data specified through the ZOI). The list of parameter values VL contains the appropriate parameter for each of these processing units.

The following example is given in the standard [32] to illustrate the relationship between ZOI, processing order, granularity level and the list of parameter values: A JPEG2000 codestream has been encoded with resolution progression (RLCP) and 3 resolution levels and 3 layers. The ZOI is defined by resolutions 0 and 1. The processing order is layer and the granularity level is resolution. Figure 2 illustrates the process, the value list VL would contain hash values (if hashing is applied).

The granularity syntax is employed by secure scalable streaming (SSS) as proposed in [1,2,62,63]. Its implementation within JPSEC is discussed in the next section.

### 3.4.3 JPSEC and bitstream-oriented encryption

The normative tools of JPSEC enable the rearrangement of JPEG2000 data (except the main header) into segments. It is possible to conduct the rearrangement across packet borders. The segments are then encrypted (see Fig. 3). Segment-based encryption enables very efficient secure

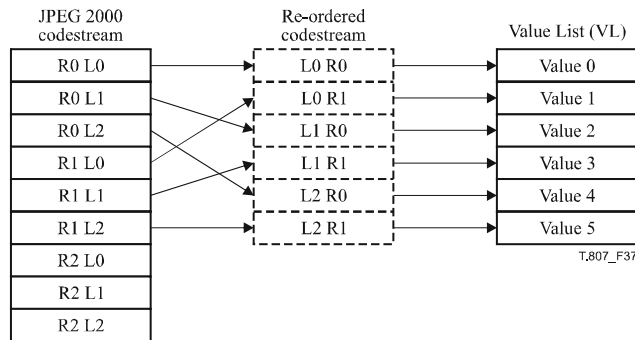


Fig. 2 Granularity level is resolution

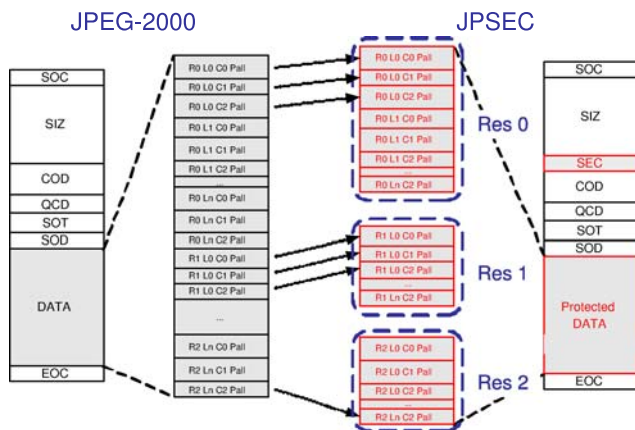


Fig. 3 Segment-based encryption

transcoding, i.e., SSS, because the meta-data of segmentation and encryption is stored in the SEC marker segment. Hence a JPSEC transcoder only needs to parse the main header for the SEC segment and truncate the JPSEC codestream at the according position. Compression performance is hardly influenced by this approach.

The rest of the JPEG2000 codestream (tile headers, packet headers and packet bodies) is reassembled into segments. The advantage of this approach is a low transcoding complexity, while a disadvantage is that rate adaption can only be done by a JPSEC-capable transcoder (but not by a transcoder that is only JPEG2000-compliant). In general, the advantages of format-compliant encryption are lost, but scalability is preserved to a definable level.

Format-compliant bitstream-oriented encryption schemes (see Sect. 5) can be implemented as non-normative tools.

### 3.4.4 JPSEC and compression-integrated encryption

JPSEC allows to specify the ZOI via image related parameters. The area specified by the ZOI may be encrypted by employing a normative tool. Normative tools allow the specification of a processing domain, e.g., pixel domain, wavelet domain, quantized wavelet domain, or codestream domain. If the wavelet domain or quantized wavelet domain is chosen, the processing domain field indicates whether the protection method is applied on the sign bit or on the most significant bit [32, p. 35]. Hence the encryption of rather freely definable portions of the wavelet transformed data is possible.

## 3.5 The interplay of JPEG2000, JPSEC and JPEG2000 encryption

JPSEC can be used to secure JPEG2000 codestreams. Annex C of the JPSEC standard [32, p. 91] elaborates in more detail on the interoperability of JPSEC and the other parts of the JPEG2000 standard suite. As a JPEG2000 Part 1 decoder will skip marker segments that it does not recognize (see [34, p. 28]), it is possible to place the SEC marker segment in the main header of JPEG2000 codestream and still preserve compliance to JPEG2000 Part 1. The term “Part 1 compliance” is defined in [32, p. 91] for JPSEC codestreams that have a strictly defined behavior for a JPEG2000 Part 1 decoder. Note that this definition of “Part 1 compliance” is stricter than the definition of “format compliance” for JPEG2000 given in Sect. 3.3. However, the definition given in Sect. 3.3 is sufficient for assessing the compliance for JPEG2000 encoders according to Part 4 of the standard and will therefore be sufficient for assessing format compliance. Many of the JPEG2000 encryption approaches will produce codestreams that are in accordance with both compliance definitions, e.g., the encryption via a random permutation of the wavelet coefficients (see Sect. 6.2.2).

In summary, JPSEC can be used to format-compliantly signal all the necessary parameters of a format-compliant encryption scheme.

The extended coding system of JPEG2000 Part 2 offers vast parameter spaces, and thus keeping the actual parameters secret (the chosen decomposition structure, or the wavelet filter) can be employed as a form of compression-integrated encryption. The advantage of this approach is that no additional decompression/decryption software is necessary, only the parameters have to be encrypted and decrypted.

Although the JPSEC standard has been tailored to JPEG2000 Part 1 codestreams, it is reasonable to employ the JPSEC syntax for the encryption of JPEG2000 Part 2 codestreams (e.g., if secret JPEG2000 Part 2 compression options are employed, the corresponding byte ranges containing these parameters are encrypted). Annex C.2 of the JPSEC standard discusses the interoperability with Part 2 and mentions that the usage of JPSEC can be signalled via Part 2 (CAP marker segment).

### 3.6 Application of the JPEG2000 standard

JPEG2000 does not yet dominate the mass market, but there are several areas where it has been widely adopted. Most interesting for the scope of this survey is the Digital Cinema Specification that defines JPEG2000 as intra-frame codec. As content and copyright protection play a major role in this area there is extensive coverage of security issues in the DCI specification.

#### 3.6.1 DCI's digital cinema specification

Despite the extensive coverage of security issues in [8], the defined encryption methods are conventional. The digital video is divided into reels of 10–20 min. These reels consist of several track files that may contain image, audio, subtitle and other meta-data [8, p. 44]. A track file starts with a file header and ends with a file footer. The track file body consists of several KLV (key length value) units. The key is an identifier for the content of the KLV unit, length specifies the length of the value. For an image track file, the image data is wrapped using KLV on an image frame boundary [8, p. 47]. For encryption, KLV units are simply mapped to new  $K^*L^*V^*$ . While  $K^*$  and  $L^*$  are the new identifier and length,  $V^*$  is composed of cryptographic options,  $K$ ,  $L$  and the encrypted  $V$ . In other words: the video is encrypted frame per frame. The application of the AES cipher in CBC mode with 128 bit keys is required.

The JPEG2000 file is fully encrypted, only its length and the fact that it is an image track file are known. Frame dropping can easily be implemented by ignoring the corresponding KLV unit. To transcode an encrypted image, its entire data has to be decrypted. Ciphertext bit errors affect one

block (16 bytes) and one bit of the JPEG2000 file [51]. The KLV, CBC and JPEG2000 systems are prone to synchronization errors, e.g., bit loss. For images this method corresponds to the naive encryption approach, while for videos it is notable that the compressed frame sizes are preserved in the encrypted domain and can potentially be used as a fingerprint.

#### 3.6.2 Software implementations

JPEG2000 Part 1 is implemented in the reference software: There is a C implementation (JasPer) and a Java implementation (JJ2000). Apart from the reference software there are several commercial implementations, e.g., Kakadu. For our experiments we employ JasPer (Version 1.900.1), JJ2000 (Versions 4.1 and 5.1), and Kakadu (Version 6.0).

## 4 Evaluation criteria for JPEG2000 encryption

In addition to the different categories discussed in Sect. 2, which relate to intended level of security, field of application and mode of operation, criteria for the evaluation of the different encryption schemes are necessary. While their diversity makes it hard to directly compare all schemes, there are some criteria common to all encryption schemes that can be used in an evaluative comparison.

### 4.1 Compression

Compression performance may suffer if encryption is applied. While most of the bitstream-oriented encryption schemes have no or only a negligible influence on the compression performance, many compression-integrated schemes may dramatically decrease the compression performance, especially if inaccurate parameters are chosen. However, the influence on compression performance may also depend strongly on the source image characteristics.

### 4.2 Security

Given the different security levels of various application scenarios (as defined above) the definition of security will vary. For high-level security every bit of information that is preserved during encryption reduces the security. However, none of the format-compliant encryption schemes discussed in this survey complies with these high standards. At least the main headers and the packet structure (packet body and header borders) are preserved in the encryption process. Thus linking a plaintext and a ciphertext is to some extent possible for all of the schemes.

For content security it has to be assessed if the image content is still discernible; the standard image quality metric

PSNR is not well suited for this task. There is a similar situation for sufficient encryption: it has to be assessed whether the image still has a commercial value. For transparent/perceptual encryption a certain image quality has to be preserved, but an attacker shall not be capable to further increase the image quality by exploiting all available information (e.g., the encrypted parts).

Hence, security for all but the high-level case may be defined by the level of resistance to increase the image quality by an attack. These attacks can exploit any of the preserved data, as well as context specific side channel information (e.g., some statistics of the source images may be known). This cryptoanalytic model for multimedia encryption has been proposed in [49] (furthermore a public low quality version is assumed here, which is not appropriate for the case of content security).

For these definitions of security the evaluation of image quality is necessary.

#### 4.2.1 Evaluation of image quality

The peak-signal-to-noise-ratio (PSNR) is no optimal choice for assessing image quality. A state-of-the-art image quality measure is the structural similarity index (SSIM) [61] and it ranges, with increasing similarity, between 0 and 1. Mao and Wu [42] propose a measure specifically for the security evaluation of encrypted images that separates luminance and edge information into a *luminance similarity score* (LSS) and an *edge similarity score* (ESS). LSS behaves in a way very similarly to PSNR. ESS is the more interesting part and ranges, with increasing similarity, between 0 and 1. We use the weights and block sizes proposed by [42] in combination with Sobel edge detection.

#### 4.3 Complexity

The proposed schemes are diverse, some need to run through the entire JPEG2000 compression pipeline (compression-integrated), others do not (or only partly). For some compression-integrated proposals the complexity of the compression process is increased (as for wavelet packets as described in Sect. 6.1.1), while for others the compression complexity remains unchanged (as for parameterized filters as described in Sect. 6.1.2).

Initially, one would assume that JPEG2000 encryption has to compete against conventional encryption (naive approach) in terms of runtime performance. However, most of the (runtime) benefits of JPEG2000 specific encryption schemes are due to the preservation of image and compressed domain properties in the encrypted domain. Probably the most important feature is scalability. If scalability is preserved, rate adaptation can be conducted in the encrypted domain, whereas otherwise the entire encrypted codestream needs to be

decrypted. The issue of key distribution is thereby greatly simplified, as the key does not need to be present for rate adaptation. All of the discussed JPEG2000 encryption schemes preserve the scalability to some extent and thus the direct comparison of the runtime with the naive approach is not representative for the actual runtime benefits.

In order to give an estimate of runtime performance of the various JPEG2000 encryption schemes, several time estimates are needed as reference. The bitstream-oriented schemes need to identify the relevant portions of the codestream. There are three possibilities: The first is to analyze the codestream in the same manner as a JPEG2000 decoder, basically the header and the packet headers need to be decoded. An alternative is to employ SOP and EPH marker sequences to identify the relevant portions. This method is extremely simple (parsing for two byte marker sequences) compared to relatively complex packet header decoding via several tag trees and contextual codes. The third possibility is to employ JPSEC as meta-language to identify the relevant parts at the decrypter/decoder side.

For compression-integrated schemes the runtime complexity of the compression pipeline is necessary.

The following numbers are based on a test set of 1,000 images ( $512 \times 512$ , 2bpp, single quality layer) and averages of 100 trials on an Intel(R) Core(TM)2 CPU 6700 @ 2.66 GHz. The results for header decoding have been obtained by modifying the reference software JasPer (see Sect. 3.6.2) and the results for SOP/EPH parsing have been obtained by a custom implementation. Additionally, for compression and decompression the results of the Kakadu implementation are given. Empirical results for:

- time of header decoding  
very low (370.92 fps, 23.18 MB/s)
- time for SOP/EPH parsing  
extremely low (1030.93 fps, 63.84 MB/s)
- time of compression  
high (JasPer: 12.89 fps, 0.81 MB/s, Kakadu with 2 threads: 41.19 fps, 2.57 MB/s, Kakadu with 1 thread: 25.00 fps, 1.56 MB/s)
- time of decompression  
high (JasPer: 21.45 fps, 1.34 MB/s, Kakadu with 2 threads): 60.18 fps, 3.76 MB/s, Kakadu with 1 thread: 40.23 fps, 2.51 MB/s)

Compared to compression and decompression, header decoding and SOP/EPH parsing are extremely computationally inexpensive. Therefore bitstream-oriented techniques are preferable if the visual data is already compressed. However, SOP/EPH parsing is significantly less expensive than header decoding (three times less according to our results).

## 5 Bitstream-oriented techniques

The basic unit of the JPEG2000 codestream is a packet, which consists of the packet header and the packet body (see Sect. 3.1). Almost all bitstream-oriented JPEG2000 encryption schemes proposed in literature target the packet bodies. Format-compliance can easily be preserved by adhering to a few syntax requirements (namely those that relate to bitstream compliance: no sequences in excess of `0xff8f` are allowed and the last byte must not equal `0xff`). Scalability is thereby preserved on a packet basis. Additionally, if the packet headers are preserved, the lengths of the plaintext parts (packet bodies) have to be preserved as well. Several bitstream-compliant encryption algorithms have been proposed and are discussed in Sect. 5.3.

Scalability at an even finer level than packets can be preserved if each CCP (or even more general, each codeword segment) is encrypted independently. If encryption modes are employed that need initialization vectors (IVs), it has to be guaranteed that the IVs can be generated at the decrypting side as well—even if allowed adaptations of the encrypted codestream have been performed during transmission. The generation of truncation and cropping invariant initialization vectors is discussed in [70]. Basically a codeblock can be uniquely identified in a JPEG2000 codestream (e.g., by specifying the component, the tile, the resolution, the subband, the precinct and the codeblock's upper left coordinates) and a codeblock contribution to a packet can be uniquely identified by the quality layer. In [70] tiles are identified by their position relative to a reference grid (which is truncation/cropping invariant) and codeword segments are identified by the first contributing coding pass.

Several contributions discuss how to enable scalable access control (e.g., a user only has access to the lowest resolution, as a preview) [32, p. 65], [27, 28, 67] with only one single master key. This is in general achieved via hash chains and hash trees.

In the following sections, we first discuss replacement attacks and their simulation by JPEG2000 error concealment in Sect. 5.1. In Sect. 5.2 we discuss format-compliant packet body encryption algorithms which require the packet header to be modified. Note that these schemes can also be applied on a CCP or codeword segment basis, preserving scalability on a finer granularity, but requiring every CCP or codeword segment length in the packet header to be changed. Then in Sect. 5.3 packet body encryption with bitstream-compliant algorithms is discussed which allows to preserve the original packet header (again these algorithms may also be applied on a CCP or codeword segment basis). Both of these schemes preserve practically all of the packet header information, which leads to serious security problems concerning content security/confidentiality. Therefore a

format-compliant packet header encryption algorithm is discussed in Sect. 5.4.

### 5.1 Security issues and attacks

If only parts of the JPEG2000 codestream are encrypted, these might be identified and replaced [45], thereby tremendously increasing the image quality of a reconstruction as compared to a reconstruction with the encrypted parts in place. A way to mimic these kinds of attacks is to exploit the JPEG2000 built-in error resilience tools [45]: while error resilience options are optional, they represent the same outcome that a possible attacker is likely to obtain by identifying the encrypted portions of the wavelet coefficients by means of a statistical analysis.

In the case of the JJ2000 decoder—preliminarily the error-correcting symbols have to be invoked by passing the `-cseg_symbol` option to the JJ2000 encoder—the erroneous bitplane and all successive bitplanes are discarded. This error concealment method protects each cleanup pass with 4 bits, as at the end of each cleanup pass an additional symbol is coded in uniform context (`0xa`). Additionally, further JPEG2000 error concealment strategies can be employed, such as predictive termination of all coding passes (invoked with `Cterminate all` and `Cterm_type predict`). Predictive termination of a coding pass protects the data with 3.5 bits in average, as error concealment information is written on the spare least significant bits of the coding pass.

It has to be noted that the JJ2000 library has minor bugs in the error concealment code (for details please cf. to [54] and [53]). The bug-fixed JJ2000 source code is available at <http://www.wavelab.at/~sources/>.

Attacks which use the error-concealment mechanisms to identify and replace the encrypted portions of the codestream are called error-concealment attacks or replacement attacks.

### 5.2 Packet body encryption with packet header modification

One of the first contributions to JPEG2000 security was made by Grosbois, et al. [25]. The packet body data is conventionally encrypted (they propose to XOR the packet body bytes with key bytes derived from a PRNG), which introduces the problem of superfluous marker generation (conventional encryption does not preserve bitstream compliance), however, this topic is not further discussed in [25]. They propose storing security information (e.g., encryption key, hash value) at the end of a codeblock's bitstream after an explicit termination marker. This method was later adapted in several contributions, namely by Dufaux and Ebrahimi [13] and Norcen and Uhl [45].

The application of this method is not as straightforward as it seems to be. "The codeword segment is considered to



be exhausted if all  $L_{\max}$  bytes (all the bytes contributing to a codeword segment) are read or if any marker code in the range of  $ff90_h$  through  $fffef_h$  is encountered.” [56, p. 483]. In practice this means that it is not sufficient to simply add explicit termination markers at the end of the codeblock’s bitstream in order to add data to the codestream, furthermore the overall length of the packet has to be adjusted in the packet header. Nevertheless, it is possible to overwrite packet body data (then the packet header does not need to be changed), but this causes noise in the reconstructed image. Only if the termination marker is placed at the end of the codestream (where the desired image quality has already been reached) the image quality is not lowered. However, it has to be taken into account that the last packets will be the first to be removed in the process of rate adaptation.

The approach can avoid special encryption schemes by storing the information about superfluous markers after the explicit termination marker. Neither in [25] nor in [45] is this topic discussed further. Norcen and Uhl [45] define the encryption process, namely AES in CFB mode.

We propose a simple method to avoid marker sequences: We use the  $0xff8f$  sequence to signal that a sequence in excess or equal to  $0xff8f$  has been produced. Hence for every generated sequence in excess of  $0xff8e$  an additional byte has to be stored. This byte can easily be appended to the packet body, the subtraction of one (all appended bytes are in excess of  $0x8e$ ) removes the possibility of marker code generation in the appended bytes.

In [66], Mao and Wu discuss several general communication-friendly encryption schemes for multimedia. Their work includes syntax-aware bitstream encryption with bit stuffing that can be applied to JPEG2000 as well. In the basic approach for JPEG2000, conventional encryption is applied and for every byte in excess of  $0xff$  an additional zero bit is stuffed in. In this way, bitstream compliance of the packet bodies is achieved. The decoder reverts this process by deleting every stuffed zero bit after a  $0xff$  byte in the ciphertext and then conventionally decrypting the resulting modified ciphertext.

It has to be considered that the JPEG2000 packet body has to be byte-aligned. Therefore, if the number of stuffed bits is not divisible by eight, additional bits have to be added. We propose simply filling up the remaining part with zero bits. Thereby no marker can be generated. The resulting encrypted packet body length then has to be updated in the packet header. To reconstruct the ciphertext, which is then decrypted, the bit stuffing procedure is reversed and the superfluous zero bits (outside the byte boundary) are ignored.

**Compression** Compression performance is negligibly reduced, depending on the method to achieve bitstream compliance. If bitstream compliance is achieved by signaling violations with  $0xff8f$ , one additional byte for approximately

every 579 bytes is generated (on average every 256th byte is a  $0xff$  byte and the next byte is in 113 of 256 cases in excess of  $0x8e$ ). If bitstream compliance is achieved via bit stuffing on average one bit every 256 bytes plus the 0–7 padding bits are added.

**Security** The headers are not encrypted and only slightly modified.

**Performance** These schemes perform very well, however, as the packet headers need to be altered, JPEG2000’s tier2 decoding and encoding has to be (at least partly) conducted.

### 5.3 Packet body encryption with bitstream-compliant algorithms

In the following bitstream-compliant encryption algorithms are presented. Bitstream-compliant encryption algorithms differ only in terms of information leakage (amount of preserved plaintext) and computational complexity. The common properties and an experimental performance analysis of the discussed schemes are discussed in Sect. 5.3.8.

#### 5.3.1 Conan et al. and Kiya et al.

The algorithm by Conan is not capable of encrypting all of the packet body data but can be implemented rather efficiently [6]. Only the 4 LSBits (least significant bits) of a byte are encrypted, and only if the byte’s value is below  $0xf0$ . In this way no sequence in excess of  $0xff8f$  is produced. It is easy to see that no bytes are encrypted to  $0xff$ , because only the lower half of the bytes below  $0xf0$  are encrypted. The byte  $0xff$  is preserved. Hence a byte sequence in excess of  $0xff8f$  could only be produced after a preserved  $0xff$  byte. However, due to the bitstream-compliance the plaintext byte after a  $0xff$  byte is not in excess of  $8f$ .

Kiya et al. [35] extend this approach to an even more lightweight and flexible scheme. They propose encrypting only one randomly chosen byte out of  $m$  bytes. In this way the choice of the parameter  $m$  can trade-off security for performance. Additionally they propose a random shift of the 4 LSBits instead of encryption. The shifting operation can be applied to the 4 LSBits of all bytes, while preserving bitstream compliance.

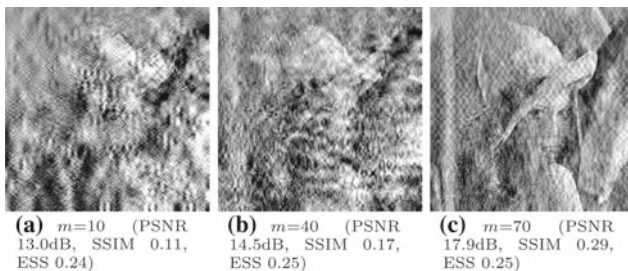
**Security** The information leakage is very high, even with the most secure settings more than half of the compressed coefficient data remains unencrypted. Assuming state-of-the-art underlying encryption techniques, the encrypted half bytes are irrecoverable by cryptographic means and therefore the whole arithmetic codeword is in general irrecoverable.

The shifting algorithm is less secure since there are only 4 possible ciphertexts for a half byte as compared to 16 for the encryption algorithm (which of course have to be different for every byte). Figures 4 and 5 show the result of the direct reconstruction, i.e., the full reconstruction of the encrypted codestream without trying to conceal the encrypted parts, and the concealment attack and different values of the parameter  $m$ . The shifting algorithm preserves more image plaintext than the encryption algorithm. Apart from error concealment options (segmentation symbol, predictive termination of all coding passes and SOP and EPH marker), the compression parameters have been set to JJ2000 default values, i.e., layer progression, 32 quality layers and no limit on bitrate (default bitrate is 100 bpp). If the 4 LSBits are encrypted, hardly any image information is visible for  $m$  up to 10. If the 4 LSBits are shifted, image information starts to become visible for  $m$  greater than one.

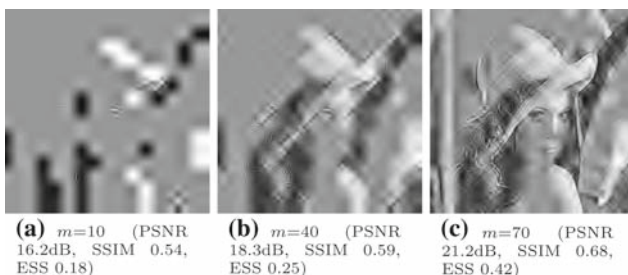
**Performance** Half byte encryption and permutation cannot be implemented much more efficiently than byte encryption on standard CPUs. For every encrypted byte one condition (is the byte below  $0\text{xff}0$ ) has to be evaluated.

### 5.3.2 Wu and Ma

Wu and Ma [65] greatly reduce the amount of information leakage compared to the algorithm of Conan and Kiya (see Sect. 5.3.1). They propose two algorithms for format-compliant packet body encryption. Both algorithms only preserve the  $0\text{xff}$  byte and its consecutive byte and can be implemented efficiently.



**Fig. 4** Kiya: Encryption of the 4 LSBits: direct reconstruction



**Fig. 5** Kiya: Encryption of the 4 LSBits: concealment attack

### Stream Cipher Based Algorithm:

Their first algorithm is based on a stream cipher (in [65] RC4 is employed). To that end a keystream is generated. By discarding  $0\text{xff}$  bytes, a modified keystream  $S$  is obtained. In the following, the term  $s_i$  denotes the  $i$ th byte of the keystream,  $m_i$  denotes the  $i$ th byte of the packet body (plaintext) and  $c_i$  denotes the  $i$ th ciphertext byte.

The encryption works byte by byte on the packet body in the following way:

```

If  $m_1$  equals  $0\text{xff}$ 
then  $c_1 = m_1$ 
else  $c_1 = m_1 + s_1 \pmod{0\text{xff}}$ 
For  $i = 2$  to length
    If ( $m_i$  equals  $0\text{xff}$ ) or ( $m_{i-1}$  equals  $0\text{xff}$ )
    then  $c_i = m_i$ 
    else  $c_i = m_i + s_i \pmod{0\text{xff}}$ ;
  
```

Every byte that is not a  $0\text{xff}$  byte or the successor of a  $0\text{xff}$  byte is encrypted to the range  $[0\text{x}00, 0\text{x}fe]$ . The decryption algorithm works similarly:

```

If  $c_1$  equals  $0\text{xff}$ 
then  $m_1 = c_1$ 
else  $m_1 = c_1 - m_1 \pmod{0\text{xff}}$ 
For  $i = 2$  to length
    If ( $m_i$  equals  $0\text{xff}$ ) or ( $m_{i-1}$  equals  $0\text{xff}$ )
    then  $c_i = m_i$ 
    else  $m_i = c_i - s_i \pmod{0\text{xff}}$ 
  
```

This algorithm avoids producing values in excess of  $0\text{xff}8\text{f}$ , since no  $0\text{xff}$ s are produced and all two byte sequences ( $0\text{xff}, X$ ) are preserved.

### Block Cipher Based Algorithm:

This algorithm preserves all two byte sequences ( $0\text{xff}, X$ ) as well, and basically works as follows:

Select all bytes of the packet body that are neither equal to  $0\text{xff}$  nor a successor of a  $0\text{xff}$  byte.  
 Split the selected bytes into blocks and encrypt them iteratively until no  $0\text{xff}$  is contained in the ciphertext.  
 Replace the selected bytes in the original block with the encrypted bytes.

When the number of selected bytes is not a multiple of the blocksize, ciphertext stealing (as described in [51]) is proposed. This algorithm does not contain any feedback mechanisms (equal packets yield equal ciphertexts, a basis for replay attacks). In [65] AES is employed.



**Security** Only (0xff, X) sequences are preserved, which renders the reconstruction of the image content unfeasible.

**Performance** The stream cipher based algorithm (as presented) needs to evaluate two conditions for every encrypted byte and needs to perform a modulo operation for every encrypted byte. Every 0xff byte has to be discarded from the keystream as well, thus slightly more encryption operations are necessary compared to conventional encryption.

For the block cipher based algorithm the probability for a ciphertext that does not contain any 0xff byte has to be assessed. In general this probability is  $(\frac{255}{256})^n$ , where  $n$  is the length of the plaintext and the encryption method is assumed to be uniformly distributed. This results in a success probability of approximately 96.9% for a block of 8 bytes, which increases the encryption time about 3.18% compared to the underlying encryption routine (AES with 16 bytes or Triple-DES with 8 byte are proposed in [65]). For a block of 16 bytes the success probability is 93.9%, which corresponds to an overhead of 6.46%. Thus the overhead induced by additional encryption is modest as well. However, this algorithm has additional copy operations (the bytes have to be copied to a buffer before the iterative encryption).

### 5.3.3 Dufaux et al.

The encryption algorithm presented by Dufaux et al. [12] is basically an improvement of the algorithm by Wu and Ma [65] in terms of reduced information leakage. Only 0xff bytes are preserved. They propose the usage of the SHA1 PRNG with a seed of 64 bit for keystream generation; however, any other cryptographically secure method of generating an appropriate keystream can also be applied. The encryption procedure is the following:

If  $m_i$  equals 0xff  
 then  $c_i = m_i$   
 If  $m_{i-1}$  equals 0xff  
 then  $c_i = m_i + s_i \pmod{0x8f+1}$ , and  $s_i \in [0x00, 0x8f]$ .  
 If  $m_{i-1}$  does not equal 0xff  
 then  $c_i = m_i + s_i \pmod{0xff}$ , and  $s_i \in [0x00, 0xfe]$ .

The proposed method to obtain a number in the right range is the iterative generation of random numbers until a number in the right range is produced. Decryption works analogously.

**Security** The information leakage is further reduced compared to the algorithm by [65], only 0xff bytes are preserved (every 128th byte, cf. Sect. 5.3.8).

**Performance** Slightly more keystream bytes have to be used, since the 0xff bytes have to be discarded for the keystream and after a 0xff byte, bytes with values in excess

of 0x8f have to be discarded. Additionally, one condition and one modulo operation have to be evaluated for every encrypted byte.

### 5.3.4 A JPSEC technology example

In [32, p. 72] a method for format-compliant encryption is sketched in Annex B.5 (Technology examples: Encryption tool for JPEG2000 access control). The document, however, does not contain all necessary details to implement the method. On the contrary, in [29] it is pointed out that the proof for the reversibility of the algorithm is still missing.

The encryption process is defined in the following way:

The packet body is split into two byte sequences. Every two byte sequence of the packet body is temporarily encrypted. If the temporary byte sequence or its *relating code* is more than 0xff8f it is not encrypted, otherwise the temporarily encrypted code is outputted as ciphertext.

If the length of the plaintext is odd it is proposed to leave the byte in plaintext or pad an extra byte. The padding of an extra byte would require the modification of the packet header. The decryption process is similarly specified:

The packet body is split into two byte sequences. Every two byte sequence is temporarily decrypted. If the temporary byte sequence or its *relating code* is more than 0xff8f it is not decrypted, otherwise the temporarily decrypted code is outputted as plaintext.

It is notable that the underlying encryption routine for two byte sequences must satisfy the following property  $e(p) = c = d(p)$  and thus  $e(e(p)) = p$ . This is met by all encryption modes that xor the plaintext with a keystream (e.g., OFB mode). The term *relating code* is not further specified. Furthermore it is possible for an encrypted packet body to end with 0xff, which might lead to problems (a marker sequence at packet borders is possibly generated).

In [15] an interpretation for the term *relating code* is given which makes the scheme reversible (a proof is given): Let  $P_j$  denote the  $j$ th plaintext two byte sequence,  $I_j$  the  $j$ th temporarily encrypted two byte plaintext sequence,  $C_j$  the  $j$ th two byte ciphertext sequence and  $D_j$  the  $j$ th temporarily decrypted ciphertext sequence. The term  $X|Y$  denotes the concatenation of the second byte of  $X$  and the first byte of  $Y$ , where  $X$  and  $Y$  are arbitrary two byte sequences. If the following conditions are met, then the ciphertext  $C_j$  is set to the temporarily encrypted sequence  $I_j$ :

- E1  $I_j \leq 0\text{xff}8\text{f}$   
Necessary to obtain a bitstream-compliant two byte ciphertext sequence.
- E2  $P_{j-1}|I_j \leq 0\text{xff}8\text{f}$   
Necessary to ensure bitstream compliance if the previous two byte sequence has been left in plaintext.
- E3  $I_{j-1}|I_j \leq 0\text{xff}8\text{f}$   
Necessary to ensure bitstream compliance if the previous two byte sequence has been replaced by the temporarily encrypted sequence.
- E4  $I_j|P_{j+1} \leq 0\text{xff}8\text{f}$   
Necessary to be able to preserve the next two byte sequence in plaintext.
- E5  $I_{j-1}|P_j \leq 0\text{xff}8\text{f}$   
Necessary to detect E4 for  $j - 1$ .

In order to decrypt the  $j$ th ciphertext the following conditions have to be met:

- D1  $D_j \leq 0\text{xff}8\text{f}$   
Detection of the violation of E1 (if E1 has not been met D1 is not met and the ciphertext is the plaintext).
- D2  $P_{j-1}|D_j \leq 0\text{xff}8\text{f}$   
Detection of the violation of E2.
- D3  $I_{j-1}|D_j \leq 0\text{xff}8\text{f}$   
Detection of the violation of E3.
- D4  $D_j|C_{j+1} \leq 0\text{xff}8\text{f}$   
Detection of the violation of E4.
- D5  $I_{j-1}|C_j \leq 0\text{xff}8\text{f}$   
Detection of the violation of E5.

All conditions referencing undefined bytes (e.g.,  $P_{-1}$ ) are by default true. Note that in the case of an even number of packet body bytes, the last two byte sequence requires special treatment. In this case the best solution (in terms of maximum encryption percentage) is to modify  $E1$  and  $D1$  such that a byte with value  $0\text{xff}$  at the end is forbidden.

**Security** Information leakage occurs whenever a two byte sequence of plaintext is preserved. Our experiments, which implement the algorithm specified in [15], reveal that about every 128th byte is preserved (cf. Sect. 5.3.8). However, the preserved two byte sequences are not distinguishable from the encrypted sequences (compared to the previous bitstream-compliant algorithms, that always preserve the  $0\text{xff}$  byte). Thus the algorithm is an improvement over the previous bitstream-compliant algorithms, as, for example, the two encrypted versions (different encryption keys) preserve totally different plaintext bytes.

**Performance** There is a slight performance overhead, due to the additional comparisons (five conditions for every two byte sequence).

### 5.3.5 Wu and Deng

The iterative encryption which works on CCPs was proposed by Wu and Deng in [68]. While the iterative encryption algorithm is capable of encrypting all of the packet body data, it cannot be implemented very efficiently. Contrary to most other schemes the iterative encryption algorithm does not preserve any plaintext information (except its length). The CCPs are recursively encrypted until they are bitstream-compliant. The basic encryption algorithm is the following: For all CCPs:

```

ccpmid = encrypt(CCP)
While (isNotBitstreamCompliant(ccpmid))
    ccpmid = encrypt(CCP)
Output ccpmid as ciphertext.

```

In [68] addition modulo  $256^n$  is proposed as encryption method for the CCPs, however, encryption with the ECB mode of a blockcipher and ciphertext stealing [51] works as well and can be expected to be more efficient (and is therefore used in our experiments, see Sect. 5.3.8). Accordingly, for decryption the ciphertext is iteratively decrypted until it is bitstream-compliant. This algorithm is fully reversible and encrypts 100% of the packet body data.

Theoretically this algorithm can easily be extended to packet bodies by iteratively encrypting the packet bodies. However, the computational complexity of this algorithm will in general prevent the application of this algorithm on a packet body basis.

**Compression** If this scheme is applied to CCPs there is no direct influence on the compression performance, but certain parameter settings will be required to reduce the CCP lengths (e.g., enough quality layers [52]) that may reduce the compression performance.

**Security** There is no information leakage for the encrypted packet bodies.

**Performance** A detailed performance analysis of this scheme has been conducted by Stütz and Uhl [52] who show that the complexity of this algorithm increases dramatically with the length of the plaintext. Thus this algorithm is only feasible for certain coding settings which guarantee short CCP lengths, e.g., the choice of enough quality layers is necessary [52]. No stream processing is possible, the entire plaintext/ciphertext has to be kept in memory.

### 5.3.6 Zhu et al.

Zhu et al. [70, 72] propose to apply their bitstream-compliant scheme on a codeword segment basis.

1. The plaintext is XOR ed with a keystream.
2. In this intermediate ciphertext every byte is checked to meet the bitstream compliance. If an illegal byte (its value concatenated with the value of the next byte is in excess of  $0\text{x}\text{ff}\text{8}\text{f}$ ) is found (at index *currIdx*), this and the next (if there is one) intermediate ciphertext byte are replaced with the plaintext bytes at the same location (same indices).
  - (a) Now it is checked whether this replacement results in a bitstream syntax violation of the decrypted intermediate ciphertext.
    - i Therefore the previous byte of the intermediate ciphertext is decrypted and together with the decrypted plaintext byte checked for bitstream compliance (are the two bytes concatenated in excess of  $0\text{x}\text{ff}\text{8}\text{f}$ ).
    - ii If this two byte sequence is illegal, this byte is also replaced with the plaintext byte in the intermediate ciphertext.
    - iii This procedure is conducted backwards until no more illegal byte sequences are found or a certain index (*lastModIdx* + 1, which is initialized with -1) is reached.
  - (b) If any byte has been replaced in (a), the intermediate ciphertext itself is checked for bitstream compliance.
    - (i) Therefore the previous byte of the last intermediate ciphertext byte that has been changed to the plaintext is checked for bitstream compliance.
    - (ii) If it is illegal, it is replaced with the corresponding plaintext byte.
    - (iii) This procedure is conducted backwards until no more illegal byte sequences are found or a certain index (*lastModIdx* + 1) is reached.
3. (a) and (b) are repeatedly executed until no illegal bytes are found in (a) and (b).
4. The index *lastModIdx* is then set to *currIdx* and the forward search for illegal bytes in the intermediate ciphertext is continued.
5. At the end the intermediate ciphertext is outputted as ciphertext.

Why is this scheme reversible and why is the decryption algorithm the same as the encryption algorithm? If no replacements have been conducted this is obviously the case as  $M \text{ XOR } S \text{ XOR } S \text{ equals } M$ . A precise argument for the

general case may be rather complex, but the scheme relies on the simple fact that a certain plaintext sequence and a certain key sequence result in an illegal ciphertext, which may be cause for illegal intermediate decrypted sequences (see (a)) or illegal intermediate ciphertexts (see (b)), which are all “switched back” to the plaintext. As for those pairs of sequences the plaintext is preserved, this property is preserved for the ciphertext and thus perfect reconstruction is possible.

**Security** According to the authors 0.36% of the plaintext are preserved [70].

**Performance** The encryption via XOR is very fast, but the searching for bitstream syntax violations is necessary. The entire plaintext/ciphertext has to be kept in memory (no stream processing).

### 5.3.7 Fang and Sun

The algorithm presented by Fang and Sun in [22] does not preserve any plaintext byte sequence and can be applied to CCPs and packet bodies. Nevertheless it is a computationally rather inexpensive procedure that concurrently works on three consequent plaintext bytes (the other schemes only consider two consequent plaintext bytes). The actual encryption and decryption algorithms are rather complex, therefore we will also give their pseudo code.

The first byte is encrypted depending on the second byte. If the second byte is in excess of  $0\text{x}\text{8}\text{f}$ , then according to the bitstream syntax, the first byte must not be encrypted to  $0\text{x}\text{ff}$ . If  $m_1 + s_1$  equals  $0\text{x}\text{ff}$ , then  $c_1$  is set to  $m_1 + 2s_1$ , which cannot yield  $0\text{x}\text{ff}$  too (only possible if  $s_1$  is zero and  $m_1$  is  $0\text{x}\text{ff}$ , but then, according to the bitstream compliance, the second byte cannot be in excess of  $0\text{x}\text{8}\text{f}$ ).

The second byte (and all following, except the last one) is encrypted depending on the previously encrypted plaintext byte, the previous cipher byte and the encryption of the previous byte (if it employed double encryption), the current plaintext byte and the next plaintext byte. There are basically three cases:

1. If  $m_{i-1}$  or  $c_{i-1}$  are  $0\text{x}\text{ff}$  then the current byte is encrypted to the range  $0\text{x}\text{00}$  to  $0\text{x}\text{8}\text{f}$  (this is possible because both facts indicate that the current plaintext byte is not in excess of  $0\text{x}\text{8}\text{f}$ ).
2. If the current byte is in excess of  $0\text{x}\text{8}\text{f}$  and the previous byte has been encrypted twice, then this property has to be preserved (encryption in the range  $0\text{x}\text{90}$  to  $0\text{x}\text{ff}$ ) to signal the double encryption in the decryption process. The next plaintext byte has to be considered as well; if it is in excess of  $0\text{x}\text{8}\text{f}$ , then the current cipher byte must not become  $0\text{x}\text{ff}$ , which is again avoided by double encryption.

3. In all other cases the byte is encrypted and double encryption is conducted if the next byte is in excess of  $0x8f$  and the cipher byte would be  $0xff$ .

The last byte is encrypted such that it is ensured that the cipher byte is in the same range (either  $0x00$  to  $0x8f$  or  $0x90$  to  $0xff$ ).

We can confirm that this encryption process is reversible (also experimentally). The pseudo code of the encryption and the decryption algorithm is given.

The encryption algorithm works in the following way:

$$c_1 = c_{mid} = (m_1 + s_1) \bmod 256$$

If  $c_{mid}$  equals  $0xff$  and  $m_2 \geq 0x90$

$$\text{then } c_1 = c_{mid} + s_1 \bmod 256$$

For  $i = 2$  to  $length - 1$

If  $m_{i-1}$  equals  $0xff$  or  $c_{i-1}$  equals  $0xff$

$$\text{then } c_i = (m_i + s_i) \bmod 0x90, c_{mid} = 0x00$$

else

if  $c_{mid}$  equals  $0xff$

$$\text{then } c_{mid} = (m_i - 0x90 + s_i) \bmod 0x70 + 0x90$$

If  $c_{mid}$  equals  $0xff$  and  $m_{i+1} \geq 0x90$

$$\text{then } c_i = (c_{mid} - 0x90 + s_i) \bmod 0x70 + 0x90$$

else  $c_i = c_{mid}$

else  $c_{mid} = (m_i + s_i) \bmod 256$

If  $c_{mid}$  equals  $0xff$  and  $m_{i+1} \geq 0x90$

$$\text{then } c_i = (c_{mid} + s_i) \bmod 256$$

else  $c_i = c_{mid}$

If  $m_{length} < 0x90$

$$\text{then } c_{length} = (m_{length} + s_{length}) \bmod 0x90$$

$$\text{else } c_{length} = (m_{length} - 0x90 + s_{length}) \bmod 0x70 + 0x90$$

If  $c_{length}$  equals  $0xff$

$$\text{then } c_{length} = (c_{length} - 0x90 + s_{length}) \bmod 0x70 + 0x90$$

The decryption algorithm works in the following way:

$$m_1 = m_{mid} = (c_1 - s_1) \bmod 256$$

If  $m_{mid}$  equals  $0xff$  and  $c_2 \geq 0x90$

$$\text{then } m_1 = (m_{mid} - s_1) \bmod 256$$

For  $i = 2$  to  $length - 1$

If  $m_{i-1}$  equals  $0xff$  or  $c_{i-1}$  equals  $0xff$

$$\text{then } m_i = (c_i - s_i) \bmod 0x90, m_{mid} = 0x00$$

else

if  $m_{mid}$  equals  $0xff$

$$\text{then } m_{mid} = (c_i - 0x90 - s_i) \bmod 0x70 + 0x90$$

If  $m_{mid}$  equals  $0xff$  and  $c_{i+1} \geq 0x90$

$$\text{then } m_i = (m_{mid} - 0x90 - s_i) \bmod 0x70 + 0x90$$

else  $m_i = m_{mid}$

$$\text{else } m_{mid} = (c_i - s_i) \bmod 256$$

If  $m_{mid}$  equals  $0xff$  and  $c_{i+1} \geq 0x90$

$$\text{then } m_i = (m_{mid} - s_i) \bmod 256$$

else  $m_i = m_{mid}$

If  $c_{length} < 0x90$

$$\text{then } m_{length} = (c_{length} - s_{length}) \bmod 0x90$$

$$\text{else } m_{length} = (c_{length} - 0x90 - s_{length}) \bmod 0x70 + 0x90$$

If  $m_{length}$  equals  $0xff$

$$\text{then } m_{length} = (m_{length} - 0x90 - s_n) \bmod 0x70 + 0x90$$

**Security** No byte sequence is preserved. However, from a cryptographic point of view there is a small weakness in this scheme.

The encryption operation  $(c_i + s_i) \bmod 0x90$  introduces a bias and therefore does not meet high cryptographic standards. The same holds for the operation  $(m_{mid} - 0x90 + s_i) \bmod 0x70 + 0x90$ . However, this bias can be removed by requiring the proper range for  $s_i$  to be  $0x00$  to  $0x8f$  in the first case and  $0x00$  to  $0x6f$  in the second case. This can easily be integrated into the algorithm by simply ignoring out-of-range keystream bytes in case of an encryption to a restricted range. In case of encryption to the range  $0x00$  to  $0x6f$  it is more efficient to halve the key byte before testing its range.

The answer to the question to which extent plaintext information is preserved in this scheme is beyond the scope of this survey paper. It is, however, intuitively dubious that every possible codeword has the same probability of becoming a plaintext's ciphertext. Nevertheless it needs to be pointed out that the information leakage is considered to be less than in all other algorithms except the iterative algorithm.

**Performance** Several conditions have to be evaluated in order to encrypt a single byte. If bias is to be prevented some keystream bytes have to be discarded.

### 5.3.8 Discussion of packet body encryption with bitstream-compliant algorithms

Several properties are shared by all approaches that employ bitstream-compliant algorithms.

**Compression** There is no influence on compression performance if bitstream-compliant encryption algorithms are applied.

**Security** The packet headers are preserved if bitstream-compliant encryption algorithms are applied (as proposed in



literature). There are known attacks (e.g., the error concealment attack) concerning selective/partial application of these bitstream-compliant schemes (see Sect. 5.1). The packet body encryption with bitstream-compliant encryption algorithms is not secure under IND-CPA (Indistinguishability under chosen-plaintext attack), as a potential attacker is very likely to successfully identify the corresponding ciphertext for a plaintext compressed image.

Most of the proposed schemes preserve plaintext bytes or properties, which is not a major concern as the headers and the packet headers already deliver a distinct fingerprint of the JPEG2000 codestream [15] (which is preserved in any case for all of the schemes, cf. Sect. 5.4).

However, only the iterative encryption algorithm of Wu and Deng is expected to be secure against IND-CPA attacks (only considering a packet body and disregarding the fingerprint obtained by the headers and packet headers).

In the following we present an empirical evaluation of information leakage of the presented bitstream-compliant encryption algorithms.

*Empiric Information Leakage:* In order to assess the amount of information leakage we give the average number of bytes until one byte is preserved.

Average number of bytes for one byte preservation	
Conan and Kiya ( $m = 1$ )	2
Mao and Wu	125.61
Dufaux	251.22
JPSEC techn. example	128.45
Zhu	312.50

The algorithm by Conan and Kiya preserves at least half of the plaintext, hence linking ciphertext and plaintext is obviously trivial. Plaintext two byte sequences starting with  $0xff$  are preserved in the ciphertext for the two algorithms by Mao and Wu, while only the  $0xff$  bytes are preserved for the algorithm by Dufaux et al. Hence plaintext and ciphertext have the same number of these sequences or  $0xff$  bytes at the same positions, which greatly simplifies the linking of the two. For the JPSEC Technology Example algorithm it is not known which sequences are preserved since the decision if a byte is preserved depends on the temporarily encrypted bytes (unknown to an attacker) as well. Hence the linking of plaintext and ciphertext has to exploit higher correlation between the two, which renders the process more complicated and less certain. Zhu's algorithm significantly reduces the information leakage (but stream processing is no longer possible). The algorithm by Fang and Sun does not preserve any plaintext byte but preserves some of the properties, which can be again exploited by a statistical analysis. In detail there are 4,789 bytes encrypted to the range  $0x90$  to  $0xff$  and 17,159 bytes encrypted to the range  $0x00$  to  $0x8f$  of a total

of 2,782,951 encrypted bytes. The certainty of the linking is expected to be further reduced.

**Performance** For all of the presented bitstream-compliant encryption algorithms, except the iterative encryption algorithm (see Sect. 5.3.5), the throughput (encrypted bytes per second) is independent of the plaintext length (disregarding the initialization overhead for the underlying encryption routine). All of the algorithms employ a cryptographic primitive (stream cipher, block cipher, secure random number generator) to obtain cryptographically secure randomness. However, the specific choice of primitives to be employed varies greatly.

In order to experimentally assess the runtime performance of the bitstream-compliant encryption algorithms discussed in this survey a single source of randomness is applied, namely AES. If a stream cipher is employed in the original contribution, AES is used in OFB mode to produce the keystream (in case of the application of a block cipher AES is used directly in ECB mode). In the following table stable results (128MB of plaintext data have been encrypted 150 times to obtain these results) for the throughput of the bitstream encryption algorithms are presented.

Throughput of bitstream-compliant encryption	
AES OFB	42.71 MB/s
Conan and Kiya ( $m = 1$ )	37.87 MB/s
Conan and Kiya ( $m = 10$ )	156.63 MB/s
Conan and Kiya ( $m = 40$ )	606.83 MB/s
Wu and Ma Stream	27.60 MB/s
Wu and Ma Block	1.53 MB/s
Dufaux	28.18 MB/s
JPSEC techn. example	37.81 MB/s
Zhu	29.30 MB/s
Fang and Sun	27.95 MB/s

For a large test set of 1,000 images the following throughputs have been achieved with SOP/EPH marker parsing (including all file reads and parsing). Results are presented for 1, 20 and 100% encryption of the packet body data, thereby covering the encryption percentage for all of the feasible application scenarios (see Sect. 5.5).

1% Encrypted	
Conan and Kiya ( $m = 1$ )	41.13 MB/s
Mao and Wu Stream	40.69 MB/s
Mao and Wu Block	40.88 MB/s
Dufaux	40.66 MB/s
JPSEC techn. example	41.01 MB/s
Zhu	40.87 MB/s
Fang and Sun	40.30 MB/s



20% Encrypted	
Conan and Kiya ( $m = 1$ )	33.41 MB/s
Mao and Wu Stream	29.24 MB/s
Mao and Wu Block	16.19 MB/s
Dufaux	29.67 MB/s
JPSEC techn. example	33.48 MB/s
Zhu	31.66 MB/s
Fang and Sun	29.73 MB/s
100% Encrypted	
Conan and Kiya ( $m = 1$ )	19.27 MB/s
Mao and Wu Stream	14.05 MB/s
Mao and Wu Block	3.83 MB/s
Dufaux	14.39 MB/s
JPSEC techn. example	19.72 MB/s
Zhu	17.24 MB/s
Fang and Sun	14.67 MB/s

The application of the iterative encryption algorithm (see Sect. 5.3.5) is not feasible in general as its complexity depends on the plaintext length. At a plaintext length of 4,000 bytes the iterative encryption algorithm achieves a throughput of only 0.07 MB/s.

Another important aspect is memory consumption. All but three algorithms (the iterative encryption algorithm by Wu and Deng, the algorithm by Zhu et al., and the block cipher based algorithm by Wu and Ma) are capable of stream processing (requiring only a few state variables), while the memory consumption of these three algorithms increases linearly with the plaintext length.

#### 5.4 Format-compliant packet header encryption

In Sect. 3.1.1 we discussed the structure of the JPEG2000 packet headers. These contain crucial (even visual) information of the source image. Especially for high-resolution images, content security/confidentiality can not be met without encrypting the leading zero bitplane information in the packet headers (see Figs. 6, 7).

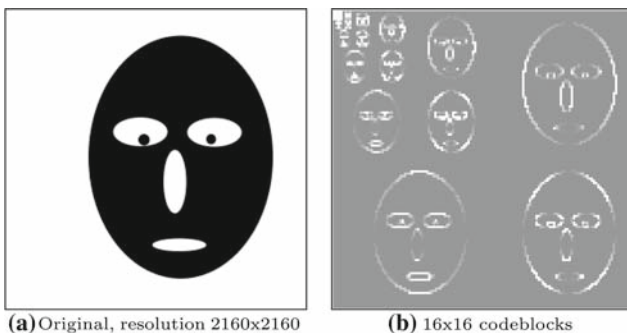


Fig. 6 The LZB information of a high resolution image

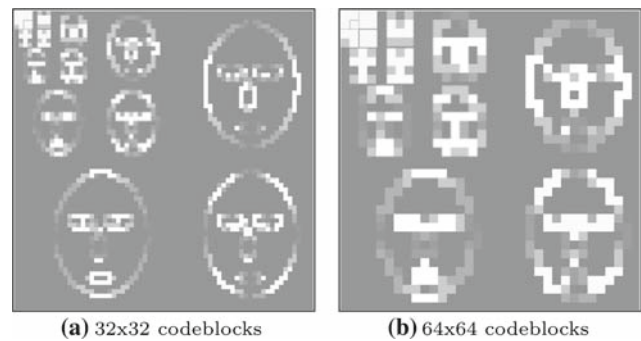


Fig. 7 The LZB information of a high resolution image

In [15], Engel et al. propose format-compliant transformations for each piece of information contained in the packet header. These transformations make use of a random key-stream, the knowledge of which allows the decoder to obtain the original packet header. The resulting codestream is format-compliant.

*CCP Lengths and Number of Coding Passes:* JPEG2000 explicitly signals both the number of coding passes and the length of each codeblock contribution.

The algorithm described in [15] redistributes lengths and coding passes among the codeblocks in a packet. The procedure in pseudo-code is given below.  $v[]$  is a vector of non-zero positive integers (indexing starts at 1).

```

shuffle (v)
borders = size(v) - 1
For i = 1 to borders
    sum = v[i] + v[i + 1]
    r = [random(0,1)*sum]
    newBorder = ((v[i] + r) mod (sum-1)) + 1;
    v[i] = newBorder;
    v[i+1] := sum - newBorder
shuffle(v)
    
```

The transformation can be reversed easily by unshuffling the input, traversing it from end to start, using the random numbers in reverse order, setting newBorder as:

```

newBorder := (v[1] - r - 1) mod (sum - 1)
if (newBorder ≤ 0)
    then newBorder = newBorder + (sum - 1)
    
```

and finally unshuffling the result again.

*Leading Zero Bitplanes:* The number of leading zero bitplanes (LZB) for each codeblock is coded by using tag trees [56]. As discussed above, this information is even more

critical than the other classes of header information, as by using the number of LZB an attacker can obtain information on the visual content of the encrypted image (for small codeblock sizes or high resolutions).

In [15] a random byte is added to the number of leading zero bitplanes modulo a previously determined maximum number. For decoding, the random byte is subtracted instead of added. The maximum number of skipped bitplanes needs to be signaled to the decoder, e.g., by inserting it into the key or by prior arrangement.

**Inclusion Information:** Each packet contains the inclusion information for a certain quality layer for all codeblocks in a precinct. There are four types of inclusions that a codeblock  $c$  can have in packet  $p$ .

The sequence of inclusion information of each codeblock is coded depending on the type of inclusion.

In [15], an algorithm is presented that allows to permute inclusion information for each packet in such a way that the original inclusion information cannot be derived without the key and that the resulting “faked” total inclusion information complies with the semantics of JPEG2000.

**Combined Format-Compliant Header Transformation:** The format-compliant transformation of the different pieces of information in the packet headers can be combined. The format compliance of the combined format-compliant header encryption has been verified experimentally by decoding the encrypted codestreams with the reference implementations JasPer and JJ2000.

**Compression** There is basically no influence on compression performance.

**Security** The visual information contained in LZB information is effectively encrypted (content security/confidentiality can be achieved). Even if packet body based encryption and format-compliant header encryption are combined, security under IND-CPA can still not be achieved as the packet borders are preserved.

**Performance** As packet header data is only a small fraction of the actual codestream, format-compliant header encryption only introduces a small overhead.

## 5.5 Application of bitstream-oriented encryption

Bitstream-oriented JPEG2000 encryption is capable of meeting most of the quality and security constraints of the different applications (cf. Sect. 2.1).

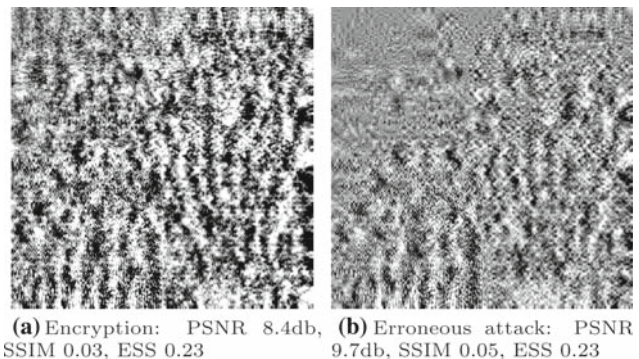
Content security/confidentiality can be achieved by encrypting all of the packet body and packet header data (this can be done format-compliantly).

Sufficient encryption can be achieved by encrypting the packet body data. Depending on the desired level of protection, partial/selective application of bitstream-oriented is feasible.

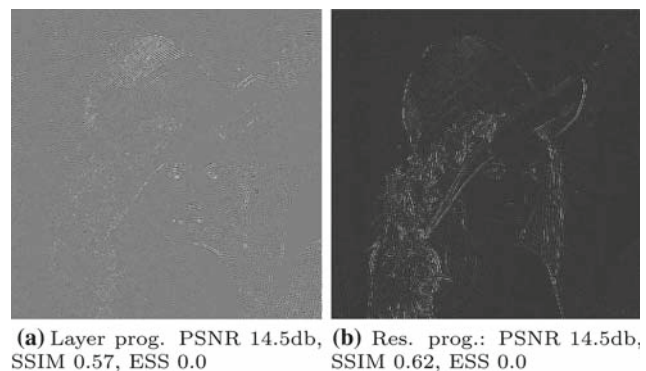
Transparent/perceptual encryption is feasible as well, via the partial/selective application of format-compliant bitstream-oriented schemes.

Only the highest level of security cannot be achieved, as certain properties of the image (e.g., its compressibility, truncation points, ...) will always be preserved.

Norcen and Uhl evaluate JPEG2000 bitstream-oriented encryption for content security/confidentiality [45] and show that it is sufficient to encrypt the first 20% of the JPEG2000 codestream (lossy at a rate of 0.25 or lossless) in order to confidentially hide all image information. Figure 8 shows the direct reconstruction (i.e., a reconstruction with the encrypted parts in place) and the corresponding erroneous error-concealment attack (without the bug-fix mentioned in Sect. 5.1). However, if correct error concealment (i.e., a successful attack) is applied, it turns out that the rule of thumb does not hold anymore, as illustrated in Fig. 9 for both layer and resolution progression. We observe that the SSIM index is capable of measuring the similarity even for very low quality images in contrast to the PSNR and the ESS (see Figs. 8, 9). These results also relativize the claim of data



**Fig. 8** Confidentiality with JPEG2000: 20% encrypted



**Fig. 9** Concealment attack: 20% encrypted, 2bpp

confidentiality for the technology example of Annex B.10 [32, p. 85] (in this example only 1% of JPEG2000 data is encrypted). In [54] Stütz et al. give a more detailed examination of this topic, where they conclude that partial/selective encryption of JPEG2000 cannot guarantee confidentiality.

Transparent/perceptual encryption via bitstream-oriented JPEG2000 encryption has been evaluated by Obermaier and Uhl [59]. The packet body data is encrypted starting from a certain position in the codestream up to the end. This procedure allows the reconstruction of a low quality version from the encrypted codestream. In [59] the impact of the choice for the start of encryption is evaluated for different progression orders, namely resolution and layer progression. A drawback of their approach is that most (more than 90%) of the JPEG2000 codestream has to be encrypted.

More efficient solutions both in terms of computational complexity and reduced deployment cost have been proposed by Stütz and Uhl [53]. Their proposed scheme optimizes the quality of the publicly available low quality version by employing JPEG2000 error concealment strategies and encrypts only a small fraction of the JPEG2000 codestream, namely 1–5%. As a consequence the gap in image quality between the publicly available low quality version and a possible attack is reduced.

## 6 Compression-integrated techniques

Numerous and diverse compression-integrated techniques have been proposed. Encryption in the compression pipeline can be viewed as a compression option (which is kept secret). All considered compression options are not covered in JPEG2000 Part 1, thus their application leads to an encrypted stream not format-compliant with respect to JPEG2000 Part 1.

A major difference among compression-integrated approaches is whether they can be implemented with compliant encoders and decoders. The application of compliant compression software/hardware is an advantage for the practical application of a compression-integrated encryption scheme. Thus the discussion on compression-integrated techniques is divided into two sections, the first discussing techniques that can be implemented with standard compression options, while the second presents various approaches that can only be implemented with non-standard options and with non-standard compression tools.

### 6.1 Secret standard compression options

The following two approaches aim at using the degrees of freedom in the wavelet transform to construct a unique wavelet domain for the transformation step. By keeping the wavelet domain secret, these approaches provide lightweight

security. This procedure can be seen as a form of header encryption, as only the information pertaining to the wavelet domain needs to be encrypted, the rest of the data remains in plaintext. In order to use secret transform domains, Part 2 of the JPEG2000 standard has to be employed. Therefore, a codec that is compliant to JPEG2000 Part 2, is required for encoding and also for decoding of the image in full quality. However, for transparent encryption, a codec compliant to JPEG2000 Part 1, is sufficient to decode the preview image.

#### 6.1.1 Key-dependent wavelet packet subband structures

The wavelet packet decomposition [64] is a generalization of the pyramidal wavelet decomposition, where recursive decomposition may be applied to any subband and is not restricted to the approximation subband. This results in a large space of possible decomposition structures.

**Isotropic Wavelet Packets (IWP)** Pommer and Uhl [47, 48] propose the use of wavelet packets for providing confidentiality in a zerotree-based wavelet framework. Wavelet packet decompositions are created randomly and kept secret. Engel and Uhl [20] transfer the idea and the central algorithm to JPEG2000 and adapt it to support transparent encryption. The aim for a lightweight encryption scheme with wavelet packets is the definition of a large set of possible bases that perform reasonably well at compression. The process that randomly selects one of the bases from this set should operate in a way that does not give a potential attacker any advantage in an attack. To provide these properties, the construction process is controlled by several parameters, e.g., maximal decomposition depth of certain subbands.

To provide transparent encryption, an additional parameter  $p$  is introduced that can be used to optionally specify the number of higher pyramidal resolution levels. If  $p$  is set to a value greater than zero, the pyramidal wavelet decomposition is used for resolution levels  $R_0$  through  $R_p$ . Non-pyramidal wavelet packets are used for the higher resolution levels, starting from  $R_{p+1}$ . With resolution-layer progressions in the final codestream, standard JPEG2000 Part 1 codecs can be used to obtain resolutions  $R_0$  to  $R_p$ .

**Anisotropic Wavelet Packets (AWP)** For the isotropic wavelet packet transform horizontal and vertical decomposition can only occur in pairs. In the anisotropic case this restriction is lifted. The main motivation to introduce anisotropic wavelet packets for lightweight encryption is a substantial increase in keyspace size: the space of possible bases is not only spanned by the decision of decomposing or not, but also by the direction of each decomposition. The amount of data that needs to be encrypted remains extremely small. The complexity of the anisotropic wavelet packet transform is the same as the complexity of the isotropic wavelet packet trans-

form. Like in the isotropic case, compression performance and keyspace size need to be evaluated.

The method for generating randomized wavelet packets has been extended for the anisotropic case by Engel and Uhl [19]. The parameters used to control the generation differ from the isotropic case to reflect the properties of the anisotropic wavelet packet transform. Most notably, the maximum degree of anisotropy is restricted to prevent excessive decomposition into a single direction, as, especially in the case of the approximation subband, this would lead to inferior energy compaction in the wavelet domain for the other direction.

**Compression** For suitable parameter settings (which facilitate energy compaction, see [19–21]), the average compression performance of the wavelet packet transform is comparable to the performance of the pyramidal wavelet transform.

**Security** There are two groups of attacks to consider: attacks that try to determine the wavelet packet structure used for encoding, and attacks that try to (partially) reconstruct the transformed image data without knowing the wavelet packet structure.

*Reconstruction of Decomposition Structure:* Possible attacks that try to determine the wavelet packet structure used for encoding are (a) breaking the cipher with which the decomposition structure was encrypted, (b) inferring the wavelet packet structure from statistical properties of the wavelet coefficients, (c) inferring the wavelet packet structure from the codestream, or (d) performing a full search.

The feasibility of attack (a) is equivalent to the feasibility of breaking the used cipher. Attack (b), inferring the decomposition structure from the codestream tries to use the inclusion metadata in the JPEG2000 codestream. JPEG2000 employs so-called tag trees [55] to signal inclusion information: In a highly contextualized coding scheme, the contributions of each codeblock contained in a packet are linked to the subband structure. Thereby the subband structure is used as context to interpret the output of the tag trees. In order to gather information on either subband structure or coefficients an attacker would have to make a large number of assumptions. However, there are cases (e.g., few quality layers combined with use of markers for packet boundaries) for which fewer possibilities exist and an attacker will have a higher chance of deciphering (some of) the headers. To prevent information leakage, the headers can be encrypted (at the cost of additional computational complexity).

The feasibility of attack (c) is linked to attack (b). If the subband decomposition structure is unknown, the attacker has no way of correctly associating the contributions of a codeblock to the correct coefficients. The attacker therefore lacks full access to the coefficient data (partial access is possible though, see below).

The feasibility of attack (d) depends on the size of the keyspace, which is the number of wavelet packet bases for the used parameters. The number of isotropic wavelet packet bases up to a certain decomposition depth  $j$  can be determined recursively, as shown by Xu and Do [69]. Based on this formula, Engel and Uhl [19, 21] determine the number of isotropic and anisotropic bases of decomposition level up to  $j$ , recursively.

For both, isotropic and anisotropic wavelet packet decompositions, the number of bases obtained with practical parameter settings (i.e., already considering restriction imposed by compression quality requirements) lies above the complexity of a brute-force attack against a 256-bit-key AES cipher.

*Partial Reconstruction:* Rather than trying to find the used wavelet packet decomposition structure, an attacker can try to partially decode the available data.

For the lower resolutions this approach is successful, prohibiting the use of secret wavelet packet decompositions for full confidentiality. This is due to the fact that the packets of the lowest resolution of any (isotropic) wavelet packet decomposition are the same as the packets produced by a pyramidal decomposition of the same image.

In contrast to encryption for full confidentiality, in a transparent encryption scheme the accessibility of the lower resolutions  $R_0$  (or up to  $R_p$ ) is desired. Security is only required for the full quality version.

In order to obtain an image of higher quality than  $R_p$ , an attacker could try to read a fraction of the coefficient data of  $R_{p+1}$  into the pyramidal structure and then attempt a full resolution reconstruction. However, typically the intersection of the randomly generated decomposition structures and the pyramidal structure is far too small to obtain data that allows reconstruction at a substantial quality gain (compared to  $R_p$ ).

When trying to reconstruct the full quality image, the attacker's problem is how to associate packet data with codeblocks, i.e., spatial location. Again it is the highly contextual coding of JPEG2000 that makes it computationally infeasible for the attacker to correctly perform this association. Engel et al. [16] discuss this issue in more detail.

**Performance** Wavelet packets bring an increase in complexity as compared to the pyramidal wavelet decomposition: The order of complexity for a level  $l$  full wavelet packet decomposition of an image of size  $N^2$  is  $\mathcal{O}\left(\sum_{i=1}^l 2^{2(i-1)} \frac{N^2}{2^{2(i-1)}}\right)$  compared to  $\mathcal{O}\left(\sum_{i=1}^l \frac{N^2}{2^{2(i-1)}}\right)$  for the pyramidal decomposition, with the randomized wavelet packet decompositions ranging in-between. With the parameters used in our empirical tests the average time needed for the transform stage increased by 45% as compared to the pyramidal



transform. The average time taken for the whole compression pipeline increased by 25%.

The anisotropic wavelet packet transform does not increase complexity compared to the isotropic case. As more bases can be constructed with lower decomposition depths, the use of the anisotropic wavelet packet transform lowers the computational demands of the scheme.

In general, wavelet packets dramatically reduce the effort for encryption compared to full encryption and other partial or selective encryption schemes. This circumstance makes encryption with a public key scheme feasible, which reduces the effort for key management considerably.

However, the considerable computational complexity that is introduced for the transform step needs to be taken into account for potential application scenarios. For some application scenarios the decrease of complexity in the encryption stage might not suffice to justify the increase of complexity in the compression stage.

### 6.1.2 Parameterized lifting schemes

Three wavelet parameterization schemes have been investigated in the context of lightweight encryption: the parameterization for a family of orthogonal wavelets proposed by Schneid and Pittner [50], the parameterization for even and odd length biorthogonal filters proposed by Hartenstein et al. [26], and the lifting parameterization of the CDF 9/7 wavelet proposed by Zhong and Jiao [71]. Köckerbauer and Uhl [36] report that in the context of JPEG2000 the first parameterization produces unreliable compression results.

Engel and Uhl [17] use the biorthogonal lifting parameterization presented by Zhong and Jiao [71] with JPEG2000 and report compression performance that is superior to the other parameterization schemes. The used parameterization constructs derivations of the original CDF 9/7 wavelet based on a single parameter  $\alpha$ .

**Compression** Compression performance of the produced filters and their utility for JPEG2000 lightweight encryption are investigated in [17]. The tests show that the range in which the parameterized filters achieve good compression results on the one hand and exhibit sufficient variation to withstand a brute-force attack is rather limited. In [18], Engel and Uhl argue that, because filters vary much more for lower absolute values of  $\alpha$ , discretization bins should not be uniform. In order to enlarge the keyspace, they further propose to use different parameters for the horizontal and the vertical wavelet decomposition on different decomposition levels. These techniques have been called “non-stationary” (varying on each decomposition level) and “inhomogeneous” (varying in vertical and horizontal orientation) in the context of adaptive compression [58]. Neither of these methods results in a significant deterioration of compression performance.

The distortion introduced by this scheme is a loss of luminance information rather than a loss of structural information. Figure 10 shows some examples of reconstructed version of the Lena image encrypted with parameter  $\alpha = 2.5$ : (a) shows the image reconstructed with the correct parameter, (b), (c) and (d) show the image reconstructed with incorrect parameters.

**Security Brute-force:** It is reported in [18] that for keys of small individual values for all parameters in each direction and each level a brute-force search for the full-quality version remains unsuccessful. However, keys exist where each parameter is of higher absolute value for which the brute-force attack comes close to the full quality version.

**Symbolic Attack:** A principal attack on parameterized lifting schemes is presented by Engel et al. in [14]. It is based on the symbolic computation of the inverse wavelet transform.

An attacker, who does not know the parameter values for the parameterized transform, can build a symbolic expression for each pixel value in the reconstructed image containing the necessary operations for the inverse transformation.



**Fig. 10** Parameterized wavelet filters: reconstructed images and quality measure results for the Lena image ( $\alpha_{\text{enc}} = -2.5$ ), rate 1 bpp



The resulting term will depend on the transform coefficients, which are known to the attacker. The only unknowns are formed by the parameters of the transform. By performing a full symbolic inverse wavelet transformation, the attacker can construct a complete symbolic description of the operations necessary to reconstruct the plaintext image.

A ciphertext only attack in this context remains largely unsuccessful. This is due to the lack of a reliable non-reference image quality metric.

Known-plaintext attacks are much more successful. If the full plaintext is known, then the symbolic representation can be used to determine the used parameters. This also works if more parameters are used (as in the case of inhomogeneous and non-stationary variation).

Also if only partial plaintext information is available, the symbolic representation yields successful attacks.

Engel et al. [14] discuss two possible scenarios in this context: For the pixel samples attack, the attacker is assumed to have obtained individual pixel samples from the reconstructed image; for the average luminance value attack only the average luminance value from a preview image is required. For both cases, the attacker can obtain a more or less accurate solution for the used wavelet parameters.

Inhomogeneous and non-stationary variation as well as higher-dimensional parameterizations of the wavelet transform increase the number of parameters and therefore make the attack more difficult. However, on a principal note, these symbolic attacks show a general problem of lightweight encryption schemes that rely on linear transforms for providing security. Such attacks severely compromise the security of encryption schemes that use a parameterized wavelet transform, even if their claim is to provide only lightweight encryption.

**Performance** The filter parameterization comes at virtually no cost: Apart from five values in the lifting scheme that have to be computed for each used parameter value, no additional complexity is introduced.

## 6.2 Secret non-standard compression options

Many proposals for compression-integrated encryption modify parts of the compression pipeline in a non-standardized fashion.

### 6.2.1 Wavelet coefficient sign encryption

The signs of the wavelet coefficients are scrambled in a number of contributions [9–13], mainly with the goal to preserve privacy. In [11], as well as in [10], flipping the signs of selected coefficients is proposed for “privacy enabling technology for video surveillance”. This scheme may also be

applied selectively in the transform domain, scrambling only parts of the image.

**Compression** There is only a small negative influence on compression performance. According to [10] the bitrate is increased by less than 10% bitrate (i.e., less than 1dB for a wide range of compression ratios).

**Security** The pseudo-random flipping of wavelet coefficient signs may be subject to specific cryptanalysis. In [49] Said shows the insecurity of DCT sign encryption; however, he uses strong assumptions for his cryptanalytic framework as well as for his attack.

**Performance** The introduced overhead is negligible [10].

### 6.2.2 Random permutations

Norcen and Uhl [43,44] have investigated the usage of random permutations applied to wavelet coefficients within the JPEG2000 coding pipeline. Both confidential and transparent encryption can be implemented by applying permutations to the appropriate subbands, however, one has to keep the inherent security concerns regarding permutations in mind (e.g., vulnerability against known plaintext attacks).

Norcen and Uhl [43] have investigated the permutation of single coefficients within wavelet subbands. In this approach the compression performance is degraded significantly, because the intra-subband dependencies of the coefficients are destroyed. They show that a key generation algorithm has to be employed, since the direct embedding of the permutation key is not feasible from a compression point of view.

In later work [44] aim at improving the rate-distortion performance of permutation based schemes by permuting and rotating differently sized blocks of coefficients (instead of single coefficients) within wavelet subbands. The best compromise with respect to the tradeoff between compression performance and security turns out to be the *blockwise-fully-adaptive* scheme where each subband is divided into the same number of blocks (e.g., 64) which are then permuted. Additionally to the permutation on a block basis, the blocks can be rotated, which increases the keyspace but does not influence compression quality.

**Compression** Compression performance may suffer from the destruction of coefficient statistics and cross correlations through permutation and rotation. In [43] it has been shown that permutation applied to single coefficients severely reduces the compression performance (up to 35%). Schemes applying permutations to blocks of coefficients have been found to be more suited with respect to compression quality [44]—the image quality is augmented with increasing block-size, however, security is decreased with increasing blocksize

(see below). In the blockwise-fully-adaptive scheme compression performance loss can be kept below 10%.

**Security** The security of the presented permutation schemes strongly relies on the blocksize used. Basically there are  $n!$  permutations of  $n$  elements, in this case, coefficients or blocks of coefficients. The keyspace of a specific subband thus is  $n!$  where  $n$  is the number of its blocks (or single coefficients). The whole keyspace is the product of the keyspaces of all subbands. For the block-based permutation the keyspace for a certain subband is  $(width \times height / blocksize^2)!$ . If additionally a random rotation is applied, then the keyspace of a certain subband is  $4^b b!$ , where  $b$  is the number of blocks. For the blockwise-full-adaptive case each subband has  $64!$  different permutations. If random rotation is applied, this number is increased to  $4^{64} 64!$  (except if the remaining block consists only of one coefficient).

In fact not all of the blocks are different (in the high frequency subbands zero coefficients are very likely). If  $k$  blocks are similar, then the number of permutations is decreased by  $k!$  For the blockwise-full-adaptive scheme for 15 similar blocks there are still  $64!/15! = 2 \times 10^{26}$  possible permutations.

However, the security of a system is not entirely determined by its keyspace. Keeping in mind that the lower subbands contain the visually most important information, it has to be pointed out that those are naturally secured by a smaller keyspace.

Moreover, the actual strength of the permutation approach is reduced since correlations among neighboring block borders can be exploited. Hence the bigger the blocks, the less secure the scheme.

An image with permuted  $16 \times 16$  blocks reveals a considerable amount of image information, mostly due to the fact that the lowest resolution subband is not modified at all (it contains exactly one  $16 \times 16$  block). In general the problem with fixed size permutations is that the visually more important subbands are not better secured, hence the full keyspace is a wrong assumption, because an attacker might be able to deduce information from the lower subbands without even considering the higher frequency parts. This is especially true if the block sizes are in excess of  $16 \times 16$ , which mostly leads to unencrypted low frequency subbands.

In the blockwise-fully-adaptive scheme, the number of blocks can be adjusted to a certain security level and the problems with fixed sized blocks are resolved.

A permutation per coefficient destroys all block correlations and can therefore be considered the most secure type of permutation. As a consequence there is a trade-off between security and compression performance.

Another important aspect is information leakage. Since the wavelet coefficients are not changed with this encryption scheme, a simple comparison between the coefficients of an

assumed plaintext image and that of the encrypted image will reveal its identity.

**Performance** The entire compression pipeline has to be run through, but the additional effort is negligibly small (according to our experimental tests).

### 6.2.3 Mixed perturbations

Lian et al. [37, 38] propose the combination of several compression-integrated encryption schemes, such as sign encryption of the wavelet coefficients, inter block permutation and bitplane permutation. Additionally they introduce a parameter  $q$ , the quality factor ranging from 0 to 100, to adjust the encryption strength and the actual image quality of a reconstruction. Hence their scheme may be employed to implement transparent encryption among other application scenarios. In more detail, the quality factor determines the percentage of coefficients for which sign encryption is conducted (for a quality factor of 0 the signs of all coefficients are encrypted, while for a quality factor of 100 no sign is encrypted), the number of intra-permuted codeblocks (bitplane permutation) and a boolean decision whether inter block permutation is employed (which is conducted on a codeblock basis). The order in which both codeblocks and coefficients are treated is from high frequency to low frequency and thus the quality decreases rather smoothly with a decrease of the quality factor.

**Compression** The compression ratio is reduced. An example is given where the degradation is less than 1.5dB for all bitrates in [37]. To put the loss of compression performance into context, JPEG2000 outperforms JPEG by about 2.5dB (PSNR) for a wide range of compression ratios (for the well-known Lena image of size  $512 \times 512$  pixels).

**Security** There are no known attacks against this scheme; however, every single perturbation may be subject to specific cryptanalysis.

**Performance** For a quality factor of 0 (lowest quality) the encryption process takes 7.5–13.2% (as reported in [37]) of the compression (details about the applied software and parameters are not known).

### 6.2.4 Randomized arithmetic coding (RAC)

Grangetto et al. [24] propose JPEG2000 encryption by randomized arithmetic coding. Although the arithmetic coder of the JPEG2000 pipeline is altered, their approach has no influence on the compression performance. The basic idea of their approach is to change the order of the probability intervals in the arithmetic coding process. For the partitioning of

the probability interval, it is a convention (agreed upon by both the encoder and the decoder) which interval (either that of the most probable or that of the least probable symbol) is the preceding one. In [24], for every encoded decision bit the ordering of the intervals is chosen securely randomly (by using a random bit from the PRNG).

Selective/partial application of this encryption approach is possible.

**Compression** There is no influence on compression performance.

**Security** Packet header information is left unencrypted and thus the same considerations as for packet body based format-compliant encryption schemes apply (cf. Sects. 5.3.8). An entire section in [24] is dedicated to the cryptanalysis of their method. It is noted that their method might be susceptible to known-plaintext attacks, but it is argued that these kinds of attacks are not relevant for the proposed encryption systems. A possible counter-argument to this assumption is that, as codeblocks in higher frequency subbands tend to be quantized to zero, it is likely that compressed codeblock contributions of higher frequencies represent bitplanes with a vast majority of zeros.

Due to performance issues, Grangetto et al. propose the usage of a weaker PRNG (with a 32 bit key) based on the standard rand function of the Linux C library. The analysis of the security of this PRNG is out of the scope of this paper, but it can be considered a possible vulnerability. A key size of 32 bit is too short for serious security anyway. Alternatively, the secure random number generator proposed in [3] is employed. However, more secure and efficient PRNG can be considered, e.g., AES in OFB mode.

**Performance** The entire compression pipeline has to be run through, the additional effort arises from the intensive usage of the PRNG. For every decision bit (one per coefficient and per bitplane) coded in the arithmetic coder, a random bit is required. This amount of randomness (basically the same as for raw encryption) induces the authors to employ a faster random number generator (encryption time of 0.33 s for the Lena image with  $512 \times 512$  pixel). Using a secure random number generator [3], the authors report an encryption time of 370.01 s for the full encryption of the Lena image with  $512 \times 512$  pixel. However, the usage of such a computationally complex PRNG is not justified and instead AES in OFB mode could be used as PRNG. Employing our implementation of the AES OFB mode, the generation of the pseudo random keystream of the appropriate length for the Lena image with  $512 \times 512$  pixel only takes 0.045 s. Thus even for secure settings the increase in complexity is not that exorbitant.

However, the computational complexity of this approach is high.

### 6.2.5 Secret initial tables in the MQ-coder

This approach has been proposed by Liu [39] and as in the previously discussed approach of RAC (see Sect. 6.2.4), the entropy coding stage is modified. The arithmetic coding engine (the MQ-coder) receives a context label and a decision (MPS, more probable symbol or LPS, less probable symbol). There are 19 context labels in JPEG2000. The estimation of current interval size for a context is conducted via a finite state machine with 47 states. At the start of entropy coding, each context label is assigned an initial state [34, p. 89]. Liu proposes to randomly select these initial states in order to prevent standard JPEG2000 decoders from correctly decoding the data.

Like the RAC approach, this approach is closely related to packet body encryption with bitstream-compliant algorithms. Selective/partial application of this encryption approach is possible as well.

**Compression** According to [39] the compression overhead is negligible (compressibility equivalent).

**Security** Packet header information is left unencrypted and thus the same considerations as for packet body based format-compliant encryption schemes apply (cf. Sects. 5.3.8).

According to [39] the approach is computationally secure as there are  $47^{19}$  (approx.  $2^{105.5}$ ) possible initial tables. However, a huge key space may not prevent specifically tailored attacks against this scheme.

**Performance** The computational complexity remains almost the same as for the standard JPEG2000 Part 1 compression pipeline; the only effort is to build the random initial table.

## 7 Discussion and overview

In this Section, we will provide a general discussion on which techniques are appropriate for the different application scenarios discussed in the introduction.

The naive encryption technique, i.e., encrypting the entire JPEG2000 codestream with a classical cipher, of course achieves a higher data throughput as compared to all format-compliant bitstream-oriented techniques at the highest level of security. Additionally, information leakage occurs neither for header information nor for packet data. Therefore, if format compliance and all associated functionalities that rely on the JPEG2000 codestream structure are not an issue in the target application, naive encryption is the method

of choice (e.g., as in the DCI security scheme discussed in Sect. 3.6.1).

The discussed format-compliant bitstream-oriented techniques can meet the demands of both on-line and off-line scenarios. Furthermore, an almost arbitrary range of confidentiality levels may be supported by employing partial/selective encryption, ranging from transparent/perceptual encryption where even a certain quality of the visual data in encrypted form has to be guaranteed, to sufficient encryption where strict security is not the major goal but only pleasant viewing has to be impossible. Of course, also high security scenarios may be supported by simply encrypting all the packet data and/or even packet header data. These facts taken together make format-compliant bitstream-oriented encryption techniques the most flexible and most generally applicable schemes discussed.

Considering all approaches including segment based encryption, the KLV approach of the DCI standard and of course all format-compliant encryption schemes one has to mention that a small fingerprint of the image (the compressed size) is preserved by all. For a single image this information can be regarded as insignificant, however, a series of these fingerprints, e.g., obtained from an encrypted movie, identifies the source data in a rather unique way. (Of course the identification only works, if the rate is dynamically adjusted.)

Compression-integrated techniques can only be applied in a sensible way in on-line application scenarios. When compared to bitstream-oriented techniques, the computational demand for encryption is significantly reduced, however, the reduction of complexity for encryption comes at the cost of a (more or less significant) rise in complexity in the compression pipeline. In all schemes considered the impact on compression performance can be kept to a minimum if applied with care.

In a certain sense, the use of key-dependent wavelet transforms in encryption is an extreme case of selective/partial encryption since encryption is limited to the actual subband structure/filter choice. The corresponding amount of data to be encrypted is so small, that this approach can directly employ public key encryption schemes and benefit from their superior key management. Another advantage is that because even though the coefficient data cannot be interpreted without the correct transform at hand, it can be used to perform signal processing in the encrypted domain. For example, this can be useful in a retrieval-based application for creating hashes of encrypted visual data, which facilitates search in the encrypted domain. With respect to the level of confidentiality that can be supported, both wavelet packet and parameterized filters based schemes are found to be restricted to the transparent/perceptual (potentially also to the sufficient) encryption application scenario—real privacy cannot be achieved. Whereas the increase in complexity that is shifted to the compression pipeline can be considered significant for the

wavelet packet case, in the case of parameterized filters there is only negligible additional cost. The successful attacks against the approach based on parameterized filters renders this technique almost useless in environments where sincere attacks are to be expected. At most, in settings that require soft encryption, e.g., in the area of mobile multimedia applications, the level of security might suffice and the extremely low computational demands could be an incentive for using parameterized wavelet filters. It has to be pointed out that the low amount of encryption required for transparent encryption in key-dependent transform techniques makes those especially attractive since the classical approach for transparent encryption in bitstream-oriented techniques requires almost the entire codestream to be encrypted (compare Sect. 5.5). However, more recent techniques [53] only require a fraction of the encryption amounts. Engel et al. [16] discuss various application scenarios for transparent encryption where one or the other approach might be of advantage.

Finally, employing permutations within the JPEG2000 pipeline is a classical case of soft encryption. The computational overhead remains negligible, however, permutations are of course vulnerable to known plaintext attacks unless the keys are exchanged frequently. Contrasting to the previous techniques also higher levels of confidentiality may be targeted (this just depends on which subbands are subject to permutations), however, the security flaws present with permutations should be kept in mind.

Information leakage is significantly higher in compression-integrated techniques as compared to bitstream-oriented ones. The entire coefficient information is available in plaintext—due to the missing context this cannot be exploited for reconstructing the original data, but all sorts of statistical analyses may be conducted on these data potentially allowing an attacker to identify an image in encrypted form. Therefore, the security level of for these compression-integrated schemes has to be assessed to be lower as compared to bitstream-oriented ones in general.

A completely different approach is randomized arithmetic coding as proposed by Grangetto et al. It is closely linked to the bitstream-compliant encryption approaches discussed in Sect. 5 as it targets the coefficient data contained in the packet bodies. However, the drawbacks of this solution are the increased complexity compared to bitstream-compliant approaches. If bitstream-compliant approaches are applied on a CCP basis there is no difference in the preserved functionality (given the appropriate key management for both schemes). Thus the randomized arithmetic coding approach has the same functionality as the bitstream-compliant approaches, but obvious disadvantages. Secret initial tables may be an interesting option; however, the security of this approach against specifically tailored attacks remains to be proven. The selective/partial application cannot gain substantial performance gains, as the encryption only takes a



**Table 1** An overview of JPEG2000 encryption approaches

Approach	Naive	Segment-based	Bitstream-compliant	Wavelet packets	Filters	Permutations	RAC
Compression decrease	None	Slightly	None	Moderately	Slightly	Moderately	None
Confidentiality	Privacy	All levels (ZOI)	All levels (ZOI)	Transparent	Transparent	All levels (ZOI)	All levels (ZOI)
Security	Very high	High–very high	High	Medium–high	Low	Medium	Medium–high
Transcodability	None	Partial, segment based	On packet basis	JPEG2000	JPEG2000	JPEG2000	JPEG2000
Transcode complexity	Very high	Very low	Very low with markers	JPEG2000	JPEG2000	JPEG2000	JPEG2000
Create complexity	Low	Low–medium	Low	High*	Very low–low*	Very low*	High*
Compression pipeline	Full	Tier2	No	Full	Full	Full	Full
Consume complexity	Low	Low–medium	Low	High	Very low–low	Very low	Medium
Uninformed consume	Impossible	Impossible	Possible	Possible	Possible	Possible	Possible
Error propagation	Avalanche eff.	Within segment	Mostly JPEG2000	JPEG2000	JPEG2000	JPEG2000	JPEG2000

negligible fraction of the entire compression and encryption system (cf. Sect. 5.3.8).

In Table 1, we provide a concise summary of the various aspects discussed in this and the preceding Sections. A “\*” indicates that the entire compression pipeline has to be conducted and the given information is with respect to the additional effort within the compression pipeline.

## 8 Conclusion

In this survey we have discussed and compared various techniques for protecting JPEG2000 codestreams by encryption technology. As to be expected, some techniques turn out to be more beneficial than others and some methods hardly seem to make sense in any application context. In any case, a large variety of approaches exhibiting very different properties can be considered useful and covers almost any thinkable multimedia application scenario. This survey provides a guide to find the proper JPEG2000 encryption scheme for a target application.

## References

1. Apostolopoulos, J., Wee, S., Dufaux, F., Ebrahimi, T., Sun, Q., Zhang, Z.: The emerging JPEG2000 security (JPSEC) standard. In: Proceedings of International Symposium on Circuits and Systems, ISCAS'06. IEEE, May 2006
2. Apostolopoulos, J.G., Wee S.J.: Supporting secure transcoding in JPSEC. In: Tescher, A.G. (ed.) Applications of Digital Image Processing XXVIII, vol. 5909, p. 59090J. SPIE (2005)
3. Blum, L., Blum, M., Shub, M.: A simple unpredictable pseudo-random number generator. *SIAM J. Comput.* **15**(2), 364–383 (1986)
4. Bradley, J.N., Brislawn, C.M., Hopper, T.: The FBI wavelet/scalar quantization standard for gray-scale fingerprint image compression. In: SPIE Proceedings, Visual Information Processing II, vol. 1961, pp. 293–304, Orlando, FL, USA, April (1993)
5. Conan, V., Sadourny, Y., Jean-Marie, K., Chan, C., Wee, S., Apostolopoulos, J.: Study and validation of tools interoperability in JPSEC. In: Tescher, A.G. (ed.) Applications of Digital Image Processing XXVIII, vol. 5909, p. 59090H. SPIE (2005)
6. Conan, V., Sadourny, Y., Thomann, S.: Symmetric block cipher based protection: Contribution to JPSEC. ISO/IEC JTC 1/SC 29/WG 1 N 2771 (2003)
7. Daemen, J., Rijmen, V.: The Design of Rijndael: AES—The Advanced Encryption Standard. Springer, Berlin (2002)
8. Digital Cinema Initiatives, LLC (DCI). Digital cinema system specification v1.2. online presentation, March (2008)
9. Dufaux, F., Ebrahimi, T.: Region-based transform-domain video scrambling. In: Proceedings of Visual Communications and Image Processing, VCIP'06. SPIE (2006)
10. Dufaux, F., Ebrahimi, T.: Scrambling for video surveillance with privacy. In Proceedings of the 2006 Conference on Computer Vision and Pattern Recognition Workshop, CVPRW '06. IEEE (2006)
11. Dufaux, F., Ouaret, M., Abdeljaoued, Y., Navarro, A., Vergnenegre, F., Ebrahimi, T.: Privacy enabling technology for video surveillance. In: Proceedings of SPIE, Mobile Multimedia/Image Processing for Military and Security Applications, vol. 6250. SPIE (2006)
12. Dufaux, F., Wee, S., Apostolopoulos J., Ebrahimi T.: JPSEC for secure imaging in JPEG2000. In: Tescher, A.G. (ed.) Applications of Digital Image Processing XXVII, vol. 5558, pp. 319–330. SPIE (2004)
13. Dufaux, F., Ebrahimi, T.: Securing JPEG2000 compressed images. In: Tescher, A.G. (ed.) Applications of Digital Image Processing XXVI, vol. 5203, pp. 397–406. SPIE (2003)
14. Engel, D., Kutil, R., Uhl, A.: A symbolic transform attack on light-weight encryption based on wavelet filter parameterization. In: Proceedings of ACM Multimedia and Security Workshop, MM-SEC '06, pp. 202–207, Geneva, Switzerland, September (2006)
15. Engel, D., Stütz, T., Uhl, A.: Format-compliant JPEG2000 encryption in JPSEC: Security, applicability and the impact of compression parameters. *EURASIP J. Inform. Secur.* (Article ID 94565), 20 (2007). doi:10.1155/2007/94565
16. Engel, D., Stütz, T., Uhl, A.: Efficient transparent JPEG2000 encryption. In: Li, C.-T. (ed.) Multimedia Forensics and Security, pp. 336–359. IGI Global, Hershey, PA, USA (2008)



17. Engel, D., Uhl, A.: Parameterized biorthogonal wavelet lifting for lightweight JPEG2000 transparent encryption. In: Proceedings of ACM Multimedia and Security Workshop, MM-SEC '05, pp. 63–70, New York, NY, USA, August (2005)
18. Engel, D., Uhl, A.: Security enhancement for lightweight JPEG2000 transparent encryption. In: Proceedings of Fifth International Conference on Information, Communication and Signal Processing, ICICS '05, pp. 1102–1106, Bangkok, Thailand, December (2005)
19. Engel, D., Uhl, A.: Lightweight JPEG2000 encryption with anisotropic wavelet packets. In: Proceedings of International Conference on Multimedia and Expo, ICME '06, pp. 2177–2180, Toronto, Canada, July 2006. IEEE (2006)
20. Engel, D., Uhl, A.: Secret wavelet packet decompositions for JPEG2000 lightweight encryption. In: Proceedings of 31st International Conference on Acoustics, Speech, and Signal Processing, ICASSP '06, vol. V, pp. 465–468, Toulouse, France, May 2006. IEEE (2006)
21. Engel, D., Uhl, A.: An evaluation of lightweight JPEG2000 encryption with anisotropic wavelet packets. In: Delp, E.J., Wong, P.W. (eds.) Security, Steganography, and Watermarking of Multimedia Contents IX. Proceedings of SPIE, pp. 65051S1–65051S10, San Jose, CA, USA, January 2007. SPIE (2007)
22. Fang, J., Sun, J.: Compliant encryption scheme for JPEG2000 image code streams. *J. Electron. Imaging* **15**(4) (2006)
23. Furht, B., Kirovski, D. (eds.): *Multimedia Security Handbook*. CRC Press, Boca Raton, (2005)
24. Grangetto, M., Magli, E., Olmo, G.: Multimedia selective encryption by means of randomized arithmetic coding. *IEEE Trans. Multimed.* **8**(5), 905–917 (2006)
25. Grosbois, R., Gerbelot, P., Ebrahimi, T.: Authentication and access control in the JPEG2000 compressed domain. In: Tescher, A.G. (ed.) Applications of Digital Image Processing XXIV. Proceedings of SPIE, vol. 4472, pp. 95–104, San Diego, CA, USA, July (2001)
26. Hartenstein, F.: Parametrization of discrete finite biorthogonal wavelets with linear phase. In: Proceedings of the 1997 International Conference on Acoustics, Speech and Signal Processing (ICASSP'97), April (1997)
27. Imaizumi, S., Watanabe, O., Fujiyoshi, M., Kiya, H.: Generalized hierarchical encryption of JPEG2000 codestreams for access control. In: Proceedings of the IEEE International Conference on Image Processing (ICIP'05), vol. 2. IEEE, September (2005)
28. Imaizumi, S., Fujiyoshi, M., Abe, Y., Kiya, H.: Collusion attack-resilient hierarchical encryption of JPEG 2000 codestreams with scalable access control. In: *Image Processing, 2007. ICIP 2007. IEEE International Conference on*, vol. 2, pp. 137–140, September (2007)
29. ISO/IEC 15444-8, Final Committee Draft. Information technology—JPEG2000 image coding system, Part 8: Secure JPEG2000. Technical report, ISO, November (2004)
30. ISO/IEC 15444-1. Information technology—JPEG2000 image coding system, Part 1: Core coding system, December (2000)
31. ISO/IEC 15444-4. Information technology—JPEG2000 image coding system, Part 4: Conformance testing, December (2004)
32. ISO/IEC 15444-8. Information technology—JPEG2000 image coding system, Part 8: Secure JPEG2000, April (2007)
33. ITU-T H.264. Advanced video coding for generic audiovisual services, November (2007)
34. ITU-T T.800. Information technology—JPEG2000 image coding system, Part 1: Core coding system, August (2002)
35. Kiya, H., Imaizumi, D., Watanabe O.: Partial-scrambling of image encoded using JPEG2000 without generating marker codes. In: Proceedings of the IEEE International Conference on Image Processing (ICIP'03), vol. III, pp. 205–208, Barcelona, Spain, September (2003)
36. Köckerbauer, T., Kumar, M., Uhl, A.: Lightweight JPEG2000 confidentiality for mobile environments. In: Proceedings of the IEEE International Conference on Multimedia and Expo, ICME '04, Taipei, Taiwan, June (2004)
37. Lian, S., Sun, J., Wang, Z.: Perceptual cryptography on JPEG2000 compressed images or videos. In: 4th International Conference on Computer and Information Technology, Wuhan, China, September 2004. IEEE (2004)
38. Lian, S., Sun, J., Zhang, D., Wang, Z.: A selective image encryption scheme based on JPEG2000 codec. In: Nakamura, Y., Aizawa, K., Satoh, S. (eds.) Proceedings of the 5th Pacific Rim Conference on Multimedia. Lecture Notes in Computer Science, vol. 3332, pp. 65–72. Springer, Berlin (2004)
39. Liu, J.-L.: Efficient selective encryption for jpeg 2000 images using private initial table. *Pattern Recognit.* **39**(8), 1509–1517 (2006)
40. Lo, S.-C.B., Li, H., Freedman, M.T.: Optimization of wavelet decomposition for image compression and feature preservation. *IEEE Trans. Med. Imaging* **22**(9), 1141–1151 (2003)
41. Macq, B.M., Quisquater, J.-J.: Cryptology for digital TV broadcasting. *Proc. IEEE* **83**(6), 944–957 (1995)
42. Mao, Y., Wu, M.: Security evaluation for communication-friendly encryption of multimedia. In: Proceedings of the IEEE International Conference on Image Processing (ICIP'04), Singapore, October 2004. IEEE Signal Processing Society (2004)
43. Norcen, R., Uhl, A.: Encryption of wavelet-coded imagery using random permutations. In: Proceedings of the IEEE International Conference on Image Processing (ICIP'04), Singapore, October 2004. IEEE Signal Processing Society (2004)
44. Norcen, R., Uhl, A.: Performance analysis of block-based permutations in securing JPEG2000 and SPIHT compression. In: Li, S., Pereira, F., Shum, H.-Y., Tescher, A.G. (eds.) Visual Communications and Image Processing 2005 (VCIP'05). SPIE Proceedings, vol. 5960, pp. 944–952, Beijing, China, July 2005. SPIE (2005)
45. Norcen, R., Uhl, A.: Selective encryption of the JPEG2000 bitstream. In: Lioy, A., Mazzocchi, D. (eds.) Communications and Multimedia Security. Proceedings of the IFIP TC6/TC11 Sixth Joint Working Conference on Communications and Multimedia Security, CMS '03. Lecture Notes on Computer Science, vol. 2828, pp. 194–204, Turin, Italy, October 2003. Springer, Berlin (2003)
46. Pommer, A., Uhl, A.: Application scenarios for selective encryption of visual data. In: Dittmann, J., Fridrich, J., Wohlmacher, P. (eds.) Multimedia and Security Workshop, ACM Multimedia, pp. 71–74, Juan-les-Pins, France, December (2002)
47. Pommer, A., Uhl, A.: Selective encryption of wavelet packet subband structures for secure transmission of visual data. In: Dittmann, J., Fridrich, J., Wohlmacher, P. (eds.) Multimedia and Security Workshop, ACM Multimedia, pp. 67–70, Juan-les-Pins, France, December (2002)
48. Pommer, A., Uhl, A.: Selective encryption of wavelet-packet encoded image data—efficiency and security. *ACM Multimed. Syst. (Special Issue Multimed. Secur.)* **9**(3), 279–287 (2003)
49. Said, A.: Measuring the strength of partial encryption schemes. In: Proceedings of the IEEE International Conference on Image Processing (ICIP'05), vol. 2, September (2005)
50. Schneid, J., Pittner, S.: On the parametrization of the coefficients of dilation equations for compactly supported wavelets. *Computing* **51**, 165–173 (1993)
51. Schneier, B.: *Applied cryptography: protocols, algorithms and source code in C*, 2nd edn. Wiley, New York (1996)
52. Stütz, T., Uhl, A.: On format-compliant iterative encryption of JPEG2000. In: Proceedings of the Eighth IEEE International Symposium on Multimedia (ISM'06), pp. 985–990, San Diego, CA, USA, December 2006. IEEE Computer Society (2006)
53. Stütz, T., Uhl, A.: On efficient transparent JPEG2000 encryption. In Proceedings of ACM Multimedia and Security Workshop,

- MM-SEC '07, pp. 97–108, New York, NY, USA, September 2007. ACM Press
54. Stütz, T., Uhl, A.: On JPEG2000 error concealment attacks. In: Proceedings of the 3rd Pacific-Rim Symposium on Image and Video Technology, PSIVT '09, Lecture Notes in Computer Science, Tokyo, Japan, January 2009. Springer, Berlin (2009, to appear)
  55. Taubman, D.: High performance scalable image compression with EBCOT. *IEEE Trans. Image Process.* **9**(7), 1158–1170 (2000)
  56. Taubman, D., Marcellin, M.W.: *JPEG2000—Image Compression Fundamentals, Standards and Practice*. Kluwer, Dordrecht (2002)
  57. Tolba, A.S.: Wavelet packet compression of medical images. *Digit. Signal Process.* **12**(4), 441–470 (2002)
  58. Uhl, A.: Image compression using non-stationary and inhomogeneous multiresolution analyses. *Image Vis. Comput.* **14**(5), 365–371 (1996)
  59. Uhl, A., Obermair, Ch.: Transparent encryption of JPEG2000 bit-streams. In: Podhradsky, P. et al. (eds.) *Proceedings EC-SIP-M 2005 (5th EURASIP Conference focused on Speech and Image Processing, Multimedia Communications and Services)*, pp 322–327, Smolenice, Slovak Republic (2005)
  60. Uhl, A., Pommer, A.: *Image and Video Encryption. From Digital Rights Management to Secured Personal Communication. Advances in Information Security*, vol. 15. Springer, Berlin (2005)
  61. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.* **13**(4) (2004)
  62. Wee, S., Apostolopoulos, J.: Secure transcoding with JPSEC confidentiality and authentication. In: *Proceedings of the IEEE International Conference on Image Processing (ICIP'04)*, Singapore, Singapore, October (2004)
  63. Wee, S.J., Apostolopoulos, J.G.: Secure scalable streaming and secure transcoding with JPEG2000. In: *Proceedings of the IEEE International Conference on Image Processing (ICIP'03)*, vol. I, pp. 547–551, Barcelona, Spain, September (2003)
  64. Wickerhauser, M.V.: *Adapted Wavelet Analysis from Theory to Software*. A.K. Peters, Wellesley (1994)
  65. Wu, H., Ma, D.: Efficient and secure encryption schemes for JPEG2000. In: *Proceedings of the 2004 International Conference on Acoustics, Speech and Signal Processing (ICASSP 2004)*, pp. 869–872, May (2004)
  66. Wu, M., Mao, V.: Communication-friendly encryption of multimedia. In: *Proceedings of the IEEE Multimedia Signal Processing Workshop, MMSP '02*, St. Thomas, Virgin Islands, USA, December (2002)
  67. Wu, Y., Deng, R., Ma, D.: Securing JPEG2000 code-streams. In: Lee, D., Shieh, S., Tygar, J., (eds.) *Computer Security in the 21st Century*, pp. 229–254 (2005)
  68. Wu, Y., Deng, R.H.: Compliant encryption of JPEG2000 code-streams. In: *Proceedings of the IEEE International Conference on Image Processing (ICIP'04)*, Singapore, October 2004. IEEE Signal Processing Society (2004)
  69. Xu, D., Do, M.N.: Anisotropic 2-D wavelet packets and rectangular tiling: theory and algorithms. In: Unser, M.A., Aldroubi, A., Laine, A.F. (eds.) *Proceedings of SPIE Conference on Wavelet Applications in Signal and Image Processing X. SPIE Proceedings*, vol. 5207, pp. 619–630, San Diego, CA, USA, August 2003. SPIE (2003)
  70. Yang, Y., Zhu, B.B., Yang, Y., Li, S., Yu N.: Efficient and syntax-compliant JPEG2000 encryption preserving original fine granularity of scalability. *EURASIP J. Inform. Secur.* (2007)
  71. Zhong, G., Cheng, L., Chen, H.: A simple 9/7-tap wavelet filter based on lifting scheme. In: *Proceedings of the IEEE International Conference on Image Processing (ICIP'01)*, pp. 249–252, October (2001)
  72. Zhu, B., Yang, Y., Li, S.: JPEG2000 syntax-compliant encryption preserving full scalability. In: *Proceedings of the IEEE International Conference on Image Processing (ICIP'05)*, vol. 3, September (2005)



## COPYRIGHT INFORMATION

TITLE: A survey on JPEG2000 encryption  
SOURCE: Multimedia Syst 15 no4 Ag 2009

The magazine publisher is the copyright holder of this article and it is reproduced with permission. Further reproduction of this article in violation of the copyright is prohibited. To contact the publisher:  
<http://www.springerlink.com/content/1432-1882/>

MTR 04B0000022

---

MITRE TECHNICAL REPORT

# Profile for 1000ppi Fingerprint Compression

## Version 1.1

April 2004

Margaret A. Lepley

**Sponsor:** DOJ / FBI  
**Dept. No.:** G036

**Contract No.:** W15P7T-04-C-D001  
**Project No.:** 0704E02X

Approved for public release; distribution unlimited.

©2004 The MITRE Corporation. All Rights Reserved.

**MITRE**  
Center for Integrated Intelligence Systems  
Bedford, Massachusetts





## **Abstract**

This document specifies a format for use in compressing 1000ppi fingerprints. This format is a profile (usage subset) of the ISO/IEC 15444-1 JPEG 2000 image compression standard. Compliance testing procedures are described for this profile.

**KEYWORDS:** JPEG 2000, JP2, fingerprint compression, 1000 ppi, WSQ, wavelets



# Table of Contents

Section	Page
<b>1 Scope</b> .....	<b>1</b>
<b>2 References</b> .....	<b>1</b>
<b>3 Definitions</b> .....	<b>2</b>
<b>4 Abbreviations and Symbols</b> .....	<b>2</b>
4.1 Abbreviations .....	2
4.2 Symbols.....	2
<b>5 Introduction</b> .....	<b>3</b>
<b>6 JP2 File Format</b> .....	<b>3</b>
6.1 FP JP2 Profile.....	3
<b>7 JPEG 2000 Codestream</b> .....	<b>4</b>
7.1 FP JPEG 2000 Profile .....	4
7.1.1 FP COM Marker Segment.....	5
7.2 FP JPEG 2000 Layers (informative) .....	5
7.3 FP JPEG 2000 Guidance (informative) .....	5
<b>8 Compliance Testing</b> .....	<b>7</b>
8.1 Syntax Tests .....	7
8.2 Visual Confirmation .....	7
8.3 Implementation Tests .....	8
8.3.1 Encoder Compliance Tests.....	8
8.3.2 Decoder Compliance Tests.....	9
8.3.3 Transcoder Compliance Tests .....	10
8.3.3.1 Transcoder Test A .....	10
8.3.3.2 Transcoder Test B .....	10
8.3.4 Test Data .....	11
<b>Appendix A Minimal FP JP2 Example (Informative)</b> .....	<b>13</b>
A.1 JPEG 2000 Signature Box .....	13
A.2 File Type Box.....	14
A.3 JP2 Header Box.....	14
A.3.1 Image Header Box .....	15
A.3.2 Color Specification Box.....	15
A.3.3 Resolution Box .....	16
A.4 Contiguous Codestream Box.....	16

<b>Appendix B Quality Metrics .....</b>	<b>17</b>
B.1 Comparative Image Metrics.....	17
B.1.1 Root Mean Square Error (RMSE) .....	17
B.1.2 Mean Absolute Error.....	17
B.1.3 Absolute Mean Error.....	17
B.2 Single-Image Metrics.....	18
B.2.1 Image Quality Measure (IQM) .....	18
<b>Appendix C Metric Bounds.....</b>	<b>21</b>
C.1 Notation.....	21
C.2 Metric Bounds Tables.....	21

## List of Figures

<b>Figure</b>	<b>Page</b>
Figure A-1. High-level FP JP2 Mandatory Content .....	13
Figure A-2. Organization of the JPEG 2000 Signature Box .....	13
Figure A-3. Contents of the JPEG 2000 Signature Box.....	13
Figure A-4. Organization of the File Type Box.....	14
Figure A-5. Contents of the File Type Box.....	14
Figure A-6. Organization of the JP2 Header Box.....	14
Figure A-7. Contents of the JP2 Header Box .....	14
Figure A-8. Organization of the Image Header Box.....	15
Figure A-9. Contents of the Image Header Box .....	15
Figure A-10. Organization of the Color Specification Box .....	15
Figure A-11. Contents of the Color Specification Box.....	15
Figure A-12. Organization of the Resolution Box.....	16
Figure A-13. Contents of the Resolution Box .....	16
Figure A-14. Organization of the Contiguous Codestream Box .....	16
Figure A-15. Contents of the Contiguous Codestream Box.....	16
Figure B-1. AuxdataFile Content for Image A.....	19
Figure B-2. Example of IQM Interpretation of the AuxdataFile.....	20

## List of Tables

<b>Table</b>	<b>Page</b>
Table 1. Codestream Requirements for FP.....	4
Table 2. Content of FP specified COM marker .....	5
Table 3. JPEG 2000 Parameter Guidance .....	6
Table 4. Sample Metric Bounds for 1000ppi Encoder Test .....	8
Table 5. Sample Metric Bounds for 500ppi Encoder Test.....	9
Table 6. Sample Metric Bounds for Decoder Compliance Test.....	9
Table 7. Sample Metric Bounds for Transcoder Test A .....	10
Table C-1. Metric Bounds for 1000ppi Encoder Test.....	22
Table C-2. Metric Bounds for 500ppi Encoder Test.....	22
Table C-3. Metric Bounds for Decoder Compliance Test.....	22
Table C-4. Metric Bounds for Transcoder Test A .....	23





## 1 Scope

The 1000ppi fingerprint JPEG 2000 profile and required content of the associated JP2 format are described in this document. The purpose for this profile is to:

- insure image quality
- insure interoperability including backward compatibility
- position criminal justice agencies to leverage commercial investment in open COTS solutions

This specification is applicable to 1000ppi continuous-tone gray-scale digital fingerprint images with a bit depth of 8 bits per pixel.

This specification

- specifies a file format for storing and transmitting compressed 1000ppi fingerprint image data
- specifies a class of encoders for converting source 1000ppi fingerprint image data to compressed image data
- specifies a class of decoders for converting compressed image data to reconstructed 1000ppi fingerprint image data
- specifies two classes of transcoders for converting between this compression specification and the FBI's compression spec for 500ppi fingerprints (WSQ)

For brevity, elements of this specification will be labeled as FP (1000ppi Fingerprint compression Profile). For example, references will be made to the FP JPEG 2000 codestream and the FP JP2 format.

All sections of this document are normative, unless explicitly labeled as informative.

## 2 References

The following Recommendations, Specifications and International Standards contain provisions that, through reference in this text, constitute provisions of this Specification.

- 1 ISO/IEC 646:1991, ISO 7-bit coded character set for information interchange.
- 2 ANSI/NIST-ITL 1-2000, NIST Special Publication 500-245, "American National Standard for Information Systems---Data Format for the Interchange of Fingerprint, Facial, & Scar Mark & Tattoo (SMT) Information," 2000.
- 3 Criminal Justice Information Services (CJIS) WSQ Gray-scale Fingerprint Image Compression Specification, Federal Bureau of Investigation document No. IAFIS-IC-0110(V3), 19 Dec 1997.
- 4 ISO/IEC 15444-1:2000, JPEG 2000 Part 1: Image Coding System: Core Coding System.
- 5 ISO/IEC 15444-1:2000-Amd1, JPEG 2000 Part 1: Image Coding System: Core Coding System, Amendment 1.
- 6 ISO/IEC 15444-1:2000-Amd2, JPEG 2000 Part 1: Image Coding System: Core Coding System, Amendment 2.
- 7 ISO/IEC 15444-4:2002, JPEG 2000 Part 4: Image Coding System: Conformance
- 8 <http://www.mitre.org/tech/mtf>, Image Quality Measure (IQM)

### 3 Definitions

For the purposes of this Specification, the definitions shown in 15444-1 [4] Section 3 and the following apply.

**transcoding:** A process that converts one compressed format to another.

### 4 Abbreviations and Symbols

#### 4.1 Abbreviations

For the purposes of this Specification, the following abbreviations apply.

**bpp:** bits per pixel

**IEC:** International Electrotechnical Commission

**IQM:** Image Quality Measure [8]

**ISO:** International Organization for Standardization

**JPEG:** Joint Photographic Experts Group. The joint ISO/ITU committee responsible for developing standards for continuous-tone still picture coding. It also refers to the standards produced by this committee.

**ppi:** pixels per inch

**PCRL:** Position-Component-Resolution-Layer

**RLCP:** Resolution-Layer-Component-Position

**RPCL:** Resolution-Position-Component-Layer

**RMSE:** Root Mean Square Error

**WSQ:** Wavelet Scalar Quantization [3]

#### 4.2 Symbols

For the purposes of this Specification, the following symbols apply

**'xyz':** Denotes a character string, included in a field exactly as shown but without the beginning and ending quote marks.

**0x---:** Denotes a hexadecimal number

**<CR>:** A single carriage return character

**<LF>:** A single linefeed character

**<SP>:** A single space character (0x20 in hexadecimal)

**COM:** Comment marker

**FP:** Profile for 1000ppi Fingerprint Compression

**POC:** Progression order change marker

## 5 Introduction

This specification for storage of 1000ppi fingerprints (or like imagery such as palm- or footprints) is based upon JPEG 2000 compression.

Since JPEG2000 is an extremely broad compression standard, a specific profile for JPEG 2000 in JP2 format has been developed for fingerprint compression. The FP JP2 file can be used as a single file, or encapsulated in an ANSI NIST card file [2]. The FP JPEG 2000 profile and required content of the FP JP2 format are described in Sections 6 and 7.

In addition to the profile restrictions, Section 8 describes a set of compliance tests that ensure a minimal degree of quality for FP JPEG 2000 encoders and decoders. Applications may find it easier to pass the compliance tests if the JPEG 2000 parameter settings used in test development are followed to some extent. Guidelines for JPEG 2000 settings are provided in Section 7.3, but are not a requirement.

As well as testing JPEG 2000 encode/decode capabilities, the compliance suite tests the ability to convert a 1000ppi FP JPEG 2000 compressed file into a 500ppi WSQ file. This conversion, referred to as transcoding, requires an understanding of the WSQ standard [3].

## 6 JP2 File Format

JP2 is a file format that allows meta-data to be packaged with the image data, using a convention called 'boxes'. Each box begins with an indication of box length and type, followed by the contents, which may be data or other boxes. Boxes contain logical groupings of meta-data or a compressed image codestream.

### 6.1 FP JP2 Profile

The JP2 specification [4] mentions mandatory, optional and customizable boxes. The FP JP2 profile increases the list of mandatory boxes to include Capture Resolution.

A FP JP2 file must include the following:

- JPEG 2000 Signature box
- File Type box
- JP2 Header superbox containing:
  - Image Header box
  - Color Specification box (populated to indicate grayscale)
  - Resolution superbox containing Capture Resolution
- Contiguous Codestream box (using the FP JPEG 2000 profile)

Other optional boxes may appear in a FP JP2 file, but the above list is mandatory. An informative example of a minimal FP JP2 file is given in Appendix A.

## 7 JPEG 2000 Codestream

The JPEG 2000 standard [4] is very flexible, but the number of choices provided can be daunting when attempting to encode an image. To help create a useful interoperable system some additional limitations and guidance are provided. The limitations, called the FP JPEG 2000 Profile, are requirements for any 1000ppi fingerprint compression. The guidance, by contrast, is not a requirement but an aid to achieving parameter settings that are known to produce adequate image quality.

### 7.1 FP JPEG 2000 Profile

The FP JPEG 2000 fingerprint profile is an additional restriction within JPEG 2000 Profile 1 as defined in ISO 15444-1 Amd 1 [5]. Table 1 below shows the FP JPEG 2000 Profile Requirements (including Profile 1 limitations).

**Table 1. Codestream Requirements for FP**

<b>Restrictions</b>	<b>1000ppi Fingerprint Profile</b>
<b>Profile 1 Requirements</b>	
<b>Profile Indication</b>	Rsiz = 2 or 1 (minimal value appropriate)
<b>Image Size</b>	Xsiz, Ysiz < 2 <sup>31</sup>
<b>Tiles</b>	Multiple tiles: XTsiz/min(XRsiz <sup>i</sup> , YRsiz <sup>i</sup> ) ≤ 1024 XTsiz=YTsiz Or one tile for the whole image: YTsiz+YTOsiz ≥ Ysiz XTsiz+XTOsiz ≥ Xsiz
<b>Image &amp; tile origin</b>	XOsiz, YOsiz, XTOsiz, YTOsiz < 2 <sup>31</sup>
<b>Code-block size</b>	xcb ≤ 6, ycb ≤ 6
<b>RGN marker segment</b>	SPrgn ≤ 37
<b>Additional FP Requirements</b>	
<b>COM marker</b>	Required COM marker indicating compression software version. See Section 7.1.1
<b>Filter</b>	9-7 irreversible
<b>Levels of Decomposition</b>	6
<b>Number of components</b>	1 for grayscale
<b>Number of Layers</b>	At least 9 layers at bitrates less than or equal to 0.55 bits per pixel. Suggestion: Include 0.55 bits per pixel to facilitate testing and some very low rates for low-resolution display.
<b>Progression Order</b>	Resolution based predominating layer order: RPCL, RLCP, or PCRL
<b>Parsability</b>	If a POC marker is present, the POC marker shall have RS <sub>POC</sub> <sup>0</sup> =0 and CS <sub>POC</sub> <sup>0</sup> =0.



### 7.1.1 FP COM Marker Segment

The FP profile number and software implementation must be identified using a 20-byte comment marker segment (COM) as specified in Table 2. The encoder may insert other COM markers at its discretion, but none of the other comments should have a Ccom string that begins 'EncID:'.

**Table 2. Content of FP specified COM marker**

Parameter	Size (bits)	Values	Notes
COM	16	0xFF64	Required for FP profile
Lcom	16	18	
Rcom	16	1	Ccom is ASCII character string
Ccom1	8*6	'EncID:'	Fixed string, Identifying this comment.
Ccom2	8*2	'1<SP>'	FP JPEG 2000 profile version number.
Ccom3	8*6	<softwareID>	Character string indicating software implementation that encoded this image. (Value assigned by the FP compliance testing body.)

## 7.2 FP JPEG 2000 Layers (informative)

The actual layer bitrates can be adjusted to meet specific program requirements. In order to have sufficient quality to allow transcoding to WSQ, the compression must contain a layer bound of at least 0.55 bpp (i.e., under 15:1 compression). Files with higher amounts of compression should not be transcoded to WSQ. To offset the cost of large file sizes, multiple layers (including some very low bitrates) should be included to facilitate progressive transmission at a variety of resolutions. If the total compression contains more than 0.55 bpp, then an encoder should include an intermediate layer at 0.55 bpp to facilitate testing.

## 7.3 FP JPEG 2000 Guidance (informative)

The FP profile for JPEG 2000 given in Section 7.1 is very broad and leaves a variety of coding alternatives unspecified. This flexibility is intentionally included to allow for future developments. However, to ensure reasonable image quality, there are compliance tests that levy an additional quality requirement that is not part of the JPEG 2000 standard. Therefore, not all JPEG 2000 codestreams that match this profile will be able to pass the quality-based certification tests. Table 3 enumerates the JPEG 2000 parameters that are used in the reference encoder, decoder, and transcoder. Implementations using these settings are more likely to pass the certification tests.

**Table 3. JPEG 2000 Parameter Guidance**

<b>Parameter</b>	<b>Test Development Settings</b>
<b>Wavelet filter</b>	9-7 irreversible *
<b>Levels of Decomposition</b>	6 *
<b>Progression</b>	RPCL
<b>Layers</b>	0.55 bpp, plus eight approximate bpp lower layers 0.35, 0.25, 0.15, 0.10, 0.06, 0.04, 0.025, 0.015
<b>Image offset</b>	0,0
<b>Subsampling (X/YRsize)</b>	1,1
<b>Components</b>	1 *
<b>Bits per sample</b>	8
<b>Tiles</b>	None
<b>Tile parts per tile</b>	1
<b>Tile offset</b>	0,0
<b>Precincts</b>	Max-size
<b>Code blocks</b>	64x64
<b>Coding alternatives</b>	
<b>Bypass mode</b>	No
<b>Reset each pass</b>	No
<b>Terminate each pass</b>	No
<b>Vertical causal contexts</b>	No
<b>Predictable termination</b>	No
<b>Segmentation symbols</b>	No
<b>Optional Markers Present</b>	None
<b>Guard Bits</b>	2
<b>Quantization Format</b>	Expounded
<b>Implementation Bit Depth</b>	32
<b>ROI's (use of RGN)</b>	None present
<b>Reconstruction Bin Position</b>	0.5

---

\* This parameter setting is required. See Table 1.

## 8 Compliance Testing

The syntax of FP JP2 files is determined by checking the presence/formatting of specific boxes/marker segments and validating that the file can be decoded with a reference decoder. Encoder, decoder, and transcoder tests are used to ensure that implementations perform coding operations accurately. In addition to the syntax and objective metric tests, visual confirmation is performed.

### 8.1 Syntax Tests

The presence and content of the mandatory box and marker segments will be checked within FP JP2 files. In addition the FP JPEG 2000 Profile parameter settings will be checked (Profile 1, 9-7 filter, progression order, LL size, etc)

Syntax Check Items
JPEG 2000 Signature box
File Type box
Image Header box
Color Specification box
Capture Resolution box
FP COM marker segment
Match Profile 1
9-7 irreversible filter
Progression order
# components
# levels of decomposition
# layers
POC restrictions

The remainder of the syntax checking is achieved by validating that the file can be decoded without warning or error messages.

### 8.2 Visual Confirmation

FP compressed images must be of sufficient quality to allow for: (1) conclusive fingerprint comparisons (identification or non-identification decision); (2) fingerprint classification; (3) automatic feature detection; and (4) overall Automated Fingerprint Identification System (AFIS) search reliability.

Compression is not expected to eliminate defects present in the source image, but it should not introduce an abundance of visual defects. Test images shall be inspected for the addition of artifacts or anomalies such as, but not necessarily limited to, the following list. Visually detected anomalies or artifacts may be quantified to further establish their degree of severity.

- boundary artifacts between tiles or codeblocks
- wavelet artifacts
- blurring

## 8.3 Implementation Tests

Three types of implementations are tested: encoders, decoders, and transcoders. Most of the tests include the computation of a few image metrics to ensure that adequate quality is maintained for each type of processing. These tests are based upon a specific set of test inputs and reference images. This section describes the form of the tests and contains metric bounds that apply to three sample images (A, B, and D) available to application builders for pre-testing. Appendix B contains information relating to the computation of these image quality metrics, and Appendix C contains the full set of metric bounds invoked during compliance testing.

### 8.3.1 Encoder Compliance Tests

The encoding process generates a 1000ppi compressed file from the original 8-bit image. To ensure adequate quality, but still allow implementation flexibility in the future, testing of the encoding process will be based upon reconstructed image metrics rather than encoded wavelet coefficients. The metric thresholds are designed for a 0.55 bpp bitrate (i.e., 14.55:1 compression ratio).

For each reference 1000ppi image, the encoder under test will generate a test 1000ppi JP2 file. This test JP2 file will then be decoded using the reference decoder<sup>1</sup>, to generate both a 1000ppi reconstruction and a 500ppi reconstruction. The test JP2 file and reconstructions will be compared to the corresponding reference JP2 (produced by a reference encoder) and the reference images, and must satisfy the following conditions.

- 1) The test JP2 file passes the syntax test.
- 2) The compressed JPEG 2000 codestream size (excluding any COM marker segments) produced by the implementation under test shall be no more than 0.1% larger than the target codestream size (0.55bpp). There is no lower limit on the codestream size. Only the size of contributions up to the layer closest to 0.55 bpp will be included in this test.

$$\frac{8S_T - 0.55N}{0.55N} 100 \leq 0.1$$

where  $S_T$  is the codestream size in bytes for the implementation under test and  $N$  is the number of pixels in the image.

- 3) The quality metrics of the 1000ppi reconstruction (at 0.55 bpp) shall conform to the bounds set out in Table C-1. Table 4 gives a small sample of the content of that table.
- 4) The quality metrics of the 500ppi reconstruction (using layers up to 0.55 bpp at the original resolution) shall conform to the bounds set out in Table C-2. Table 5 gives a small sample of the content of that table.
- 5) The test reconstructions are confirmed visually.

---

<sup>1</sup> The reference decoder used for this encoder test is JJ2000 v5.1 Available at <http://jpeg2000.epfl.ch>

**Table 4. Sample Metric Bounds for 1000ppi Encoder Test**

Original	1000ppi Reconstruction	RMSE (orig 1000, test 1000) is less than	IQM (test 1000ppi reconstruction) is greater than <sup>2,3</sup>
A.img	A.tst.1000.der	13.11	0.0268
B.img	B.tst.1000.der	9.656	0.0810
D.img	D.tst.1000.der	6.475	0.0092

**Table 5. Sample Metric Bounds for 500ppi Encoder Test**

Ref 500ppi	500ppi Reconstruction	RMSE (ref 500, test 500) is less than	IQM (test 500ppi reconstruction) is greater than <sup>3,4</sup>
A_500.img	A.tst.500.der	7.285	0.0117
B_500.img	B.tst.500.der	5.658	0.0360
D_500.img	D.tst.500.der	3.975	0.0039

### 8.3.2 Decoder Compliance Tests

Decoders must not only be able to decode FP JP2 files to sufficient quality, but also demonstrate ability to decode any JPEG 2000 Profile 1 codestream and any grayscale JP2 file.

- 1) The implementation under test has demonstrated JPEG2000 conformance [7] for Profile 1 Cclass 1 and JP2 grayscale.
- 2) The implementation under test shall decode each test input FP JP2 file fully and the resultant 8-bit image will be compared to a reference 1000ppi image. The quality metrics of the 1000ppi reconstruction shall conform to the bounds set out in Table C-3. Table 6 gives a small sample of the content of that table.

**Table 6. Sample Metric Bounds for Decoder Compliance Test**

Original	Test Image	RMSE (orig 1000, test 1000) is less than	IQM (test 1000ppi reconstruction) is greater than <sup>2,3</sup>
A.img	A.jp2.1000.dec	12.76	0.0258
B.img	B.jp2.1000.dec	9.37	0.0786
D.img	D.jp2.1000.dec	6.20	0.0090

- 3) The test reconstructions are confirmed visually.

<sup>2</sup> For 1000ppi reconstructions the threshold values in these tables are at least 73% of the original image IQM. The proportion of original image quality maintained varies with image content.

<sup>3</sup> See Appendix B for IQM preferences and auxiliary file contents required for these tests.

<sup>4</sup> For 500ppi reconstructions the threshold values in this table are at least 95% of the reference 500ppi image IQM.



### 8.3.3 Transcoder Compliance Tests

The transcoding process is used to create a 500ppi WSQ file from a FP JP2 file. Various implementation avenues exist, but they will all use some portion of a JPEG2000 decoder along with a WSQ encoder. Two alternative tests are available.

- 1) If the transcoder is separable into two segments with an 8-bit character image as a result of the JPEG 2000 decoder segment, then transcoder test A is applied.
- 2) If the JPEG 2000 decoder section uses a reconstruction factor of 0.5 and passes floating-point data to the WSQ encoder, then transcoder test B is applied.

#### 8.3.3.1 Transcoder Test A

The 8-bit image that is created by the JPEG 2000 decoder segment must be made available for testing. In addition, it must be possible to test the WSQ encoder implementation with inputs that are not derived from the JPEG 2000 decoder. Transcoder test A is then broken into two segments: a JPEG 2000 transcoder decoding test, and the WSQ encoder compliance test described in [3].

The JPEG 2000 transcoder decoding test is described here. For each test input FP JP2 file, the implementation under test will generate the 500ppi 8-bit image that would be passed to the WSQ encoder segment. This is called the test reconstruction. The test reconstruction will be compared to a reference 500ppi reconstruction. [The reference reconstruction is created using the reference JP2 decoder.]

The difference between the test reconstruction and the reference 500ppi reconstruction at any pixel shall be at most 1 gray level. The absolute value of the mean error and the mean absolute error between the test reconstruction and the reference 500ppi reconstruction shall be no more than the values set out in Table C-4. Table 7 gives a small sample of the content of that table. [The tolerances in this test are unlikely to be met without using a 0.5 reconstruction factor and greater than 16-bit implementation precision.]

**Table 7. Sample Metric Bounds for Transcoder Test A**

Reference	Test Image	Mean Absolute Error	Mean Error
A.der	A.jp2.500.dec	0.01	0.005
B.der	B.jp2.500.dec	0.01	0.005
D.der	D.jp2.500.dec	0.01	0.005

Transcoder Test A is not complete until the WSQ encoder compliance is also tested [3].

#### 8.3.3.2 Transcoder Test B

Transcoder compliance test B is nearly identical to the WSQ encoder compliance test, with the transcoder being tested as a single unit.

For each test input FP JP2 file, the transcoder will generate a 500ppi test WSQ file. The test WSQ file will be compared to the corresponding reference WSQ file and must meet the following conditions:

- 1) It is a correctly formatted WSQ fingerprint file.
- 2) The compressed file size (excluding comments) produced by the implementation under test shall be within 0.4% of the reference compressed file size.

$$\frac{|S_T - S_R|}{S_R} 100 \leq 0.4$$

where  $S_T$  and  $S_R$  are the file size for the implementation under test and the reference WSQ file respectively.

- 3) All quantization bin widths (including the zero bins) shall be within 0.051% of the corresponding bin widths contained in the quantization table within the reference compressed image.

$$\frac{|Q_{k,T} - Q_{k,R}|}{Q_{k,R}} 100 \leq 0.051 \quad \text{and} \quad \frac{|Z_{k,T} - Z_{k,R}|}{Z_{k,R}} 100 \leq 0.051 \quad 0 \leq k \leq 59$$

where  $Q_{k,R}$  and  $Q_{k,T}$  are the quantization bin widths for the  $k$ th subband in the reference and test WSQ files respectively.  $Z_{k,R}$  and  $Z_{k,T}$  are the corresponding zero bin widths.

- 4) At least 99.99% of the bin index values,  $p_k(m,n)$ , within the test implementation WSQ file shall be the same as the corresponding values in the reference WSQ file and no bin index value shall differ by more than 1.

### 8.3.4 Test Data

The following test data and information can be obtained by contacting the cognizant government office:

Federal Bureau of Investigation,  
Systems Engineering Unit, CJIS Division  
(Attn: Tom Hopper, Room 11192E)  
935 Pennsylvania Avenue, N.W.  
Washington, D.C. 20537-9700  
Telephone (voice): (202) 324-3506  
Telephone (fax): (202) 324-8826  
Email: [thopper@leo.gov](mailto:thopper@leo.gov)

- Sample test images, codestreams, and image metric bounds for compliance testing, such as those shown in the tables in Section 8.3.1 through Section 8.3.3. [Note: MITRE has prepared a CD containing this data.]
- Issuance of FP COM marker SoftwareIDs.
- Information on formal compliance certification with comprehensive test set.



## Appendix A Minimal FP JP2 Example (Informative)

Although the JP2 format is fully described in ISO 15444-1, this annex provides an informative example of a minimum content FP JP2 file for the sake of clarity. This example is not normative, since additional meta-data may appear in FP JP2 files. A JP2 decoder will be able to interpret data from any valid JP2 file.

A JP2 file is constructed of information containers called boxes. Each box begins with an indication of box length and type, followed by the contents, which may be other boxes. Boxes contain logical groupings of meta-data or a compressed image codestream.

The minimal FP content of each box is described in the following sections. For further information about box content and the different box representations in this appendix see ISO 15444-1 [4].

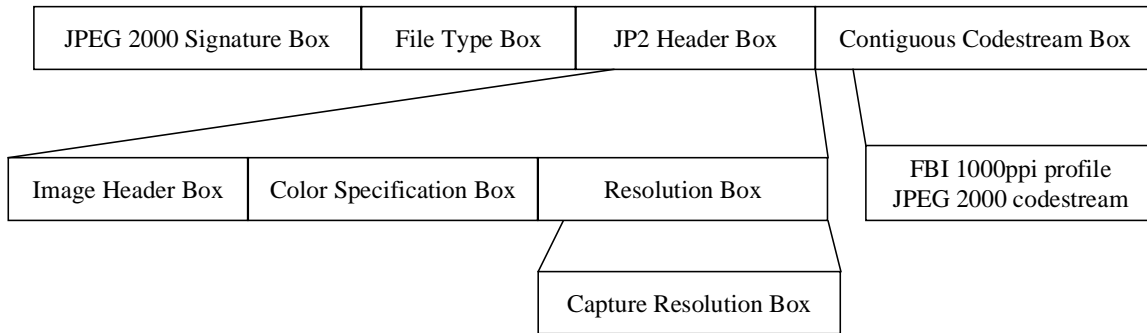


Figure A-1. High-level FP JP2 Mandatory Content

### A.1 JPEG 2000 Signature Box

The JPEG 2000 Signature box has the following format and contents.

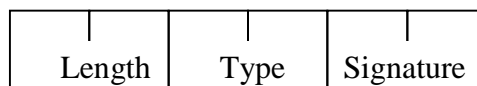


Figure A-2. Organization of the JPEG 2000 Signature Box

The Type field is written using ISO 646 [1] (ASCII) and includes the space character, denoted <SP>. In hexadecimal, a correctly formed JPEG 2000 signature box will read 0x0000 000C 6A50 2020 0D0A 870A.

Field	Value	Size (bytes)	Hexadecimal
<i>Length</i>	12	4	0000 000C
<i>Type</i>	'jP<SP><SP>'	4	6A50 2020
<i>Signature</i>	'<CR><LF><0x87><LF>'	4	0D0A 870A

Figure A-3. Contents of the JPEG 2000 Signature Box

### A.2 File Type Box

A minimal FP JP2 file type box has the following contents. More complex versions of this box are possible, but not required for encoders. Decoders shall be able to properly interpret any JP2 file type box. See ISO 15444-1 for a complete description of this box, and what additional options are available.



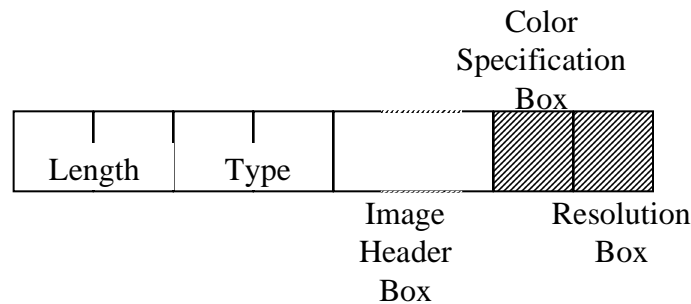
**Figure A-4. Organization of the File Type Box**

Field	Value	Size(bytes)	Hexadecimal
<i>Length</i>	20	4	0000 0014
<i>Type</i>	‘ftyp’	4	6674 7970
<i>Brand</i>	‘jp2<SP>’	4	6A70 3220
<i>Minor Version</i>	0	4	0000 0000
<i>CL</i>	‘jp2<SP>’	4	6A70 3220

**Figure A-5. Contents of the File Type Box**

### A.3 JP2 Header Box

A minimal FP JP2 Header box is a superbox with the following format and contents. More complex versions of this box are possible, but not required for encoders. Decoders shall be able to properly interpret any JP2 file type box. See ISO 15444-1 for a complete description of this box, and what additional options are available.



**Figure A-6. Organization of the JP2 Header Box**

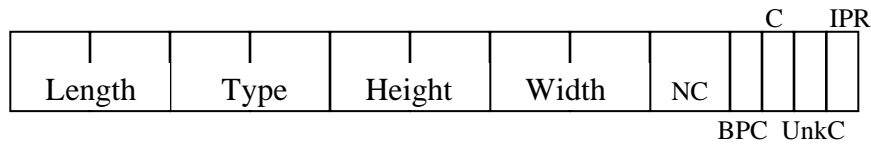
Field	Value	Size (bytes)	Hexadecimal
<i>Length</i>	71	4	0000 0047
<i>Type</i>	‘jp2h’	4	6A70 3268

**Figure A-7. Contents of the JP2 Header Box**



### A.3.1 Image Header Box

The Image Header box has the following format and contents.



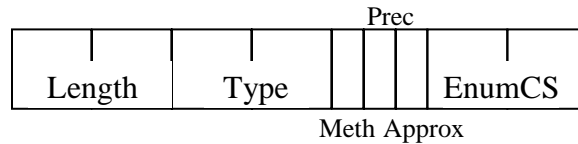
**Figure A-8. Organization of the Image Header Box**

Field	Value	Size (bytes)	Hexadecimal
<i>Length</i>	22	4	0000 0016
<i>Type</i>	'ihdr'	4	6968 6472
<i>Height</i>	Ysiz-YOsiz	4	
<i>Width</i>	Xsiz-XOsiz	4	
<i>NC (# components)</i>	1	2	0001
<i>BPC (bit depth minus one and sign of all components)</i>	7	1	07
<i>C (compression type)</i>	7	1	07
<i>Unknown Colorspace Flag</i>	0	1	00
<i>IPR</i>	0	1	00

**Figure A-9. Contents of the Image Header Box**

### A.3.2 Color Specification Box

The Color Specification box has the following format and contents for a grayscale fingerprint. If color data is allowed in the future, then see ISO 15444-1 for a complete description of alternatives available for color.



**Figure A-10. Organization of the Color Specification Box**

Field	Value	Size (bytes)	Hexadecimal
<i>Length</i>	15	4	0000 000F
<i>Type</i>	'colr'	4	636F 6C72
<i>Method</i>	1	1	01
<i>Precedence</i>	0	1	00
<i>Approximation</i>	0	1	00
<i>Enumerated Colorspace</i>	17 (= grayscale)	4	0000 0011

**Figure A-11. Contents of the Color Specification Box**

### A.3.3 Resolution Box

A minimal FP Resolution box is a superbox with the following format and contents. The presence of this box is mandatory for FP JP2 files. More complex versions of this box are possible, but not required for FP encoders. Decoders shall be able to properly interpret any resolution box. See ISO 15444-1 for a complete description of this box, and what additional options are available. [If this format is used for fingerprints at resolutions different from 1000ppi, then the Capture Resolution fields must be modified to indicate the appropriate resolution.]

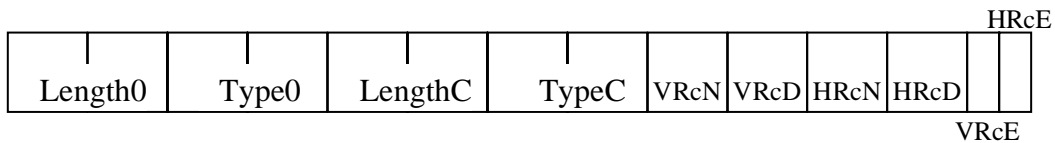


Figure A-12. Organization of the Resolution Box

Field	Value	Size (bytes)	Hexadecimal
<i>Length0</i>	26	4	0000 001A
<i>Type0</i>	'res<SP>'	4	7265 7320
<i>Length Capture Res</i>	18	4	0000 0012
<i>Type Capture Res</i>	'resc'	4	7265 7363
<i>VRcN (pixels / meter)</i>	39370	2	99CA
<i>VRcD</i>	1	2	0001
<i>HRcN (pixels / meter)</i>	39370	2	99CA
<i>HRcD</i>	1	2	0001
<i>VRcE</i>	0	1	00
<i>HRcE</i>	0	1	00

Figure A-13. Contents of the Resolution Box (indicating a capture resolution of 1000ppi)

### A.4 Contiguous Codestream Box

The contiguous codestream box consists of the box length and type indications followed by a JPEG 2000 codestream.



Figure A-14. Organization of the Contiguous Codestream Box

Field	Value	Size (bytes)	Hexadecimal
<i>Length</i>	Codestream length +8	4	
<i>Type</i>	'jp2c'	4	6A70 3263

Figure A-15. Contents of the Contiguous Codestream Box

## Appendix B Quality Metrics

There is no single image metric that is known to exactly match human perception of image quality. Instead a variety of metrics are used in the literature to test various aspects of image ‘quality.’ Several of them are used for the purposes of FP compliance testing.

### B.1 Comparative Image Metrics

Many image metrics compare a test image against a fixed reference. RMSE, Mean Absolute Error, and Absolute Mean Error are all of this type. In each of these metrics a difference is computed between the gray values in the two images at each of the  $N$  pixel positions, and then the resulting difference image is incorporated into a particular formula.

#### B.1.1 Root Mean Square Error (RMSE)

The root mean square error is computed using the following formula:

$$RMSE(reference, test) = \sqrt{\frac{1}{N} \sum_{i=1}^N (test_i - reference_i)^2}$$

Where  $test_i$  and  $reference_i$  are gray values in the corresponding images at pixel position  $i$ . The sum is computed over all ( $N$ ) pixel positions.

#### B.1.2 Mean Absolute Error

The mean absolute error is computed using the following formula:

$$MeanAbsoluteError = \frac{1}{N} \sum_{i=1}^N |test_i - reference_i|$$

When the maximum difference between two images is less than or equal to one, the mean absolute error becomes a measure of the number of image pixels which vary between the test and reference image.

#### B.1.3 Absolute Mean Error

The absolute mean error is computed using the following formula:

$$|MeanError| = \left| \frac{1}{N} \sum_{i=1}^N (test_i - reference_i) \right|$$

A large absolute mean error is indicative of an overall shift in image brightness, with the test image either brighter or darker than the reference image.

## B.2 Single-Image Metrics

In contrast to a comparative image metric, a single-image metric only relies upon data from the test image itself. The goal of this type of metric is to give a sense of image quality without reference to another image.

### B.2.1 Image Quality Measure (IQM)

IQM is a single-image quality metric. An executable implementation, a user's guide, and a paper describing IQM in more detail, can be found at: <http://www.mitre.org/tech/mtf>

The IQM code requires 3 input files to run:

- image file
- auxiliary data pertaining to the image (auxdatafile)
- preferences file that specifies IQM run parameters (prefsfile)

IQM is computed using the following formula (eq.14 in Opt.Eng. paper at above website):

$$IQM = \frac{1}{M^2} \sum_{\theta=-180^{\circ}}^{180^{\circ}} \sum_{\rho=0.01}^{0.707} S(\theta_1) W(\rho) A^2(T\rho) P(\rho, \theta)$$

For the specific application to fingerprints in this document, the parameters in the above formula have the following values:

- M = number of pixels across width of square image to which IQM is applied (specified in auxdatafile)
- $S(\theta_1) = 1.0$  when specifying sensor #4 in auxdatafile
- $W(\rho) = 1.0$  with the noise-related parameter values defined in the default prefsfile, IQM should compute a value of 1.0 for  $W(\rho)$ ; if it computes a different value, signified by "problem code" 5 or 6 appearing in IQM output file it implies the fingerprint image is far off-the-mark, e.g., very noisy
- $A(T\rho)$  = visual response modulation transfer function, with peak of MTF set to 0.5 cy/ pixelwidth when using default prefsfile values: spot=0.6, viewdist=351.3288 (T=internal constant)
- $P(\rho, \theta)$  = power spectrum of image, normalized by zero frequency power (in default prefsfile: psntype=DC)
- $\rho$  = radial spatial frequency in units of cycles per pixelwidth; lower & upper limits defined in default prefsfile (freqmin=0.01, freqmax=0.707107);  $p_{max}$  is bounded by maximum cartesian coordinate values:  $x_{max}=y_{max}=0.50$
- $\theta$  = angle around the two-dimensional power spectrum

For 8 bpp images, IQM expects pure white in the image to be gray level 255. The user should always verify that the polarity of the images, in combination with the polarity parameter value set in the auxdatafile, either "L" or "B", results in IQM reading a near-white area of the image as near gray level 255, and a near-black area as near gray level 0.

This can be verified for an individual image by noting the gray levels for 4 pixels displayed during IQM runtime, or the gray level for 1 pixel printed to the output file, and comparing to what is known to be correct for the given image, in the given pixel locations. [For more details, see IQM\_Guide, section 3-Image Formats.]

The IQM computation is only applied to square image areas. The location and size of this square area for each input image is part of the auxdatafile. The auxdatafile used on reconstructed images for encoder and decoder tests based on sample test image A is given in Figure B-1<sup>5</sup> (the IQM executable can automatically generate this file via interactive user input). Figure B-2 shows an example of how IQM interprets this auxdatafile.

```
# AuxDataFile TEMPLATE for IQM run of Fingerprint Images @ 1000ppi & 500ppi
# Use with IQM's DefaultPrefsFile
# IQM is applied to square subimage in each case; subimage size is dependent on image
# Subimage width for 500ppi image is always 1/2 of subimage width for corresponding 1000ppi image
# All cases: sensor 4, mag=1.0 for 1000ppi image, mag=0.5 for 500ppi image
# User should Always verify polarity (L or B) of actual images, as read by IQM on his/her computer !

A.tst.1000.der.pgm  GRAY      1000ppi reference reconstruction for encoder test
0 0 0 8L 0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 4 0 1.000
72 56 1024

A.tst.500.der.pgm  GRAY      500ppi reference reconstruction for encoder test
0 0 0 8L 0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 4 0 0.500
36 28 512

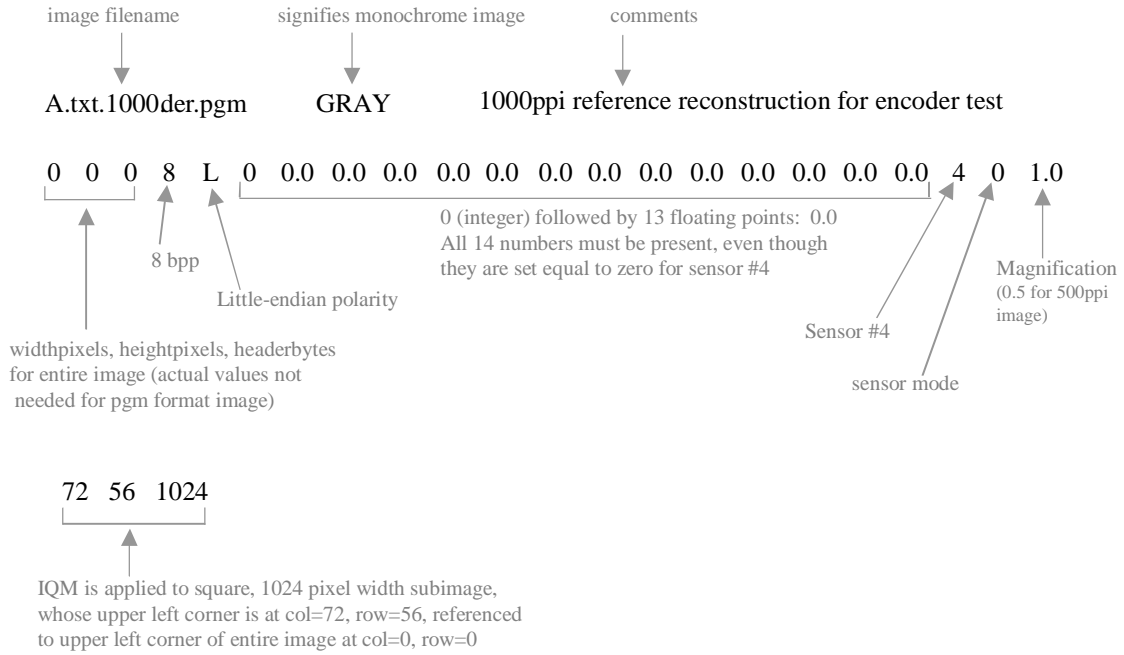
A.jp2.1000.dec.pgm  GRAY      1000ppi reconstruction for decoder test
0 0 0 8L 0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 4 0 1.000
72 56 1024
```

**Figure B-1. AuxdataFile Content for Image A**

The sample B and D images repeat this pattern, with differences only in the final line denoting the horizontal and vertical pixel offset and size of the square subimage to which the IQM is applied. The final lines for B and D are:

```
B 1000ppi: 144 68 900
B 500ppi: 72 34 450
D 1000ppi: 37 39 800
D 500ppi: 19 20 400
```

<sup>5</sup> The auxdatafile is free format, using spaces to separate inputs. Line breaks must occur as shown.



**Figure B-2. Example of IQM Interpretation of the AuxdataFile**



## Appendix C Metric Bounds

This appendix contains the full tables of metric bounds used for the compliance tests described in Section 8.3. This information is provided for use by a testing agency with access to the reference imagery. Tables in Section 8.3 contain bounds for sample data and reference imagery provided to application builders.

### C.1 Notation

The following key may be helpful in interpreting the table content.

Original	Original 1000ppi image prior to any compression
Ref 500ppi	500ppi version of the original image. Created using the 9-7 irreversible JPEG 2000 transform, rounded back to 8-bit integers without any other quantization
Reference	Image used as a basis for comparison. Not necessarily the original
Col 1	Column 1 entry
Col 2	Column 2 entry
*.der	An image created with the reference decoder The reference decoder used for this encoder test is JJ2000 v5.1 Available at <a href="http://jj2000.epfl.ch">http://jj2000.epfl.ch</a>
*.dec	Decoded image created by the product under test
*.jp2	An FP JP2 provided as test input
*.tst	An FP JP2 created by encoder software under test
*.1000.*	A full resolution decode
*.500.*	A half resolution decode

For encoder compliance, the software under test will create \*.tst files. The testing organization will then decode these files both at full resolution and at half resolution using the reference decoder to generate \*.tst.1000.der and \*.tst.500.der.

For decoder compliance, the software under test will fully decode the \*.jp2 files to create \*.jp2.1000.dec images.

For transcoder compliance test A, the software under test will perform a half resolution decode to generate \*.jp2.500.dec images.

These output \*.der or \*.dec images are compared to the appropriate reference images shown in the first column of the testing tables shown here.

### C.2 Metric Bounds Tables

The metrics contained in these tables (RMSE, Mean Absolute Error, Absolute Mean Error and IQM) are described in Appendix B. See Appendix B.2.1 for details on the preferences and auxiliary data required to compute IQM. The only IQM parameters which vary by image content across the compliance tests are horizontal and vertical pixel offset and size for the square subimage to which IQM is applied. The IQM Parameters column indicates the offset and size values required for those tests.

**Table C-1. Metric Bounds for 1000ppi Encoder Test**

<b>Original</b>	<b>1000ppi Reconstruction</b>	<b>RMSE(Col 1,Col2) is less than</b>	<b>IQM Parameters</b>	<b>IQM (Col 2) is greater than</b>
A.img	A.tst.1000.der	13.11	72 56 1024	0.0268
B.img	B.tst.1000.der	9.656	144 68 900	0.0810
D.img	D.tst.1000.der	6.475	37 39 800	0.0092
enc001.img	enc001.tst.1000.der	7.61	366 340 800	0.0346
enc002.img	enc002.tst.1000.der	4.76	290 534 700	0.0028
enc003.img	enc003.tst.1000.der	5.22	754 666 600	0.0466
enc004.img	enc004.tst.1000.der	6.06	64 322 460	0.0392
enc005.img	enc005.tst.1000.der	6.53	116 274 680	0.0108
enc006.img	enc006.tst.1000.der	6.44	362 666 500	0.1050

**Table C-2. Metric Bounds for 500ppi Encoder Test**

<b>Ref 500ppi</b>	<b>500ppi Reconstruction</b>	<b>RMSE(Col 1,Col 2) is less than</b>	<b>IQM Parameters</b>	<b>IQM (Col2) is greater than</b>
A_500.img	A.tst.500.der	7.285	36 28 512	0.0117
B_500.img	B.tst.500.der	5.658	72 34 450	0.0360
D_500.img	D.tst.500.der	3.975	19 20 400	0.0039
enc001_500.img	enc001.tst.500.der	3.79	183 170 400	0.0143
enc002_500.img	enc002.tst.500.der	2.73	145 267 350	0.00123
enc003_500.img	enc003.tst.500.der	3.07	377 333 300	0.0206
enc004_500.img	enc004.tst.500.der	3.48	32 161 230	0.0176
enc005_500.img	enc005.tst.500.der	3.83	58 137 340	0.00488
enc006_500.img	enc006.tst.500.der	3.93	181 333 250	0.0483

**Table C-3. Metric Bounds for Decoder Compliance Test**

<b>Original</b>	<b>Test Image</b>	<b>RMSE(Col1,Col2) is less than</b>	<b>IQM Parameters</b>	<b>IQM (Col2) is greater than</b>
A.img	A.jp2.1000.dec	12.76	72 56 1024	0.0258
B.img	B.jp2.1000.dec	9.37	144 68 900	0.0786
D.img	D.jp2.1000.dec	6.20	37 39 800	0.0090
dec001.img	dec001.jp2.1000.dec	4.052	622 348 700	0.0146
dec002.img	dec002.jp2.1000.dec	4.442	420 378 800	0.00642
dec003.img	dec003.jp2.1000.dec	4.532	386 470 680	0.00326
dec004.img	dec004.jp2.1000.dec	2.998	334 860 460	0.0408
dec005.img	dec005.jp2.1000.dec	3.916	366 652 680	0.00601
dec006.img	dec006.jp2.1000.dec	5.436	550 590 600	0.0455

**Table C-4. Metric Bounds for Transcoder Test A**

<b>Reference</b>	<b>Test Image</b>	<b>Mean Absolute Error</b>	<b> Mean Error </b>
A.der	A.jp2.500.dec	0.01	0.005
B.der	B.jp2.500.dec	0.01	0.005
D.der	D.jp2.500.dec	0.01	0.005
tns001.der	tns001.jp2.500.dec	0.01	0.005
tns002.der	tns002.jp2.500.dec	0.01	0.005
tns003.der	tns003.jp2.500.dec	0.01	0.005
tns004.der	tns004.jp2.500.dec	0.01	0.005
tns005.der	tns005.jp2.500.dec	0.01	0.005
tns006.der	tns006.jp2.500.dec	0.01	0.005

# JPEG2000 and Google Books

Jeff Breidenbach



Google's mission is to organize the world's information and make it universally accessible and useful.



## Mass digitization

- broad coverage
- iteratively improve quality (reprocess, rescan)
- more than  $XXM$  books out of  $XXXM$  since 2004
- $XXX$  nominal pages per book
- billions of images, petabytes of data

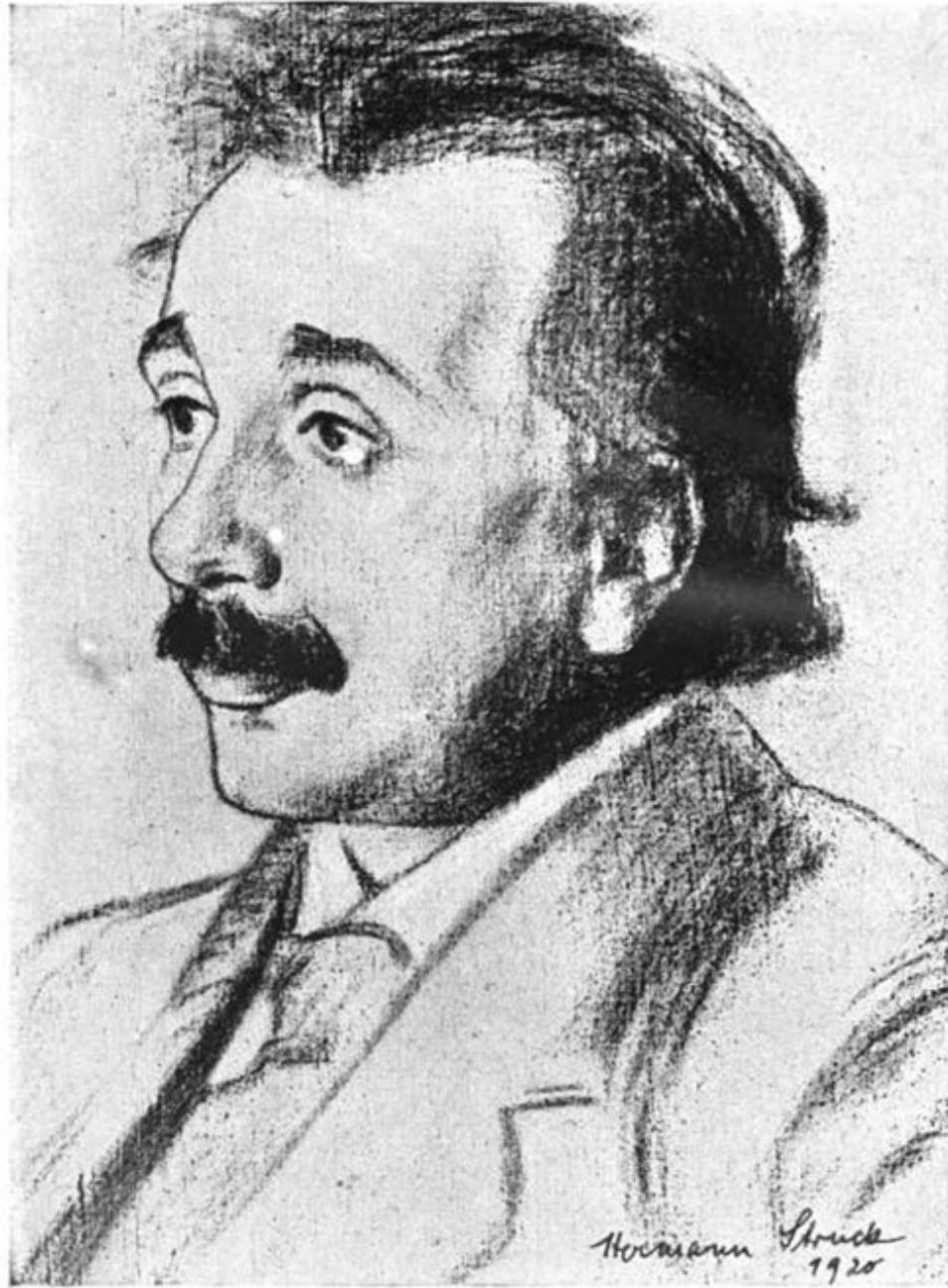
## JPEG2000

- pre-processed images
- processed illustrations and color images
- library return format
- illustrations inside PDF files





EPHESUS.



*A. Einstein*



Dem aller durchleuch-  
tigsten/ Großmechtigsten Fürsten vnd  
Herrn/ Herrn Ferdinando/ Römischen/ zu  
Ungern vnd Böhem 1c. König/ Ertz-  
herzog zu Oesterreych 1c. Meinem  
aller Gnedigsten Herrn.



Aler Durchleuch-  
tigster/ Großmechtig-  
ster Römischer König  
1c. Aller Gnedigster  
HERR/ Wiewol vil  
hochgelerte züvor/ mit  
höchstem fleiß nicht on  
nuz vnnnd frucht der  
glaubigen/ von der hai-  
ligen Mes geschriben/  
hab ich doch den güt-

hertzigen zü trost/ dise meine schrifft vnd Predig  
von dem Ampt der heyligen Mes/ auß Christen-  
lichem eyffer auch wöllen mittailen / Dieweil die  
verachtung der Mes bey etlichen vngehorsamen  
nicht abnimbt/ sonder täglich wechs/ so auch die  
ainfeltigen Catholischen mit irriger lehr/ als ob  
die Mes ein verdamlicher grewol were/ zü berede  
vnd zü verwirren ganz geschwind vñ abgericht.

\* ij Disen

Dem aller durchleuch-  
tigsten/ Großmechtigsten Fürsten vnd  
Herrn/ Herrn Ferdinando/ Römischen/ zu  
Ungern vnd Böhem 1c. König/ Ertz-  
herzog zu Oesterreych 1c. Meinem  
aller Gnedigsten Herrn.

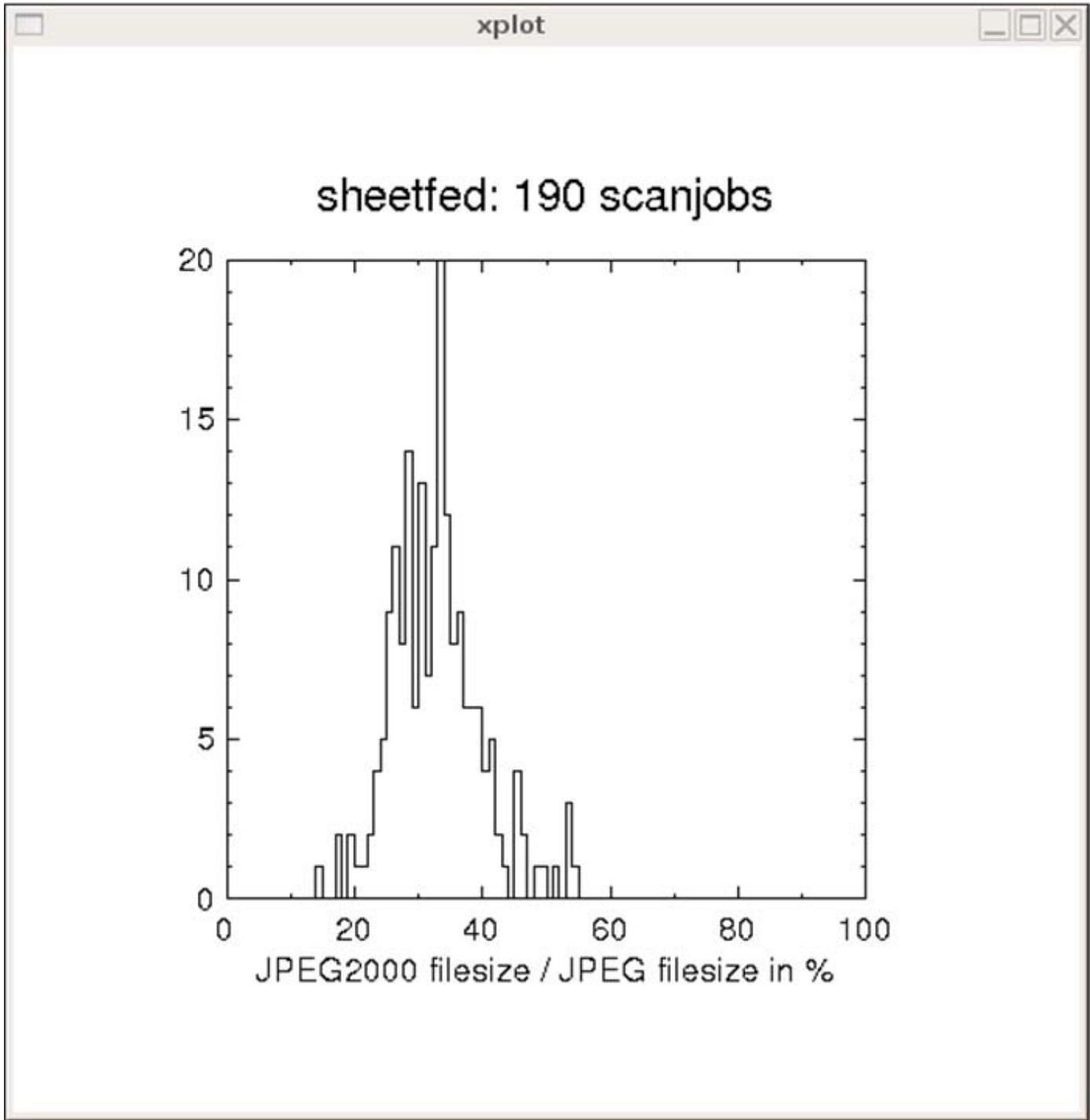


Aler Durchleuch-  
tigster/ Großmechtig-  
ster Römischer König  
1c. Aller Gnedigster  
HERR/ Wiewol vil  
hochgelerte züvor/ mit  
höchstem fleiß nicht on  
nuz vnnnd frucht der  
glaubigen/ von der hai-  
ligen Mes geschriben/  
hab ich doch den güt-

hertzigen zü trost/ dise meine schrifft vnd Predig  
von dem Ampt der heyligen Mes/ auß Christen-  
lichem eyffer auch wöllen mittailen / Dieweil die  
verachtung der Mes bey etlichen vngehorsamen  
nicht abnimbt/ sonder täglich wechs/ so auch die  
ainfeltigen Catholischen mit irriger lehr/ als ob  
die Mes ein verdamlicher grewol were/ zü berede  
vnd zü verwirren ganz geschwind vñ abgericht.

\* ij Disen







**JPEG  
Quality**

**JPEG**

**JPEG2000**

**100**

no. 180

no. 180

**93**

no. 180

no. 180

**85**

no. 180

no. 180

**75**

no. 180

no. 180



Jhove (Rel. 1.4, 2009-07-30)

Date: 2011-05-03 20:06:36 PDT

RepresentationInformation: 00000001.jp2

ReportingModule: JPEG2000-hul, Rel. 1.3 (2007-01-08)

LastModified: 2006-09-22 11:01:00 PDT

Size: 231249

Format: JPEG 2000

Status: Well-Formed and valid

SignatureMatches:

JPEG2000-hul

MIMEtype: image/jp2

Profile: JP2

JPEG2000Metadata:

Brand: jp2

MinorVersion: 0

Compatibility: jp2

ColorspaceUnknown: true

ColorSpecs:

ColorSpec:

Method: Enumerated Colorspace

Precedence: 0

Approx: 0

EnumCS: sRGB

UUIDs:

UUIDBox:

UUID: -66, 122, -49, [...]

Data: 60, 63, 120, [...]

Codestreams:

Codestream:

ImageAndTileSize:

Capabilities: 0

XSize: 1165

YSize: 2037

XOSize: 0

YOSize: 0

XSize: 1165

YSize: 2037

XTOSize: 0

YTOSize: 0

CSize: 3

SSize: 7, 1, 1

XRSize: 7, 1, 1

YRSize: 7, 1, 1

CodingStyleDefault:

CodingStyle: 0

ProgressionOrder: 0

NumberOfLayers: 1

MultipleComponentTransformation: 1

NumberDecompositionLevels: 5

CodeBlockWidth: 4

CodeBlockHeight: 4

CodeBlockStyle: 0

Transformation: 0

QuantizationDefault:

QuantizationStyle: 34

StepValue: 30494, 30442, 30442, 30396, 28416, 28416, 28386, 26444, 26444, 26468, 20483, 20483, 20549, 22482, 22482, 22369

NisImageMetadata:

MIMEtype: image/jp2

ByteOrder: big-endian

CompressionScheme: JPEG 2000

ImageWidth: 1165

ImageLength: 2037

BitsPerSample: 8, 8, 8

SamplesPerPixel: 3

Tiles:

Tile:

TilePart:

Index: 0

Length: 229827

# Jhove



```
<?xpacket begin=" id='W5M0MpCehiHzreSzNTczkc9d'?>
<x:xmpmeta xmlns:x='adobe:ns:meta/'>
<rdf:RDF xmlns:rdf='http://www.w3.org/1999/02/22-rdf-syntax-ns#'>
<rdf:Description rdf:about=" xmlns:tiff='http://ns.adobe.com/tiff/1.0/'>
<tiff:ImageWidth>1165</tiff:ImageWidth>
<tiff:ImageLength>2037</tiff:ImageLength>
<tiff:BitsPerSample>
<rdf:Seq>
<rdf:li>8</rdf:li>
<rdf:li>8</rdf:li>
<rdf:li>8</rdf:li>
</rdf:Seq>
</tiff:BitsPerSample>
<tiff:Compression>34712</tiff:Compression>
<tiff:PhotometricInterpretation>2</tiff:PhotometricInterpretation>
<tiff:Orientation>1</tiff:Orientation>
<tiff:SamplesPerPixel>3</tiff:SamplesPerPixel>
<tiff:XResolution>300/1</tiff:XResolution>
<tiff:YResolution>300/1</tiff:YResolution>
<tiff:ResolutionUnit>2</tiff:ResolutionUnit>
<tiff:DateTime>2004-04-27T00:00:00+08:00</tiff:DateTime>
<tiff:Artist>Google, Inc.</tiff:Artist>
<tiff:Make>MDP</tiff:Make>
<tiff:Model>Photostation v1</tiff:Model>
<tiff:Software>Photostation v1 scanning software</tiff:Software>
</rdf:Description>
<rdf:Description rdf:about=" xmlns:dc='http://purl.org/dc/elements/1.1/'>
<dc:source>jp2k/0345430573/00000001.jp2</dc:source>
</rdf:Description>
</rdf:RDF>
</x:xmpmeta>
<?xpacket end='w'?>
```

## Embedded XMP



# The Joys of Slope Rate Distortion

// Lower number == less distortion == higher fidelity

const int [kJp2kLosslessQuality](#) = 0;

// decent

const int [kJp2kOceanDefaultNDRawQuality](#) = 50980;

// same space as JPEG-75

const int [kJp2kOceanDefaultNDCleanQuality](#) = 51180;

// removes JPEG artifacts

const int [kJp2kOceanDefaultSFRawQuality](#) = 51315;

// GEFGW

const int [kJp2kOceanDefaultSFCleanQuality](#) = 51350;

// acceptable

const int [kJp2kOceanGRINImagePageQuality](#) = 51492;

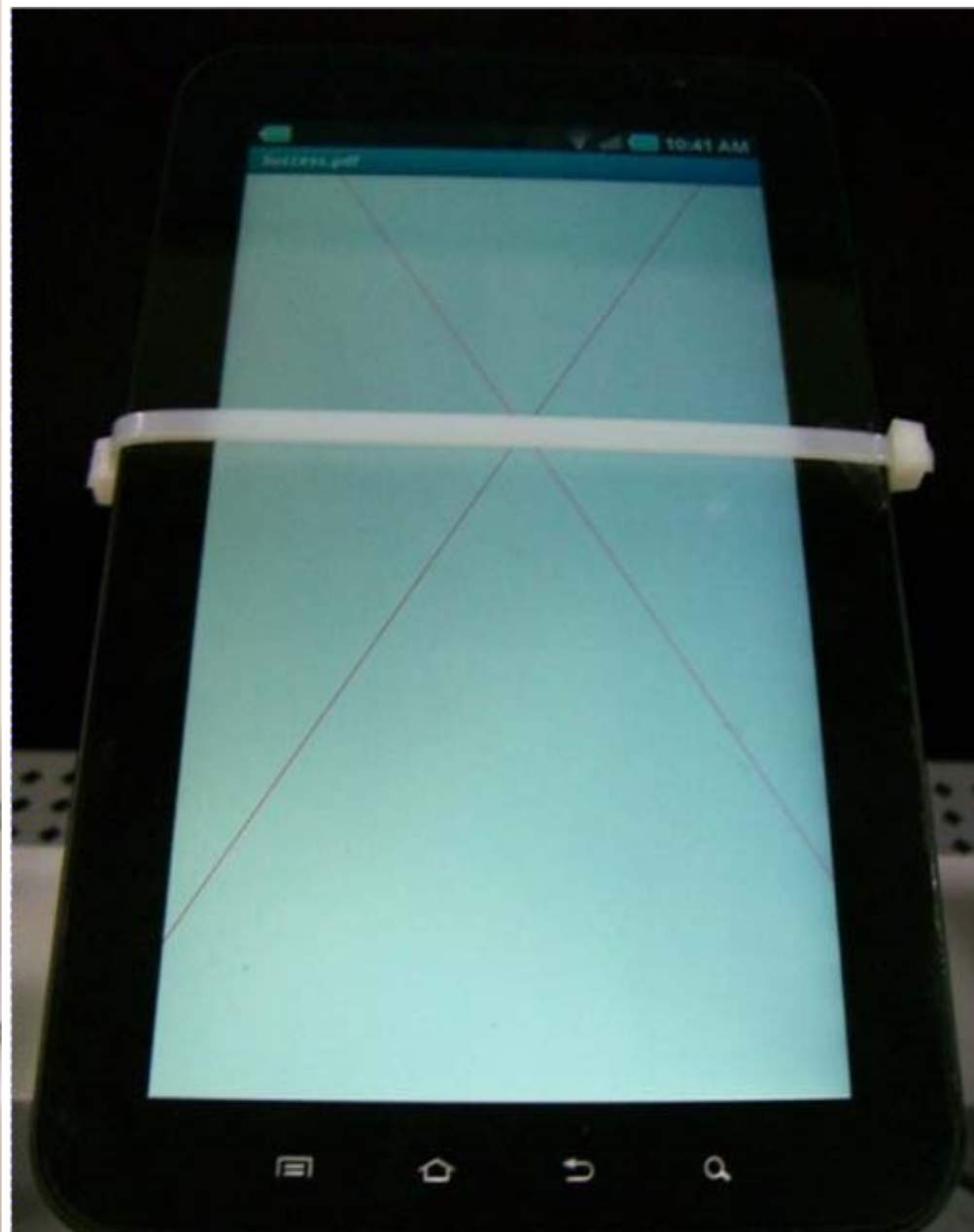
// marginal

const int [kJp2kOceanGRINTextPageQuality](#) = 52004;



# JP2K/PDF Compatibility

[Credit: Mike Cane](#)



Thank you / questions

# Backup Slides



## •WebP

mainly for publishing on the Web  
very efficient coding (e.g: segmentation) esp.  
at low bitrate. Comparable to h264.  
block-based: decoding footprint is very light  
(memory scales with the width)  
fast decoding: 2x-3x slower than jpeg (9x for  
JP2K with Kakadu)  
encoding still slow, being worked on  
royalty-free  
evolving quickly with container features

## ORIGINAL ARTICLE

# Evaluation of JPEG and JPEG2000 compression algorithms for dermatological images

KH Gulkesen,<sup>†,\*</sup> A Akman,<sup>‡</sup> YK Yuca,<sup>†</sup> E Yilmaz,<sup>‡</sup> AA Samur,<sup>†</sup> F Isleyen,<sup>†</sup> DS Cakcak,<sup>‡</sup> E Alpsoy<sup>‡</sup>

<sup>†</sup>Department of Biostatistics and Medical Informatics, <sup>‡</sup>Department of Dermatology, Akdeniz University, Medical Faculty, Antalya, Turkey

\*Correspondence: KH Gulkesen. E-mail: hgulkesen@akdeniz.edu.tr

## Abstract

**Background** Some image compression methods are used to reduce the disc space needed for the image to store and transmit the image efficiently. JPEG is the most frequently used algorithm of compression in medical systems. JPEG compression can be performed at various qualities. There are many other compression algorithms; among these, JPEG2000 is an appropriate candidate to be used in future.

**Objective** To investigate perceived image quality of JPEG and JPEG2000 in 1 : 20, 1 : 30, 1 : 40 and 1 : 50 compression rates.

**Methods** In total, photographs of 90 patients were taken in dermatology outpatient clinics. For each patient, a set which is composed of eight compressed images and one uncompressed image has been prepared. Images were shown to dermatologists on two separate 17-inch LCD monitors at the same time, with one as compressed image and the other as uncompressed image. Each dermatologist evaluated 720 image couples in total and defined whether there existed any difference between two images in terms of quality. If there was a difference, they reported the better one. Among four dermatologists, each evaluated 720 image couples in total.

**Results** Quality rates for JPEG compressions 1 : 20, 1 : 30, 1 : 40 and 1 : 50 were 69%, 35%, 10% and 5% respectively. Quality rates for corresponding JPEG2000 compressions were 77%, 67%, 56% and 53% respectively.

**Conclusion** When JPEG and JPEG2000 algorithms were compared, it was observed that JPEG2000 algorithm was more successful than JPEG for all compression rates. However, loss of image quality is recognizable in some of images in all compression rates.

Received: 15 September 2009; Accepted: 10 November 2009

## Keywords

data compression, dermatology, photography

## Conflicts of interest

None.

## Introduction

Since digital images have been used in health domain, picture archiving and transmission has become pretty easy. The clinical applications of digital photography are numerous.<sup>1-4</sup> Digital images, including dermatoscopic images, can be used to document clinical information.<sup>5</sup> Changes in skin lesions can readily be documented and monitored through serial imaging.<sup>6</sup> Clinical photography may also help histopathological diagnosis.<sup>7</sup> Approximately 85% of the dermatologists in New York City use camera and ratio of digital cameras is increasing.<sup>8</sup> Digital photography is also useful in the relatively new area of teledermatology.<sup>9</sup>

Although digital imaging is cheaper than conventional methods, digital image archiving and transmission still has a cost. Some

compression methods are used to reduce the disc space needed for the image, store and transmit the image efficiently.<sup>10</sup> However, as image quality may have critical value in medicine, each compression method and ratio must be evaluated.

Joint Photographic Experts Group (JPEG, JPG) is an image compression standard, which was declared by Joint Photographic Expert Group in 1992. Since then, it has been the predominant image file format, which is used in wide spectrum of applications including World Wide Web and digital photography. It is the most frequently used algorithm of image compression in medicine also.<sup>11</sup> JPEG compression can be performed at various qualities. All digital cameras currently in the market support JPEG format and almost all compact digital cameras are capable of saving

pictures only in JPEG format. Although there are differences in some technical details in JPEG file specifications, they share many common features. In JPEG's compression algorithm, the image is divided into  $8 \times 8$  pixel blocks and image information in each block is summarized by a mathematical function called discrete cosine transform. JPEG compression can be performed at various qualities. However, dividing the image into blocks can be held responsible for pixelization artefact and blurring, which can be observed especially in high compression rates. In spite of its wide use, JPEG file format has some weaknesses and research continues to find better compression methods.<sup>12</sup>

There are many other compression algorithms; among these, JPEG2000 (JP2) may be an appropriate candidate to be used for medical image compression in future.<sup>13</sup> Joint Photographic Expert Group announced this algorithm and its file format in 2000. JPEG2000 compression algorithm uses a different mathematical function called discrete wavelet transform to summarize image information and it does not need to divide the image into blocks. It is generally accepted that it achieves higher image quality compared with JPEG, specifically in high compression rates, which is attributed to the use of discrete wavelet transform. Another advantage of JPEG2000 over JPEG comes with its multi-resolution decomposition structure. JPEG2000, as an output of its progressive sub-band transform, i.e. discrete wavelet transform, a multi-resolution image is obtained. In other words, a JPEG2000 file, along with the original file, contains the same image at different resolutions. Also, it is capable of generating both lossy and lossless image compressions. Its main drawback is the need for higher computer performance during encoding and decoding.<sup>14</sup>

There are several studies on both methods in the medical literature. JPEG compression was reported to be useful for histopathological microscopic images.<sup>15</sup> Fifty times magnified digital videomicroscope melanocytic lesion images have shown no significant loss of image quality by 1:30 JPEG compression.<sup>16</sup> For ultrasonographic images, nine times JPEG compression is possible without recognizable loss of quality.<sup>17</sup> In a study on mammography images, the authors have reported 50 times compression without loss of diagnostic quality using JPEG2000 compression.<sup>18</sup> JPEG2000 algorithm has been reported to be more successful than JPEG algorithm in a radiology study on various image types.<sup>19</sup> Evaluation of retinal images with JPEG and wavelet compression of 1.5 MB images resulted in acceptable image quality for 29 kB JPEG and 15 kB wavelet compressed files.<sup>20</sup> Performance of classic JPEG and JPEG2000 algorithms is equivalent when compressing digital images of diabetic retinopathy lesions from 1.26 MB to 118 kB and 58 kB. Higher compression ratios show slightly better results with JPEG2000 compression.<sup>21</sup>

In spite of frequent use of digital images in dermatology, only a couple of studies to evaluate the effect of image compression on clinical macroscopic images have been performed up to date. It was reported that dermatologist's diagnostic performance was the same for both JPEG and fractal image formats up to 1 : 40 com-

pression.<sup>22</sup> According to the other study, lossy JPEG2000 compression was slightly less efficient than JPEG, but preserved image quality much better, particularly at higher compression factors. It was concluded that for its good image quality and compression ratio, JPEG2000 appears to be a good choice for clinical/video-microscopic dermatological image compression.<sup>10</sup>

The aim of this study was to investigate perceived image quality of JPEG and JPEG2000 in various compression rates.

## Materials and methods

Ethical Committee approval was obtained for the study. In total, 90 photographs in digital negative (RAW) format were taken from patients who came to dermatology outpatient clinics. The lesions which have educational value had been selected for taking photographs. Educational value had been determined by faculty of Dermatology Department. The photographs were taken by Canon EOS 40D body, Canon EF 50 mm f/2.5 Compact Macro lens, Canon EF 28–200 mm f/3.5–5.6 USM Standard Zoom lens and Canon Speedlite 580EX flash (Canon Inc., Tokyo, Japan).

Uncompressed images were processed using Adobe Photoshop CS3 software (Adobe Systems Inc., San Jose, CA, USA). For JPEG2000 recognition of Adobe Photoshop, LEADJ2K plug-in (Lead Technologies Inc, Charlotte, NC, USA) was installed. RAW files were converted to uncompressed Tagged Image File Format (TIFF) file, a file format which is equivalent to Bitmap image file format (BMP) and recognized by most of the imaging software. Subsequently,  $3888 \times 2592$  images were resampled by bicubic sharper method. Horizontal images were resampled to  $972 \times 648$  pixels and vertical images were resampled to  $648 \times 972$  pixels. The sizes of the images were determined according to resolution of the monitors. Each image was saved in JPEG and JPEG2000 file formats, by 1 : 20, 1 : 30, 1 : 40 and 1 : 50 compression rates. So, for each patient (image), a set which is composed of eight compressed images and one uncompressed image has been prepared.

Images were shown to dermatologists on two separate  $1280 \times 1024$  resolution 17-inch LCD monitors (IBM Thinkvision) at the same time as pairs with no specific order pattern, i.e. being one is compressed and the other is uncompressed or *vice versa*. For instance, image pairs were like the following, (uncompressed, JPEG-n), (JPEG2000-n, uncompressed), (JPEG-n, uncompressed), where n stands for a compression rate. The dermatologists were told that one of the images was uncompressed, but they did not know which one. By the help of Irfanview plug-ins 4.10 for JPEG2000 visualization, slideshow function of Irfanview 4.10 (<http://www.irfanview.com>) was used for demonstration. The monitors were calibrated using Monitor Calibration Wizard 1.0 (<http://www.hex2bit.com>).

Each dermatologist defined whether there was a difference between two images by quality. If there was a difference, they reported a better one. Among four dermatologists, each has evaluated 720 image couples in total.

By the evaluations of dermatologists, a quality rate was calculated for each compression rate and method. Quality rate was calculated as 100%-(% of selected compressed images-% of selected uncompressed images).

Rater agreement was tested by multirater kappa test (<http://justus.randolph.name/kappa>). Detection of quality difference was tested by chi-square for each compression rate and algorithm, using SPSS software (Statistical Package for Social Sciences, v11.0, SPSS Inc., Chicago, IL, USA). The data were preprocessed to equally distribute 'they are same' decisions to compressed and uncompressed groups.

**Results**

Details of evaluation of dermatologists are presented in Table 1. They have similar opinions, but dermatologist-4 has irregular decisions and has quality rates over 100%, a result of preferring compressed images. Kappa value for rater agreement is 0.390. When Dermatologist-4 is excluded, kappa value rises to 0.471. The mean of quality rates of the other three dermatologists are shown in Fig. 1. Quality rates for JPEG compressions 1 : 20, 1 : 30, 1 : 40 and 1 : 50 were 69%, 35%, 10% and 5% respectively. Quality rates for corresponding JPEG2000 compressions were 77%, 67%, 56% and 53% respectively. JPEG2000 compression has significantly better quality than JPEG algorithm ( $P < 0.001$  for all compression rates).

**Discussion**

Joint Photographic Experts Group compression is frequently used in medicine and reported to be useful in low compression rates.<sup>17,20,21</sup> In last couple of decades, several studies comparing it with other compression algorithms have appeared in biomedical literature. The aim of these studies can be summarized as to find another algorithm which yields better image quality in high compression rates. Among other compression algorithms, JPEG2000 is possibly the most frequently studied one and most of the studies reported that JPEG2000 is more efficient than JPEG in high compression rates.<sup>10,19-21</sup> On the other hand, it is

seen that the result of the studies are not very consistent with each other. There are significant methodological differences between the studies. The source of the image may be digital camera,<sup>10</sup> a radiological modality<sup>18</sup> or scanned transparent.<sup>20,22</sup> It may be coloured<sup>10</sup> or black and white.<sup>20</sup> Compression rates show variability, and most importantly, evaluation method of 'quality' and statistical methods differ significantly. The images were visualized by using variable monitors and even standard light box.<sup>18</sup> In some studies, the raters have been asked to give a quality score to each image;<sup>22</sup> some studies were based on diagnostic value of the image<sup>18</sup> and some were based on comparison of images.<sup>17</sup> In some studies, mathematical analyses of compressed files were performed.<sup>10,19</sup>

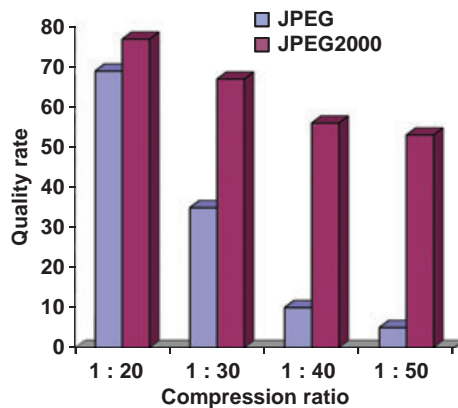
Some studies report that there is no significant quality loss for 1 : 30 compressed JPEG and 1 : 50 compressed JPEG2000 images.<sup>16,18</sup> However, our raters report detectable loss of quality nearly 20% of JPEG2000 and 30% of JPEG compressed images when they compressed in 1 : 20 ratio. In all compression rates, JPEG2000 had higher quality rates than JPEG images. In several studies, JPEG2000 algorithm has been reported more successful than JPEG algorithm.

However, most of these differences reflect different approaches to digital image. Will it be used for diagnosis, education, follow-up, print, analysing by computer software or legal issues? Images for each of these requirements may have different features. In our study, we have tried to look at the problem from human perception point, to detect recognizable quality differences by experienced dermatologists. Naturally, the same compressed image may be sufficient for one purpose, insufficient for another purpose.<sup>23</sup> For example, the image may be useful for slide presentations, but may have insufficient quality for printing. If a multipurpose image archive is considered, sufficient quality for all the purposes is desired. Our study shows that there is a recognizable loss of image quality in dermatological images even in 1 : 20 compression rates for both for JPEG2000 and JPEG, even though the former method is better. In study of Guarneri *et al.*, perceived image quality is nearly equal to uncompressed files for 1 : 5 JPEG and 1 : 14

**Table 1** Decision of four dermatologists on image couples

Compression	Dermatologist-1 Age: 34, 10 YE				Dermatologist-2 Age: 50, 21 YE				Dermatologist-3 Age: 44, 18 YE				Dermatologist-4 Age: 31, 3 YE			
	U	S	C	QR	U	S	C	QR	U	S	C	QR	U	S	C	QR
JPEG 1 : 50	86	3	1	6	82	8	-	9	88	2	-	2	86	4	-	4
JPEG 1 : 40	85	5	-	6	69	20	1	24	89	1	-	1	58	16	16	53
JPEG 1 : 30	54	36	-	40	53	35	2	43	69	21	-	23	73	9	8	28
JPEG 1 : 20	14	75	1	86	24	59	7	81	54	35	1	41	12	41	37	128
JPEG2000 1 : 50	33	57	-	63	43	41	6	59	59	28	3	38	56	22	12	51
JPEG2000 1 : 40	30	60	-	67	38	43	9	68	60	29	1	34	32	26	32	100
JPEG2000 1 : 30	14	76	-	84	38	45	7	66	45	42	3	53	61	24	5	38
JPEG2000 1 : 20	11	79	-	88	23	62	5	80	33	57	-	63	13	33	44	134

YE, years of dermatology experience; U, uncompressed image is better; S, they are the same; C, compressed image is better; QR, quality rate.



**Figure 1** Mean quality rates according to three dermatologists for JPEG and JPEG2000 compression methods.

JPEG2000 compression rates.<sup>10</sup> In future studies, other compression algorithms and lower compression rates of JPEG and JPEG2000 algorithms should be investigated for finding a fully satisfying compression method for digital dermatology images.

In the present study, four dermatologists were our raters. However, one of our raters was statistically inconsistent with the other raters. In previous studies, inconsistency of raters has been reported and it seems that there is a personal variation for perception of image quality.<sup>24</sup> Interestingly, inconsistent rater in this study was the youngest one, who can be considered as more close to the computer-age culture. However, the rater has the shortest dermatology practice, and the least experience in using digital dermatology images for education and presentation among the raters. So the experience may be an important factor for rater reliability.

In conclusion, when JPEG and JPEG2000 algorithms were compared, it was seen that JPEG2000 algorithm was more successful than JPEG for all of the compression rates in dermatological images. However, 1 : 20 compressed images of both algorithm have recognizable loss of quality and lower compression rates should be considered for the images which are considered for multipurpose use.

## References

- Papier A, Peres MR, Bobrow M, Bhatia A. The digital imaging system and dermatology. *Int J Dermatol* 2000; **39**: 561–575.
- Aspres N, Egerton IB, Lim AC, Shumack SP. Imaging the skin. *Australas J Dermatol* 2003; **44**: 19–27.
- Fawcett RS, Widmaier EJ, Cavanaugh SH. Digital technology enhances dermatology teaching in a family medicine residency. *Fam Med* 2004; **36**: 89–91.
- Kaliyadan F. Digital photography for patient counseling in dermatology – a study. *J Eur Acad Dermatol Venereol* 2008; **22**: 1356–1358.
- Rushing ME, Hurst E, Sheehan D. Clinical Pearl: the use of the handheld digital camera to capture dermoscopic and microscopic images. *J Am Acad Dermatol* 2006; **55**: 314–315.
- Chamberlain AJ, Dawber RP. Methods of evaluating hair growth. *Australas J Dermatol* 2003; **44**: 10–18.
- Fogelberg A, Ioffreda M, Helm KF. The utility of digital clinical photographs in dermatopathology. *J Cutan Med Surg* 2004; **8**: 116–121.
- Scheinfeld NS, Flanigan K, Moshiyakhov M, Weinberg JM. Trends in the use of cameras and computer technology among dermatologists in New York City 2001–2002. *Dermatol Surg* 2003; **29**: 822–825.
- Eedy DJ, Wootton R. Teledermatology: a review. *Br J Dermatol* 2001; **144**: 696–707.
- Guarneri F, Vaccaro M, Guarneri C. Digital image compression in dermatology: format comparison. *Telemed J E Health* 2008; **14**: 666–670.
- Siegel DM. Resolution in digital imaging: enough already? *Semin Cutan Med Surg* 2002; **21**: 209–215.
- Miano J. *Compressed image file formats: JPEG, PNG, GIF, XBM, BMP*. Addison Wesley Longman Inc, Reading, MA, 1999.
- Puniene J, Punys V, Punys J. Medical image compression by cosine and wavelet transforms. *Stud Health Technol Inform* 2000; **77**: 1245–1249.
- Acharya T, Tsai PS. *JPEG2000 standard for image compression: Concepts, algorithms and VLSI architectures*. Wiley Interscience, Hoboken, NJ, 2005.
- Foran DJ, Meer PP, Papatomas T, Marsic I. Compression guidelines for diagnostic telepathology. *IEEE Trans Inf Technol Biomed* 1997; **1**: 55–60.
- Seidenari S, Pellacani G, Righi E, Di Nardo A. Is JPEG compression of videomicroscopic images compatible with telediagnosis? Comparison between diagnostic performance and pattern recognition on uncompressed TIFF images and JPEG compressed ones. *Telemed J E Health* 2004; **10**: 294–303.
- Persons KR, Hangiandreou NJ, Charboneau NT et al. Evaluation of irreversible JPEG compression for a clinical ultrasound practice. *J Digit Imaging* 2002; **15**: 15–21.
- Penedo M, Souto M, Tahoces PG et al. Free-response receiver operating characteristic evaluation of lossy JPEG2000 and object-based set partitioning in hierarchical trees compression of digitized mammograms. *Radiology* 2005; **237**: 450–457.
- Shiao YH, Chen TJ, Chuang KS et al. Quality of compressed medical images. *J Digit Imaging* 2007; **20**: 149–159.
- Eikelboom RH, Yogesan K, Barry CJ et al. Methods and limits of digital image compression of retinal images for telemedicine. *Invest Ophthalmol Vis Sci* 2000; **41**: 1916–1924.
- Conrath J, Erginay A, Giorgi R et al. Evaluation of the effect of JPEG and JPEG2000 image compression on the detection of diabetic retinopathy. *Eye* 2007; **21**: 487–493.
- Sneiderman C, Schosser R, Pearson TG. A comparison of JPEG and FIF compression of color medical images for dermatology. *Comput Med Imaging Graph* 1994; **18**: 339–342.
- Tanaka M. Minimum requirements for digital images in dermatological publications. *Clin Exp Dermatol* 1999; **24**: 427.
- Ween B, Kristoffersen DT, Hamilton GA, Olsen DR. Image quality preferences among radiographers and radiologists. A conjoint analysis. *Radiography* 2005; **11**: 191–197.

This document is a scanned copy of a printed document. No warranty is given about the accuracy of the copy. Users should refer to the original published version of the material.



# Analysing the Impact of File Formats on Data Integrity

Volker Heydegger; Universität zu Köln; Cologne, Germany

## Abstract

*The concept of file format is fundamental for storage of digital information. Without any formatting rules, bit sequences would be meaningless to any machine. Due to various reasons there exists an overwhelming mass of file formats in the digital world, even though only a minority has a broader relevance. Particularly in regard of specific domains like long-term preservation of digital objects, the choice of the appropriate format can be a problematic case. Thus, one of the basic questions an archivist needs to get an answer for is: Which file format is most suitable for ensuring the longevity of its information?*

*In this study a particular criteria for long-term preservation suitability is picked up: the robustness of files according to their bit error resilience. The question we address is: Up to what extent does a file format, as a set of formatting rules, contribute to the long-term maintainability of the information content of digital objects? Or in other words: Are there any file format basing factors promoting the consistency of digital information?*

## Introduction

Among several other criteria [9], one considered to be crucial for the decision which file format to choose for digital preservation refers to the capability of file formats to keep its information, as it is, over a long period against the evil of bit rot. The single reasons for corrupted files are manifold. Nevertheless there are two main categories: First, bit errors in files occur in consequence of degradation of the storage medium, e.g. caused by poor physical storage conditions, just as a natural decay of the medium or as a consequence of massive usage. This is especially true for storage of data on optical disks [5]. Hard disks are also exposed to such errors although less severe [7][12]. Second, bit errors result from transmission procedures. However, e.g., in case of data migration, these errors can be prevented if methods for checking the integrity of the data are implemented.

The nature of the corruption of files can also be manifold. Bit errors can be located to special areas of the file, they can also be distributed [5]. The actual location of bit errors within a file strongly depends on the underlying reason for corruption: E.g. consider a DVD which was damaged by the influence of strong heat. In this case the distribution of bit errors may vary according to the strength of direction of the heat source. On the other hand, files can be corrupted in a way that not only single bits are flipped from zero to one or vice versa but also that they totally get lost. In such cases, the effect on data integrity increases dramatically. In this study we focus on bit errors in the sense of flipped bits and on equally distributed errors. In fact there is actually no general tendency of error location in files as a consequence of the manifold reasons for corruption we mentioned before.

The current strategy to get the problem of file corruption under control targets at hardware-sided solutions. Determined by the storage medium, data is usually stored according to particular methods which again follow international standards. Specific

codes for error correction are adapted to the processes of reading and writing data from and to the storage medium. The devices which deal with the medium are constantly refined in their ability to handle it with higher precision, thus improving the quality of the data as well. New technologies using different methods and materials for storage media, e.g. holography, promise to push on the durability of the medium while increasing the storage capacity at the same time. However, all of these efforts are not primarily the result of a basic sense for the necessity of keeping data as safe as possible; most notably they arise from the necessity to cope with the advancing technological complexity of such devices and storage media.

Even if it would succeed to get a grip on the problems of storage technology in terms of durability and capacity of storage media more accurately: If it comes to make long-term preservation of data also feasible in an economical sense, there is no doubt to follow up additional strategies for improving data integrity. The proposal to keep data by redundant copies, additionally locally distributed, is a simple and useful approach but may suffer from additional cost effects [1][10].

The study on hand takes up this necessity to find backing solutions and moves away from the problem of physical and technical restrictions of storage media. The focus is now on logical representation and organisation of data as files, which is determined by a set of given rules, commonly called file format. The concept of file format is the fundament for data to become meaningful. Data interpreted by a machine according to the underlying file format is not only raw data but information.

So the question we address is: Up to what extent does a file format, as a set of formatting rules, contribute to the long-term maintainability of the information content of digital objects? Or in other words: Are there any file format basing factors promoting the consistency of digital information?

The consequence of clarifying this question is obvious: If there is indeed a significant relation between file format and information constancy, it will be possible, in due consideration of the revealed determinants, to improve the long-term preservation of digital objects: E.g., existing file formats could be optimized, newly created file formats could be, with the help of the updated knowledge, conceived including aspects of longevity.

Studies in this area focused one specific aspect of representing data in files: Data compression [2][3][8]. This is not surprising, for data compression is a major feature of file formats, especially in terms of data integrity of files. It arises from certain technological facts which originate in the information technologies past, for capacity of storage media and efficiency of data processing systems were formerly quite more limited than they are now. Nevertheless these are still factors to be considered. Though technological progress may lessen these limitations, the mass of digital data still increases. After all this will be more and more a domain specific question. The question if to store data as compressed or not does not arise for digital objects like movies;

however, for an archivist who wants to keep his images for long-term preservation, this may be a question worth asking.

Indeed, especially in terms of long-term preservation, one was sceptical about the usage of data compression for a long time: Compressed data is extremely prone to consecutive faults caused by bit errors. Therefore, besides other reasons, JPEG 2000 was also developed with the goal in mind to make compressed data more robust against bit errors. Since then the discussion on usage of compressed data for preservation purpose is sparked again [2][4][8].

Although this study takes up this special point, we also focus on other aspects of file format, namely which kind of data is captured by the file format and how data is structured and related among each other.

Additionally we concentrate on image files as our practical subject of research here. Therefore the following remarks have a strong relation to image file formats.

## General Implications on File Format Data

### *Usage and Processing*

In the context of this study, a file format is a set of rules constituting the logical organisation of data and indicating how to interpret them. The quantity of set of rules may vary to a great extent and depends strongly on the information intended to be represented. In the context of this study we call all information that can be described through one or more files, their formatted data respectively, a digital object.

The complexity of digital objects may variegate also in a certain span. But even within similar categories of information, digital objects can be described by file formats in an extremely different level of complexity. A digital object of domain 'image' may be modelled in a raw data format, using quite few formatting rules. If it is intended to be transferred and represented through a specific software like a web browser, the functionality of a raw data format usually does not last anymore. Or as another example: An image intended to be represented not only statically as a whole but from which certain parts of it are matter of interest may be expressed best way using JPEG 2000 file format. The question on which format to choose for a digital objects data is in terms of temporary usage a question concerning the scope of application of that object.

An essential conclusion that can be drawn from these considerations is: Every digital object is provided with a basic content of information. This is directly reflected through the data which represent that information. Additionally the basic content of information can also be modified and enriched by added functionality.

Information is exactly that in what humans as the users of data are interested in. Exaggerated: A user does not care about data. From the users point of view a perfectly preserved digital object presents the same information to him or her as originally intended. With respect to a categorization of file format data, this should be seen from a different perspective : A perfectly preserved digital object presents the same information as originally intended after its data has been processed by a file format data processing software following the rules given by a file format specification.

The relation is now contrariwise: The software does not care about information but data.

Software which has to cope with the task of transforming data to useful information needs to rely on the readability of data. Data must be processible according to the underlying file format.

Which conclusions can be drawn out of this regarding bit error corrupted files? For simplification of the following example let us presume that a given file format defines as smallest processible unit one byte (as it is indeed usual in most formats). If so, a single bit error causes a one byte error, this is an error rate of 1:1. We call this plain information loss. In this case, the actual change in the bit state corresponds to the actual information loss (given one byte as smallest processible unit) since it affects the information which is represented by exactly one (the corrupted) byte. Consider a comparison of two files A (this is the original, uncorrupted file) and B (the original version as corrupted file), where B differs from A in exactly one byte. A program that is able to perform a pairwise comparison of the byte values of the files then recognizes exactly one different byte. In a sequence of unformatted bits every change in the bit state is definite and irreversible. For data described by a file format this is not necessarily so. E.g., file formats which allow for error correction codes within the data potentially enable the processing software to recapture the original byte (bit) value.

Sometimes a file format specification defines a byte value as fixed value. In such cases it is also possible to recapture the original byte value from the affected byte. However, such format specific definitions must be implemented by the processing software. Conventional software applications which implement a file format compliant to its rules should not accept such an error (by the way: this is exactly what a file format recovery tool does not).

Simple bit errors do not always cause plain information loss with 1:1 error rates as shown in this example. The error rate is expected to be multiplied if a file format defines logical information units by more than one single byte. We call this kind of information loss, logical information loss. E.g., for the case of a file format assigning four byte for representation of big numbers: the information loss for an one bit error then increases to an error rate of 1:4 (again in terms of byte as reference unit).

A third kind of information loss is much more effective regarding information loss. We term it conditional information loss. Such information loss produces error rates of much higher extent than those discussed so far. In the extreme case it causes the content of the entire file to not being processible, with the result of error rates increasing up to 100%. TIFF file format for example allows for placing the pixel data of an image at any position within a file except the first eight positions which are always fixed for special usage. This file format rule necessitates to set an offset to that position within the file where the pixel data can be found. This is done in the so called image file directory, which also can be placed arbitrarily within the file (again except starting at one of the first eight positions in the file). It is once more necessary to set an offset that tells the processing software where to find the beginning of the image file directory. A bit error occurring in the offset data to the start of the pixel data, not only causes an error, in the sense of a logical information loss, within the offset data per se. As an aftereffect, at least any 'conventional' processing software does not find out anymore where exactly the pixel data is located within the file. In this case we have a conditional information loss to the amount of the number of those data

indicated via the offset. More worse, such bit errors raise the conditional information loss to the maximum if, like in this example, the error already occurs in the offset to the image file directory. Repairing such an error is even for a file format recovery tool a hard job to do. To adjust such errors in corrupted files is a real challenge for file format recovery tools.

### **Functionality and Categorization**

Data, organized according to a file format, is in its basic function an information carrier. The primary task of data-processing software is to read data with respect to the file format and to capture its information content. Such processed data can then finished according to the aspired purpose. An image viewer for example reads data as defined by the image file format from a file to transfer it to one of the image viewers concurrent purposes.

Even though all of the data described through a file format always represent some kind of information, the nature of this information is different, at least in terms of functionality. That is why a file format assigns functional meaning to data, according to its information content.

Which kind of information is represented by file format data? We generally differentiate between two main categories, which are also the basis for the robustness indicators described in the following section. The first category relates to aspects relevant for usage, the other to data related to processing tasks.

The basic content of information of a digital object that was already discussed in the previous section is reflected in the first of these two main categories. Such information and its carrying data respectively is essential for representation of the object.

Data relevant for usage can be distinguished in three sub-categories. Those of the data relevant for usage which carry the basic information of a digital object are called basic data. In case of a raster image rendered to a display, the carrier of these information are the processed pixel data. Or, in case of a simple text encoded in HTML, this is those data which map the text as, for example, accessible via a web browser. In case of an audio file, this is all of the data interpretable as sound, basically all sample data.

A second sub-category of data relevant for usage can be characterised as not directly carrying a digital objects information; nevertheless this data represents information which is indirectly necessary for adequate representation of the information content of the base data. We call that kind of data derivative data. Data on picture coordinates, bit depth or compression method are representatives from the image domain. In case of text domain, this can be data relating to text formatting information, for example font style, font size or space settings.

Another sub-category of data relevant for usage is commonly known as descriptive metadata. It adds such information to digital objects that is irrelevant for basic representation of the object. Data about creator, author, date of creation or producing software are examples for that sub-category. We call it supplemental data.

The second main category of file format data introduced here is data concerned with the structural organisation and the technical processibility of any other file format data, i.e., in its core this is data relevant for processing tasks.

At a first glance, such kind of data seems to play a minor role opposite to the object-related information carrying data. However, this is not the case. Often such data is essential for the

processibility of the entire file. The example for TIFF file format we discussed in the previous section deals with data of that category.

Processing-relevant data is distinguished in two sub-categories. Such data supporting the structural configuration of the entire data is called structural data. Structural file format data describe the logical units of the file organisation. Examples for this category are the tag numbering in TIFF files, offsets to the position of certain related data, or data that functions as filler data. Structural data is directly related to the structure of the data described by the file format.

Another sub-category includes data giving information on the validity of subsequent data units. We call it definitional data. By its application on target data, data of that kind gives an answer to the question if a certain sequence of data units (the target data) are valid or not according to the parameter defined through that data. Error correcting codes or indications on the data-type to be used are two examples for that category. In contrary to structural data, definitional data asks for an interpretation on any target data.

The advantage of such a categorization should be evident. Bit errors can now related to a categorization scheme. A close analysis of the distribution of these categories on different file formats can indicate which kind of data loss is to be expected. The results of quantities analysis of errors in corrupted files can be discussed by means of a distinct vocabulary. It is also possible to derive measures for information loss using these categories. Recently, even though in a slightly different context, the assumption of general file format data categories has led to the development of new comprehensive practical approaches to the characterisation of file format data [14].

## **Measuring Information Loss**

### **Robustness Indicators**

Building on the theoretical foundations we examined in the previous chapters, metrics for measuring information loss in corrupted image files were derived. These metrics are called robustness indicators (according to reflections in [13]). They give us a hint on the robustness of a file format in terms of the categorized file format data. Thus, in difference to similar existing metrics (e.g., RMSE, simple match coefficient), these metrics explicitly refer to our categorization of file format data.

The robustness indicators can not be interpreted as image quality measures. They are prepared for giving information on information loss that is caused by data which has changed or which original information content can not be captured anymore; this can be the case if a byte, as the information carrying unit, is directly corrupted (plain information loss) or because a certain number of bytes can not be processed adequately (logical or conditional information loss). Again: Information loss is always reflected by data as carrier of information. In the following, we present those Robustness Indicators which we applied to the test corpus in the next section.

RB is defined as a robustness indicator for file formats which relates to the basic data of usage:

$$R_B = \Delta(b_0, b_1) / m \quad (1)$$

where

$b_0$  is the basic data of usage before being corrupted,

$b_1$  is the basic data of usage after being corrupted,

$m$  is the number of corruption procedures.

$R_{Bt}$  additionally includes the relation to the total number of basic data of usage:

$$R_{Bt} = R_B / n \quad (2)$$

where

$n$  is the total number of basic data of usage.

### A Test Implementation for Measuring Information Loss

We have implemented a software tool that is able to simulate data corruption, which can recognize data according to the file format data categories we defined, that is able to process and translate the relations between the data categories and that finally computes the robustness indicators.

In its core procedure it analyses files (which represent the underlying file format) in several subsequent processing parts, using both the original (error-free) file and a manipulated (bit-corrupted) version of it. The latter is prepared by the manipulation module of the software, also taking compressed data into account by trying to decompress the corrupted files. After that, the tool analysis the original file as to the data categories defined in the model. Another module transforms the data of both files into an internal normalised representation, processing file format specific data allocations as described in the file format specification. In a last move, the data of the normalised corrupted representation is used to compute RIs.

We have also built a corpus of test files for a number of image file formats. In this study we report on the results for four of them: TIFF, PNG, BMP (windows) and JP2. The corpus comprises files which consider various basic characteristics and features of each file format. The results reported in here relate to a 'real world image', i.e. a colored image, standard 24-Bit RGB. For some of the file formats we created different test files reflecting potentially important characteristics in terms of the expected data effects on data integrity. In this case we added compression characteristics (for details see table 1). As already discussed, they so far played a leading role in the discussion of file formats robustness and their potentiality for long-term preservation respectively.

Table 1 shows the results for Robustness indicator  $R_{Bt}$ . For better readability the results for  $R_{Bt}$  are transformed to base 100 (i.e. expressed in percentage). The single file formats and compression characteristics are put in the first column. The given ratios relate to compression ratio understood as ratio between uncompressed size and compressed size of the files. The indication in brackets is the compression ratio in terms of space savings. The other columns contain the single results for  $R_{Bt}$ . We have performed test series on the base of byte errors with corruption rates of exactly one byte (which results in individual percentage corruption rates based on the original file size (second column, indication in brackets) and three more for percentage corruption rates of 0.01, 0.1 and 1.0, since they seem to be sufficient enough to clearly show the effects on file corruption in general and with respect to  $R_{Bt}$  in specific. For each file type and corruption rate we

performed the corruption procedures 3000 times always using a different set of random numbers per single corruption, generated by Mersenne Twister algorithm [6] that guarantees equal distribution of errors, as we intended to have for this part of the study. We also made sure that none of the single random numbers per set occurred twice or more to avoid imprecision of  $R_{Bt}$  values. We also cross-checked the results with confidence intervals indicating a deviation of the  $R_{Bt}$ s of less than three percentage in all cases.

**Table 1: Results for  $R_{Bt}$  (in percentage) for various file formats**

	1 Byte	0.01	0.1%	1.0%
<b>TIFF</b>				
uncompressed	0.00 (0.00063)	0.56	6.64	48.83
JPEG compressed, ratio 1:2.60 (62%)	2.14 (0.00166)	13.03	-	-
JPEG compressed, ratio 1:10.72 (90%)	2.44 (0.00505)	13.32	-	-
LZW compressed, ratio 1:1.01 (2%)	1.37 (0.00064)	18.79	77.95	99.34
ZIP compressed, ratio 1:1.28 (22%)	27.12 (0.00081)	84.92	98.47	-
<b>PNG</b>				
ZLIB compressed, unfiltered	18.21 (0.00074)	79.15	97.63	-
ZLIB compressed, filtered	25.05 (0.00085)	81.83	98.08	-
<b>BMP (windows)</b>				
uncompressed	0.00 (0.00063)	0.14	1.92	15.29
<b>JP2</b>				
lossless, ratio 1:1.36 (27%)	17.53 (0.00086)	76.22	94.29	-
lossy, ratio 1:7.42 (87%)	33.31 (0.00166)	51.86	95.03	-
lossy, ratio 1:2.64 (62%)	22.61 (0.00468)	72.93	95.62	-

### Discussion of the Results

The results reveal a strong correlation between usage of compression and data integrity. As compression is a widely used feature in many file formats, for some explicitly dedicated to (e.g., JP2), compression can be considered as one of the most important features of file formats and therefore is one of the crucial factors for a file formats impact on data integrity. In almost all cases of compression usage, 0.1 percentage of byte corruption is enough to produce  $R_{Bt}$  values of more than 90 percentage (in case of TIFF with JPEG compressed data we were not able to compute  $R_{Bt}$  with sufficient exactness since the errors provoked serious software crashes). For example TIFF with ZIP compressed data : More than 98 percentage of the basic data of the corrupted file is changed compared to the original data. Or in other words: More than 98

percentage of single information units have changed according to the change in the data which carries this information.

Almost more amazing are the results for one byte corruptions. In case of JP2, a one byte error causes, as a consequence of conditional information loss, a change in basic data of about 17 percentage for lossless compressed data (corruption rate: 0.00086), up to 33 percentage for lossy compressed data (corruption rate: 0.00166) in moderate compression ratio (JP2 is able to produce much higher compression ratio). Conditional information loss is symptomatic to compressed data and seems to not depending on whether data is compressed lossless or in a lossy mode.

**Table 2: Totally failed test files (in percentage)**

	1 Byte	0.01%	0.1%	1.0%
<b>TIFF</b>				
uncompressed	0.00	0.36	3.60	32.00
JPEG compressed, ratio 1:2.60 (62%)	0.13	0.67	-	-
JPEG compressed, ratio 1:10.72 (90%)	0.11	5.63	-	-
LZW compressed, ratio 1:1.01 (2%)	0.03	1.20	13.43	72.40
ZIP compressed, ratio 1:1.28 (22%)	0.07	0.50	3.77	-
<b>PNG</b>				
ZLIB compressed, unfiltered	0.00	0.70	4.30	-
ZLIB compressed, filtered	0.00	0.10	4.30	-
<b>BMP</b>				
uncompressed	0.00	0.10	1.67	11.07
<b>JP2</b>				
lossless, ratio 1:1.36 (27%)	0.40	0.40	11.10	-
lossy, ratio 1:7.42 (87%)	0.20	2.00	12.10	-
lossy, ratio 1:2.64 (62%)	0.10	1.30	10.40	-

Particularly for the JP2 results,  $R_{Bt}$  may be a convenient measure for reflecting the characteristics of JP2 files after being corrupted. Already for low corruption rates, the rendered versions of corrupted JP2 files can be extremely different (Figure 1). This is not a JP2 specific issue. Nevertheless JP2 compression is, compared to other compressions, quite successful in producing images which keep their visual quality, especially in case of low corruption rates, although there are moderate differences in pixel data (see Figure 1 also). However, the effects of bit corruption on the rendered files can vary to a great extent. Right due to that,  $R_{Bt}$  values reflect the actual information loss, not influenced by the deficiency of humans visible system. If it is our task to make a clear statement on whether the data of a file is in danger to be

Figure 1: Two JP2 images, both with the same degree of corruption (one single byte); the second image shows no visual difference to the rendered version of the original uncorrupted file (not illustrated) although there are actual changes in pixel data (as shown in the third pseudo- image, where different pixel data is marked in red).



changed after a bit corruption, the visual appearance of the object after being rendered is not a matter of interest.

Again, this is the task of quality measures. So while considering JP2 as a candidate for long-term storage, this still remains a point for discussion, at least if one decides that error resilience should be an important issue for long-term preservation.

Those files not using compression (TIFF uncompressed, BMP uncompressed), proved to be much more stable. For one-byte errors, none of the two file formats showed serious problems ( $R_{Bi}$  values of 0.00). Table 2 shows the number of files that totally failed during processing (also in percentage). The reasons for such a phenomenon can be found in corruption of extremely significant data. This always the case for derivative data, structural data or definitional data. As a result, this causes destructive conditional information loss. We already discussed another example for conditional information loss in TIFF files in the section before. Expectedly, the values for  $R_{Bi}$  increase according to the corruption rate.

Nevertheless there are differences, especially with increasing corruption rates. Just for TIFF and BMP uncompressed, there is a clear tendency. BMP uncompressed appears to be quite stable in its file format structure. A closer look at its file format structure shows that BMP is in deed quite simple in it. Most of the lengths and positions of the data fields are predefined. In contrary, TIFF allows for advanced features like to stripe pixel data or to freely choose positions of data fields within the file. File formats which support advanced features tend to be more complex in their structure. This is not surprising since this requires concessions to the processing software. However, complex file formats tend to get into trouble with keeping their data against bit errors.

## Conclusions and Outlook

With the results of this study we give some direction for all those people who are concerned with question of file format and its usability, especially for long-term storage. The choice they make surely depends on factors which are widespread and not only depending on error resilience. To great extent they are often a matter of organizational needs. Despite all that, we regard robustness of file formats against bit corruption as a main factor: As long as it is possible to constantly check files for data integrity, error resilience may be less a hard problem. But consider a scenario where the keepers of the data are not able to do so anymore, may it be because of financial, technical, societal or whatever else shortcomings; then, robustness of file formats against bit corruption is in deed the more crucial.

Robustness Indicator is a simple measure for quantitative analysis of file format data. It does not claim to be a measure of quality analysis. The results we reported in here are part of a larger study. In the future we will focus on enlarging the set of measures for file format, also including measures which are already proven as useful for such issues. This will enable us to additionally refine the model of file format data categorization as well as the findings so far.

We will also refine the analysis of the exact data categories responsible for the specific kind of information loss we diagnosed. This is done by in depth analysis of file formats supported by additional test implementation features. This will help us to find a close understanding of the relation between file format and its error resilience.

We also will extend our research on file formats from other domains especially for formats of text or hybrid content. This will validate and/or improve our given data categorization towards a common model of file format data. Additionally, it is expected to reveal so far unidentified impact of file formats on data integrity for those domains.

## References

- [1] Bradley, K., Risks Associated with the Use of Recordable CDs and DVDs as Reliable Storage Media in Archival Collections: Strategies and Alternatives. UNESCO (2006). <http://unesdoc.unesco.org/images/0014/001477/147782E.pdf>
- [2] Buckley A., JPEG2000, A Practical Digital Preservation Standard? DPC Technology Watch Series Report 08-01 (2008). <http://www.dpconline.org/docs/reports/dpctw08-01.pdf>
- [3] Buenora, P., Long lasting digital charters. Storage, formats, interoperability, Presentation held on Digital Diplomats, Munich (2007). <http://www.cflr.beniculturali.it/Progetti/FixIt/Munich.ppt>
- [4] DPC/BL Joint JPEG 2000 workshop, June 2007. <http://www.dpconline.org/graphics/events/0706jpeg2000wshop.html>
- [5] Iraci J., The relative stabilities of Optical Disk Formats, Restaurator, Vol.26, Number2 (2005).
- [6] Matsumoto M., Nishimura T., Mersenne Twister: A 623-dimensionally equidistributed uniform pseudorandom number generator, ACM Trans. on Modeling and Computer Simulation Vol. 8, No. 1, January, pp.3-30 (1998).
- [7] Panzer-Steindel, B., Data Integrity, CERN/IT (2007). <http://indico.cern.ch/getFile.py/access?contribId=3&sessionId=0&resId=1&materialId=paper&confId=13797>
- [8] Rog, J., Compression and Digital Preservation: Do They Go Together?, Proceedings of Archiving 2007 (2007).
- [9] Rog, J., van Wijk, C., Evaluating File Formats for Long-term Preservation, National Library of the Netherlands, (2008). [http://www.kb.nl/hrd/dd/dd\\_links\\_en\\_publicaties/publicaties/KB\\_file\\_format\\_evaluation\\_method\\_27022008.pdf](http://www.kb.nl/hrd/dd/dd_links_en_publicaties/publicaties/KB_file_format_evaluation_method_27022008.pdf)
- [10] Rosenthal, D. S., Reich, V., Permanent Web Publishing (2000). <http://lockss.stanford.edu/freenix2000/freenix2000.html>
- [11] Santa-Cruz D., Ebrahimi T., Askelof J., Larsson M., Christopoulos C.A., JPEG 2000 still image coding versus other standards, Proceedings of SPIE, Vol. 4115, pp. 446-454 (2000).
- [12] Schroeder B., Gibson G.A., Disk Failures in the Real World, 5th USENIX Conference on File and Storage Technologies, San Jose, CA (2007). <http://www.cs.cmu.edu/~bianca/fast07.pdf>
- [13] Thaller, M., Preserving for 2106, 3006. Or: Is there a life for an object outside a digital library? Presentation held at 'DELOS Conference on Digital Libraries and Digital Preservation', Tallin, Estonia (2006). <http://www.hki.uni-koeln.de/events/tallinn040906/tallinn040906.ppt>
- [14] XCL. Extensible Characterisation Language (2007). <http://planetarium.hki.uni-koeln.de/XCL/>

## Author Biography

*Volker Heydegger is a PhD researcher at the University of Cologne, department of Computer Science for the Humanities (Historisch-kulturwissenschaftliche Informationsverarbeitung). Since 2004 he is working in European research projects in the field of digital preservation, currently for PLANETS. His research focus is on characterisation of file format content for digital preservation and on preservation aspects of file formats.*



---

# XEROX® GLOBAL SERVICES

Technology | Document Management | Consulting Services

## JPEG 2000 Profile for the National Digital Newspaper Program

Library of Congress  
Office of Strategic Initiatives

Date: April 27, 2006

Prepared by: Robert Buckley, Research Fellow, Xerox Innovation Group 585.422.1282  
Roger Sam, Business Development Executive, 202.962.7708



# Table of Contents

---

1	Executive Summary .....	3
2	Introduction .....	4
3	JPEG2000 Codestream Parameters .....	5
3.1	Compressed File Size.....	5
3.2	Progression Order and Resolution Levels .....	6
3.3	Layers .....	7
3.4	Tiles and Precincts.....	8
3.5	Coding Efficiency.....	9
4	JP2 File Parameters .....	10
5	Metadata.....	11
6	Tools.....	13
	Appendix A: JP2 File Format Overview .....	14
	Appendix B: Sample Files .....	15
	Appendix C: Sample Content Areas .....	18
	Appendix D: OCR Results .....	20
	Appendix E: Test Target Measurements.....	21
	Appendix F: Sample Metadata .....	23
	Bibliography.....	24

---

# 1 EXECUTIVE SUMMARY

---

This report documents the JPEG2000 file and codestream profiles for use in production masters during Phase 1 of the National Digital Newspaper Program (NDNP). NDNP, a joint collaborative program of the National Endowment for the Humanities and the Library of Congress, is intended to provide access to page images of historical American newspapers. Web access will be provided through the use of JPEG2000 production masters. For these masters, this report recommends using a visually lossless, tiled JPEG2000-compressed grayscale image, with multiple resolution levels and multiple quality layers, encapsulated in a JP2 file with Dublin Core-compliant metadata.

This report was prepared for the Office of Strategic Initiatives at the Library of Congress by Xerox Global Services, with inputs from the Imaging and Services Technology Center of the Xerox Innovation Group. A preliminary version of this report was available in April 2005 and its results were shared with awardees at a technical review meeting in May 2005. The profile described here has been in use since then.

## **Client Information**

### *LOC Contact:*

**Name** Michael Stelmach  
**Company** Library of Congress  
**Address** 101 Independence Ave SE, MS LA-G05  
Washington, DC 20540-1310  
**Telephone** 202-707-5726  
**Fax**  
**E-Mail** [mste@loc.gov](mailto:mste@loc.gov)

### *Xerox Contact:*

**Name** Roger Sam  
**Company** Xerox Global Services  
**Address** 1301 K Street NW Suite 200  
Washington, DC 20005  
**Telephone** 202-962-7708  
**Fax** 202-962-7926  
**E-Mail** [roger.sam@xerox.com](mailto:roger.sam@xerox.com)

---

## **Change History**

<i>Version</i>	<i>Date</i>	<i>Comment</i>
1.0	Jan 18, 2006	Version 1.0 of the report issued
2.0	Apr 27, 2006	Clarified format use in Section 3.1; added text on coding efficiency in Section 3.5; clarified the effect of code-block size on coding efficiency; added this Change History

## 2 INTRODUCTION

---

The National Digital Newspaper Program (NDNP) is a collaborative program of the National Endowment for the Humanities and the Library of Congress. It is intended to enhance access to historically significant American newspapers. NDNP will provide web access to a national directory of US newspaper holdings and to millions of page images.

According to the Technical Guidelines for Applicants [1], each newspaper page image will be supplied in two raster formats:

- An uncompressed grayscale TIFF 6.0 file, usually at 400 dpi
- A compressed JPEG2000 (JP2) file

JP2 is the image file format defined in Part 1 of the JPEG2000 standard [2]. A JP2 file contains a JPEG2000 codestream along with the image parameters, file information and metadata needed to interpret and render the JPEG2000-compressed image data. Appendix A contains an overview of the JP2 file format.

According to the Library of Congress and descriptions in the Technical Guidelines, the uncompressed TIFF file will be the master page image and the JP2 file will serve as the surrogate for day-to-day use and client access. The JP2 file will provide a high-quality, low-bandwidth, scalable version of the same image that is stored in uncompressed form in a large TIFF file. During the initial phase of NDNP, JP2 file access will use services created with a software development kit from Aware Inc.

Since the TIFF files are the archival masters, the JP2 production masters will be derived from them by means of a conversion process. This process will include image compression and file export using the compression options and file format parameters described in this report. The JP2 files that each awardee institution will provide should be compatible with the profile defined in this report.

This profile was derived with reference to TIFF file samples provided by the Library of Congress. The content and characteristics of typical files from this set are given in Appendix B.

The profile defined in this document covers:

1. Codestream: JPEG 2000 coding parameters, such as wavelet filter, number of decomposition levels and progression order. These are the parameters specified when applying JPEG 2000 compression to the image.
2. File: Image coding parameters, such as color space, spatial resolution and image size, which are independent of JPEG 2000. To the extent that the conversion will start with existing TIFF image files, these parameters in many cases are already defined and not design parameters
3. Metadata: Metadata, such as image description and keywords for search. These are based on current practices.

The next three sections describe these aspects of the NDNP JPEG2000 profile. They are followed by a section that describes the tools used in the development of the profile. Appendices A through F provide supporting information on the file format and sample images.

### 3 JPEG2000 CODESTREAM PARAMETERS

This profile defines a lossy compressed image with the goal of no objectionable visual artifacts and acceptable OCR performance, based on results obtained using the sample images. The JPEG2000 codestream in the production master shall contain a single-component 8-bit image with the same image size as the corresponding TIFF archival master. The compression ratio shall be eight to one. The codestream shall have 6 decomposition levels and multiple layers; in particular, this profile recommends 25 layers. The progression order shall be RLCP, which is resolution major. The codestream shall be tiled. The codestream does not use precincts or contain regions of interest.

Table 1 gives the codestream restrictions and characteristics for the JPEG2000 Production Master with reference to the markers defined in Annex A of Part 1 of the JPEG2000 standard [2]. This profile corresponds to Profile-1 of the JPEG2000 standard and would require a Cclass 2 decoder to process. The remaining parts of this section explain the rationale for the choices made.

<b>Table 1: JPEG 2000 Production Master Codestream Profile</b>	
<b>Parameter</b>	<b>Value</b>
<b>SIZ marker segment</b>	
Profile	Rsiz=2 (Profile 1)
Image size	Same as TIFF master
Tiles	1024 x 1024 (Section 3.4)
Image and tile origin	XO <sub>siz</sub> = YO <sub>siz</sub> = XT <sub>Osiz</sub> = YT <sub>Osiz</sub> = 0
Number of components	Csiz = 1
Bit depth	Ssiz = 8
Subsampling	XR <sub>siz</sub> = YR <sub>siz</sub> = 1
<b>Marker Locations</b>	
COD, COC, QCD, QCC	Main header only
<b>COD/COC marker segments</b>	
Progression Order	RLCP
Number of decomposition levels	N <sub>l</sub> = 6 (Section 3.2)
Number of layers	Multiple (Section 3.3)
Code-block size	xcb=ycb=6
Code-block style	SP <sub>cod</sub> , SP <sub>coc</sub> = 0000 0000
Transformation	9-7 irreversible filter
Precinct size	Not used (Section 3.4)
<b>Compressed File size</b>	About one-eighth of TIFF master or 1 bit per pixel (Section 3.1)

#### 3.1 COMPRESSED FILE SIZE

One of the first choices to be made is how much compression should be used in the production master. Since the uncompressed TIFF file from the scan will be retained as the archival master during Phase 1 of the National Digital Newspaper Program, there is no a priori requirement for the JP2 production master to be lossless; visually lossless may be sufficient. Visually lossless means that while it is not possible to exactly reconstruct the original from the compressed image, the differences are either not noticeable, not significant or do not adversely affect the intended uses of the production master. In



this case, the intended uses of the production master are viewing, printing and possibly text recognition.

To judge the effect of compression on visual screen appearance, a series of images was generated by applying different compression ratios to selected test images. In particular, representative image (halftone), line art and text areas were selected from agreed upon test images. For each area, ten images were provided for viewing and evaluation: an uncompressed original and nine variants compressed to 2, 1.33, 1, 0.75, 0.67, 0.5, 0.4, 0.32 and 0.25 bits per pixel, corresponding to compression ratios of 4, 6, 8, 10.67, 12, 16, 20, 25 and 32 to 1<sup>1</sup>. Appendix C shows the images of the test areas.

The compressed image samples were delivered to the Library of Congress for their review and to establish quality thresholds in terms of their application. The 4:1 and 6:1 compressed images were judged visually lossless when viewed on a screen; only experienced viewers could locate compression artifacts in these images. The image quality exhibited by 8:1 and 10.67:1 was judged preferable. Image quality at 16:1 was acceptable, although it was felt that the artifacts and loss of resolution could make extended reading uncomfortable. Even the image quality at 32:1 was judged useable for some purposes.

The evaluations focused on text quality; halftone quality was not judged to be as important. Very little difference was noted in the printouts over the varying quality levels. The conclusion was that print quality was adequate, pending further analysis, but that it was less important than visual screen presentation quality.

As a result of these observations, the decision was made to use 8:1 compression in the profile for the JP2 production masters. This was judged a good compromise between file size and image quality. Since it was noted that higher compression ratios may be acceptable for some purposes, layers were introduced to make it possible to easily obtain images at a range of compression ratios between 8:1 and 32:1 (bit rates from 1 to 0.25 bits per pixel).

While screen viewing is the most important application, applying OCR to the production master may be an option. Some simple OCR studies were performed on text areas from the sample images. Differences were found when comparing the results of OCR applied to uncompressed and compressed images. In most cases, the OCR results from the 8:1 compressed images were *better* than were those obtained using the uncompressed images. The results are reported in Appendix D. More definitive conclusions require a study using a wider range of sample images and especially the same OCR tools as the Library of Congress uses or recommends.

### 3.2 PROGRESSION ORDER AND RESOLUTION LEVELS

The Library of Congress viewer was built assuming resolution-major progression. Further, it was assumed the codestream would be organized to make it easy to extract low resolution images that could be used for thumbnails or as a navigation pane. The two resolution-major progression orders defined in the JPEG2000 standard are RLCP (Resolution level-layer-component-position) progression and RPCL (Resolution level-position-component-layer) progression. Of these two, the profile specifies

---

<sup>1</sup> Compressing an area extracted from a page image to a target bit rate will not give the same results as compressing the page image to the same target bit rate and then extracting the area for evaluation. The differences however were found to be negligible at 8:1 compression for the selected areas and much less than the differences between the 8:1 and either the 6:1 or 10.67:1 compression ratio images.



RLCP progression so that within a resolution layer, the codestream is progressive by layer, i.e. by quality.

The number of resolution levels was selected so that the lowest resolution level gives a thumbnail of the desired size for a typical-size page image. This profile assumes that the lowest resolution level will generate QVGA-sized or smaller image: a QVGA or Quarter VGA image is 320 pixels wide by 240 pixels high. For a sample image that is 6306 by 8997, like the sample image in Appendix B.2, specifying five resolution levels means that the smallest resolution level image will be about 280 pixels high by 200 pixels wide. Because there can be page images larger than this sample, this profile specifies six resolution levels.

### 3.3 LAYERS

The Library of Congress observed that higher compression ratios may be acceptable for some purposes. Using layers makes it possible to obtain reduced-quality, full-resolution versions of the production master with compression ratios between 8:1 and 32:1, equivalent to bit rates between 1 and 0.25 bits per pixel. In particular, layers are introduced that correspond to compressed bit rates of 1, 0.84, 0.7, 0.6, 0.5, 0.4, 0.35, 0.3, 0.25 bits per pixel, which are equivalent to compression ratios of 8, 9.5, 11.4, 13.3, 16, 20, 22.9, 26.7 and 32.

At maximum quality and maximum resolution, all layers would be decompressed. However, at lower resolutions, higher compression ratios are possible without objectionable visual artifacts. Additional layers optionally allow higher compression ratios at lower resolutions.

Altogether, this profile specifies 25 layers that cover the range from 1 to 0.015625 bits per pixel, or the equivalent compression ratio range of 8:1 to 512:1. The layers are specified in terms of bit rate; the bit rates and corresponding compression ratios (CR) are given in Table 2.

Layer	Bit rate	CR
1	1.0	8.0
2	0.84	9.5
3	0.7	11.4
4	0.6	13.3
5	0.5	16.0
6	0.4	20.0
7	0.35	22.9
8	0.3	26.7
9	0.25	32.0
10	0.21	38.1
11	0.18	44.4
12	0.15	53.3
13	0.125	64.0
14	0.1	80.0
15	0.088	90.9
16	0.075	106.7
17	0.0625	128.0
18	0.05	160.0

19	0.04419	181.0
20	0.03716	215.3
21	0.03125	256.0
22	0.025	320.0
23	0.0221	362.0
24	0.01858	430.6
25	0.015625	512.0

The bit rates for the layers were selected so that the logarithms of the bit rates (or the compression ratios) are close to being uniformly distributed between the maximum and minimum values. Figure 1 plots the bit rates and compression ratios in Table 2 against layers.

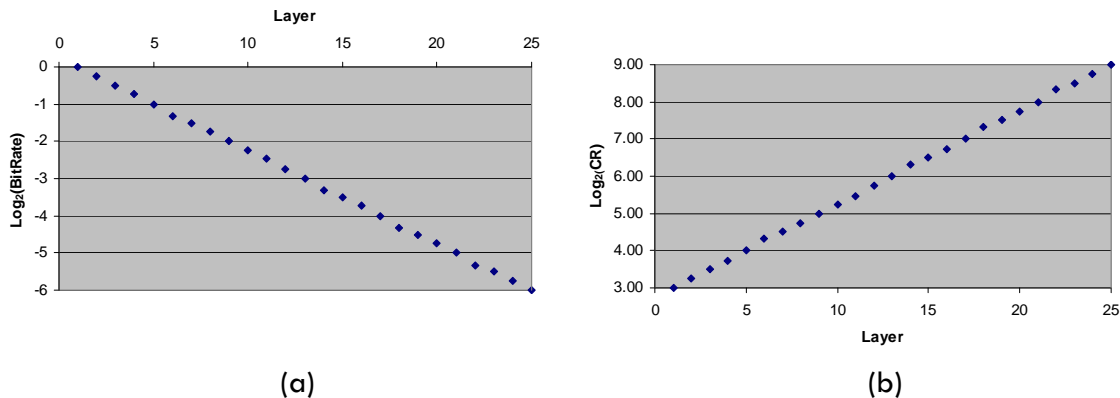


Figure 1: Plots of (a) Log<sub>2</sub> of Bit Rate and (b) Log<sub>2</sub> of Compression Ratio (CR) vs. Layer

The exact bit rate values or compression ratios are not critical. What is more important is the range of values and there being sufficient values to provide an adequate sampling within the range.

Within a fixed bit budget, the more layers there are, the more bits are needed to signal the layer structure and the fewer bits are available to represent the compressed image data. As a result, the greater the difference between the original and the image decompressed from the compressed image. However, for the sample images in Appendix B, the differences between using 1 layer and using 25 were negligible. Compressing with 25 layers instead of 1 layer was about the same as using a compression ratio of a little less than 8.1 instead of a compression ratio of 8. This was judged a relatively small price to pay to obtain the advantages of quality scalability.

### 3.4 TILES AND PRECINCTS

Tiles and precincts are two ways of providing spatial addressability within the codestream. With spatial addressability, it is possible to extract and decompress a portion of the codestream, corresponding to a region in the image. In effect, this means an application can crop and decompress the codestream, which is more efficient than decompressing the codestream and then cropping out the desired region from the decompressed image. Along with resolution and quality scalability, spatial addressability is an important feature of JPEG2000.

Test images were generated with tiles and with precincts. In the images that had tiles, the tile size was 1024x1024. In the images that had precincts, the precinct size was 256x256 for the two highest resolution levels and 128x128 for the remaining levels.

In tests conducted by the Library of Congress, it was found that the Aware codec decoded images with tiles significantly faster than images with precincts. As a result, tiling was judged to be the preferred solution for decoding with Aware. Therefore this profile specifies the use of 1024x1024 tiles. The tile X and Y origins as well as the image X and Y origins are set to 0. The main header contains the Coding style default (COD), Coding style component (COC), Quantization default (QCD) and Quantization component (QCC) marker segments. Because these marker segments do not occur in the tile-part headers, the quantization and coding parameters are the same for all tiles. Also, the progression order is the same for all tiles.

### 3.5 CODING EFFICIENCY

Coding efficiency is a measure of the ability of a coder or coder option to compress an image. The more efficient a coder or option is, the smaller the file size for a given quality, or the higher the quality for a given file size. An objective measure of quality is PSNR, the peak signal-to-noise ratio, which is typically reported in dB<sup>2</sup>. Noise in this case is the error or difference between the original uncompressed image and the decompressed image.

Code Blocks: The quantized wavelet coefficients are coded independently by code block. The JPEG2000 standard limits the maximum size of a square code block to 64x64. There is nothing to recommend smaller code-block sizes, which are less efficient since smaller code blocks mean more overhead information in the file and less opportunity for the adaptive arithmetic coder to adapt to the statistics of the signal it is compressing. Therefore, this profile specifies 64x64 code blocks.

Bypass mode: The JPEG2000 coder uses an arithmetic coder that operates bit plane by bit plane, making multiple coding passes over each bit plane. For some images, the statistics of the least significant bit planes are such that there is little compression to be had with the arithmetic coder. In these cases, bypassing the arithmetic coder for some coding passes in less significant planes can speed up the coding (and decoding) with little loss in coding efficiency. However, for the sample images in Appendix B, the loss in efficiency was between 0.16 and 0.20 dB, which is up to twice the loss that comes from using 25 layers instead of 1. This profile does not specify the use of Bypass mode for coding, although the use of bypass mode can be revisited after more experience has been gained during the initial phases of NDNP.

Wavelet Filter: Part 1 of the JPEG2000 standard [2] defines two transformation types: a 9-7 irreversible filter and a 5-3 reversible filter. While the 5-3 filter is simpler to implement than the 9-7 filter, it is also less efficient. For the sample images in Appendix B, the loss in efficiency is a little over 1 dB. This was judged too high a cost in quality for only a slight gain in decoding speed. Therefore, this profile specifies the use of the 9-7 filter.

---

<sup>2</sup> PSNR in dB is 20 times the logarithm base 10 of the ratio of the peak signal (255 in this case) to the root mean squared error, where the error is the difference between the original and decompressed images.

## 4 JP2 FILE PARAMETERS

In a JP2 production master, the codestream specified in the previous section is embedded in a JP2 file. The JP2 file format is specified in Annex I of the JPEG2000 standard [2]. Appendix A contains an overview of the file format. A JP2 production master shall contain a JPEG2000 Signature box, a File Type box, a JP2 Header box and a Contiguous Codestream box. Table 3 gives the values for the data fields of these JP2 file boxes. The file shall also contain at least one metadata box, whose contents are described in the next section.

<b>JP2 Box/Data Field</b>	<b>Value</b>
JPEG2000 Signature Box	<CR><LF><0x87><LF> <sup>3</sup>
File Type Box	
Brand	JP2
Version	0
Compatibility	'jp2 '
JP2 Header Box	
Image Header Box	
HEIGHT	TIFF ImageLength
WIDTH	TIFF ImageWidth
NC	1
BPC	Unsigned 8 bit
C	JPEG2000
UnkC	0
IPR	0
Colour Specification Box	
Method	Enumerated Colour Space (1) or Restricted ICC Profile (2)
PREC	0
Approx	0
EnumCS	17 (Greyscale) if Method = 1
Profile	Monochrome Input Profile if Method = 2
Resolution Box	
Capture Resolution Box	
Vertical	TIFF YResolution, converted to pixels/m
Horizontal	TIFF XResolution, converted to pixels/m
Contiguous Codestream Box	Codestream specified in Section 3

The JP2 production master shall contain an image with the same size and resolution as the image in the corresponding TIFF archival master. In particular, the JP2 production master file will be prepared after any image processing or clean-up and will correspond with the image that is used for OCR.

The image data in the JP2 production master is expected to have the same photometry (TIFF photometric interpretation) as the corresponding scanned TIFF file.

<sup>3</sup> In hexadecimal notation, the value of this field is 0x0D0A 870A.

To assess the photometry used in scanned images, the Library of Congress provided the image of a test target scanned on a microfilm scanner. A separate source provided the nominal densities of the patches on the target. The results of analyzing the scanned image of the test target are given in Appendix E. They show that the plot of target reflectance against pixel value is a straight line up to a pixel value of 250<sup>4</sup>. Encoding this relationship in a JP2 file requires using an ICC Monochrome Input Profile, as defined in Section 6.3.1.1 of the ICC Profile Format Specification [3].

The Monochrome Input Profile uses a one-dimensional lookup table to relate the input device values, in this case, the pixel values, to the luminance value Y of the XYZ Profile Connection Space. The JPEG2000 standard references the 1998 version of the ICC Profile Format Specification. There have been four major and minor revisions to the specification since, although there has been little change to the definition of the Monochrome Input Profile.

If gamma-corrected image data is available or can easily be generated, then a preferred method for representing the gray values in the JP2 production master is the enumerated grayscale color space defined in Annex I.5.3.3 of the JPEG2000 standard [2]. This grayscale color space is the gray analog of sRGB and codes the luminance using the same non-linearity as sRGB. The 8-bit pixel value D is converted to luminance Y using the following equations:

$$Y' = D / 255$$

$$Y = Y' / 12.92 \quad Y' \leq 0.04045$$
$$= ((Y' + 0.055) / 1.055)^{2.4} \quad Y' > 0.04045$$

These equations assume that a pixel value of 255 corresponds to a luminance value of 1.0, which is white. On this scale, a luminance value of 0.4 corresponds to a pixel value of about 170, which would provide a facsimile of the somewhat gray microfilmed image. If some scaling is used to make use of the full pixel value range, then the scaled signal would not follow the definition of the enumerated grayscale color space, but nevertheless could be represented by a Monochrome Input Profile.

## 5 METADATA

---

This profile requires that a JP2 production master contain metadata identifying the image content and provenance. This metadata would be in an XML box and would use Dublin Core elements to identify:

- File format
- Title of the newspaper
- Location of publication
- Date of Publication
- Page Label
- Description, including LCCN (Library of Congress Control Number)

---

<sup>4</sup> All the pixel values of the sample images in Appendix B are less than 250; in the case of the image in Appendix B.1, they are less than 240.

The template for the data field of the XML box with this metadata is shown below. This template was proposed by the Library of Congress for images scanned from microfilm. The application that writes the JP2 production master would supply the information between the sharp (#) characters in the template. An example of the use of this template is given in Appendix F.

```
<?xml version="1.0"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdfsyntax-ns#">
  <rdf:Description rdf:about="urn:library-
ofcongress:ndnp:mets:newspaper:page://#The normalized LCCN#/#Date of publication
in CCYY-MM-DD#/#Edition order#/#Page sequence number#"
  xmlns:dc="http://purl.org/dc/elements/1.1/">
    <dc:format>image/jp2</dc:format>
    <dc:title>
      <rdf:Alt>
        <rdf:li xml:lang="en">#The title of the newspaper#.(#Location of
publication#) #Date of publication in CCYY-MM-DD# [p #page label#]</rdf:li>
      </rdf:Alt>
    </dc:title>
    <dc:description>
      <rdf:Alt>
        <rdf:li xml:lang="en">Page from #The title of the newspaper#
(newspaper). [See LCCN: #The normalized LCCN# for catalog record.]. Prepared on
behalf of #responsible organization#. </rdf:li>
      </rdf:Alt>
    </dc:description>
    <dc:date>
      <rdf:Seq>
        <rdf:li xml:lang="x-default">#Date of publication in CCYY-MM-DD#
        </rdf:li>
      </rdf:Seq>
    </dc:date>
    <dc:type>
      <rdf:Bag>
        <rdf:li xml:lang="en">text</rdf:li>
        <rdf:li xml:lang="en">newspaper</rdf:li>
      </rdf:Bag>
    </dc:type>
    <dc:identifier>
      <rdf:Alt>
        <rdf:li xml:lang="en">Reel number #The reel number#. Sequence
number #The sequence number#</rdf:li>
      </rdf:Alt>
    </dc:identifier>
  </rdf:Description>
</rdf:RDF>
```

It would be useful if a JP2 production master also contained a reference to the TIFF archival master from which it was derived.

This profile also recommends using technical elements from the NISO Z39.87 standard [4]. Besides the mandatory elements defined in that standard, the metadata should include JPEG2000-specific information. This metadata would be contained in an XML box using the MIX schema [5].



The current draft of the NISO Z39.87 standard defines a container with JPEG2000 format-specific data. The JPEG2000 information is comprised of two containers of data elements: CodecCompliance and EncodingOptions.

- CodecCompliance
  - codec: Specific software implementation of JPEG2000 used to compress the file or codestream
  - codecVersion: version of codec used
  - codestreamProfile: P1 (Profile 1)
  - complianceClass: C2 (Cclass 2)
- Encoding Options
  - tiles: 1024x1024
  - qualityLayers: 25
  - resolutionLevels: 6

At the time this was written, the draft NISO Z39.87 standard referred to here was in the process of being reballoted.

## 6 TOOLS

---

The JP2 files used to develop this profile were generated using Kakadu Version 4.2. While Kakadu was also used for decompression, some decompression and analysis were performed using LuraWave SmartDecompress Version 2.1.05.02.

The Kakadu command line that generates a JP2 file with the codestream described in Section 3 is:

```
kdu_compress -i in.pgm -o out.jp2 -rate  
1,0.84,0.7,0.6,0.5,0.4,0.35,0.3,0.25,0.21,0.18,0.15,0.125,0.1,0.088,0.  
075,0.0625,0.05,0.04419,0.03716,0.03125,0.025,0.0221,0.01858,0.015625  
Clevels=6 Stiles={1024,1024} Corder=RLCP Cblk={64,64} Sprofile=1
```

## APPENDIX A: JP2 FILE FORMAT OVERVIEW

The JP2 file format is defined in Annex I of Part 1 of the JPEG2000 Standard [2]. JP2 is a file format for encapsulating JPEG2000-compressed image data. Applications that can read the JP2 format and access the JPEG2000 codestream it contains are in a position to take advantage of the features and capabilities of JPEG2000, which enable progressive display, scalable rendering and “Just-In-Time” imaging.

A JP2 file consists of a series of elements called “boxes.” Each box has 3 fields: a length field, a type field and a data field, which is interpreted according to the value of the type field. A JP2 file compatible with the profile defined in this report contains the following boxes:

- JPEG2000 Signature, which identifies the file as a member of the JPEG2000 file format family; it has a fixed-value data field.
- File Type, which identifies the file as a JP2 file and contains the version number along with any applicable compatibility and profile information.
- JP2 Header, which specifies image parameters, such as image size, bit depth, spatial resolution and color space. The JP2 Header box is a superbox, a box whose data field consists of the following boxes:
  - Image Header, which gives the height, width, number of components and bits per component of the image and identifies the compression type.
  - Colour Specification, which defines how an application should interpret the color space of the decompressed image.
  - Resolution, which it itself a superbox, whose data field contains the Capture Resolution box, which specifies the resolution at which the image was captured.
- Contiguous Codestream, which contains a single JPEG2000 codestream, compliant with Part 1 of the JPEG2000 standard
- XML box, which contains XML-encoded metadata

The structure and order of boxes in the JP2 file documented in this report is shown in Figure A-1.

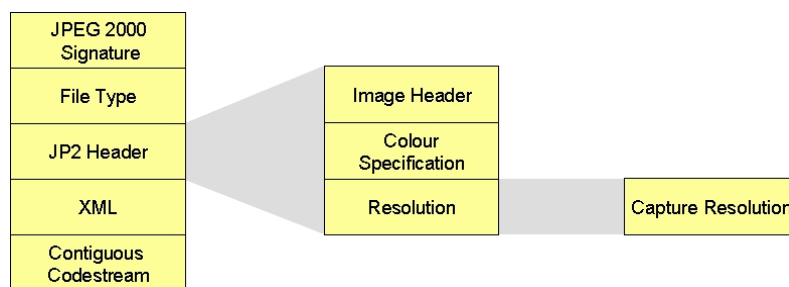


Figure A-1. JP2 File Structure

The JPEG2000 standard (Annex I.2.2 of [2]) requires that the JPEG 2000 Signature box be the first box in the JP2 file and that the File Type box immediately follow it. It also requires that the JP2 Header box come before the Contiguous Codestream box. This profile recommends that, for faster access to metadata, XML boxes come before the Contiguous Codestream box as well.

## APPENDIX B: SAMPLE FILES

---

96 sample TIFF files were provided for testing purposes. They represented two sample sets of 24 newspaper pages each, scanned from microfilm by two vendors. From this image set, a subset was selected for developing the profile described in this report. Two representative images of this subset, one from each vendor, are shown here, along with the listings of the original TIFF files.

**B1. SAMPLE FILE 1**

Sample: \oclc\sn82015056\00000013.TIF



Tag	Value
41728 (10 ASCII)	microfilm
42016 (13 ASCII)	00000013.TIF
SubFileType (1 Long)	Zero
ImageWidth (1 Short)	5231
ImageLength (1 Short)	6861
BitsPerSample (1 Short)	8
Compression (1 Short)	Uncompressed
Photometric (1 Short)	MinIsBlack
DocumentName (27 ASCII)	Reel 00100493068
Make (9 ASCII)	NextScan
Model (8 ASCII)	Eclipse
StripOffsets (6861 Long)	8, 5239, 10470, 15701, 20932, 26163, 31394,...
Orientation (1 Long)	TopLeft
SamplesPerPixel (1 Short)	1
RowsPerStrip (1 Short)	1
StripByteCounts (6861 Long)	5231, 5231, 5231, 5231, 5231, 5231, 5231,...
XResolution (1 Rational)	300
YResolution (1 Rational)	300
PlanarConfig (1 Short)	Contig
ResolutionUnit (1 Short)	Inch
Software (20 ASCII)	Fusion Version 1.22
DateTime (20 ASCII)	2004:06:09 17:52:44
Artist (57 ASCII)	Library of Congress; OCLC Preservation Servic...

**B2. SAMPLE FILE 2**

Sample iarchives\sn82015056\00000003.tif



Tag	Value
ImageWidth (1 Long)	6306
ImageLength (1 Long)	8997
BitsPerSample (1 Short)	8
Compression (1 Short)	Uncompressed
Photometric (1 Short)	MinIsBlack
DocumentName (41 ASCII)	shington-evening-times-19050717-19050826
Make (9 ASCII)	NextScan
Model (24 ASCII)	Phoenix Rollfilm Type 2
StripOffsets (1125 Long)	9458, 59906, 110354, 160802, 211250, 261698,...
Orientation (1 Short)	TopLeft
SamplesPerPixel (1 Short)	1
RowsPerStrip (1 Long)	8
StripByteCounts (1125 Long)	50448, 50448, 50448, 50448, 50448, 50448,...
XResolution (1 Rational)	400
YResolution (1 Rational)	400
ResolutionUnit (1 Short)	Inch
Software (31 ASCII)	iArchives, Inc. imgPrep v3.001
DateTime (20 ASCII)	2004:10:12 11:18:33
Artist (16 ASCII)	iArchives, Inc.
41728 (10 ASCII)	microfilm
42016 (25 ASCII)	SN82015056/1905/00000003

## APPENDIX C: SAMPLE CONTENT AREAS

---

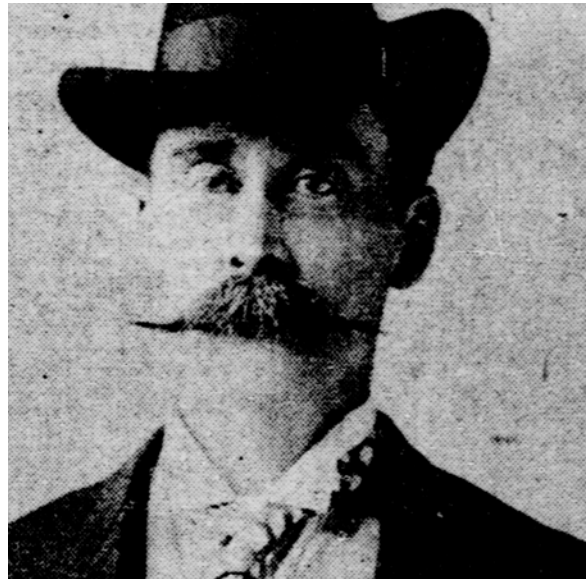
Quality assessments were made with respect to the image (halftone), line art and text areas, selected from the sample images in Appendix B and shown in this appendix.

### Image (Halftone) Area

Source: iarchives\sn82015056\00000003.tif

Crop coordinates:

Left	3088
Top	752
Right	4112
Bottom	1776



### Line art Area

Source: iarchives\sn82015056\00000003.tif

Crop coordinates:

Left	2656
Top	7408
Right	3680
Bottom	8432





Text Area

Source: \oclc\sn82015056\00000013.TIF

Crop coordinates:

Left	1047
Top	1329
Right	2391
Bottom	3378

It is built high above the cultivated land on either side and rises gradually upward till the fields give place to thickets of dwarf oaks and holly and thorn, with an undergrowth on the stony soil of bracken and bramble. Behind rolls a swelling stretch of weald, and out at one side, just below, lie the faint outlines, as shown by the height of the bushes, of a one time stupendous earthwork. Now it is a series of mere ridges, with dykes on the outer edges, reaching across the countryside; simply a long length of green mounds upon which penetrating roots and earth worms have persevered for centuries in their slow leveling process, to mark where once stood a famous Roman camp. But the white, straight road is as good as ever; it has outlived the camp, it will outlive many things of to-day—an eternal monument to the great road makers of history.

Down below on the lower modern road snorts a motor car; I can see it enveloped in a white cloud of dust; but up here I am alone with the shades of forgotten legions. There is no sense of solitude such as men experience in the unknown wilderness, but you find more a feeling of restfulness and satisfaction that you are following in the footsteps of men who, by building roads, shaped the destiny of nations.

Climbing steadily upward the edge of the ridge turns the square side of the camp and I wondered vaguely under what conditions those dykes were dug. Did our forebears hang onto the flanks of their conquerors, harassing them from every thicket of oak and thorn, or were they willing serving men? But they are dust again and the spring wind whispers of a new life as it rustles the dead fronds of the ruddy bracken, and there is the sharp blow of the axe on the green wood that tells of other life.

A little further on, in the midst of a clump of ash plants, I found it in the shape of an old hedger plying his bill hook with the vigorous skill that comes with experience and is not lost with years. I passed the time of day quired.

"No; I've heerd folks say as they never 'ud heer 'em any more. Us have it that Carter Maunders broke the spell. He wor plowin' back in a bit o' meadow land as lies on the edge o' the camp an' was goin' easy when his share hit up against sum't hard. The horses pulled in the collar a bit an' the plow went fair through it. When he looked down to see what it wor he'd plowed through a big brown pot cram full o' money, an' the coins lay all over the plow an' scattered 'round the furrows. He looked at it for a moment, an' he got that frit that he ran home. Ha, ha! Left his team standin' an' ran home, he did, an' put all the folks in a fluster by sayin' as how he'd plowed up pots o' money more'n he could lift. Us all went back wi' him, an' there they wor, thousands o' coins, little green bronze bits, all over the place. The pars'n took an' mended the pot an' sorted out the coins, an' now the squire has 'em. People as know say that they were the money o' the same folk as bullded this gre't road, an' how mebbe the camp paymaster buried the money, an' wor killed. Still it broke the spell, an' us'll never hear the clank o' armor and the marchin' any more. Young folks laugh at it, but there, they never believe owt but their own eyes an' ears."

He finished with an air of finality, a little contemptuous in his tone of the degeneracy of the rising generation.

Taking up his bill hook again, he moistened his palms in the primitive fashion of the laborer, and with the skill of an expert broke away the smooth barked staves crowned with the clustering buds. He cut carefully at their roots, saving the twists at the bottom for the handle crooks and leaving bare a hideous jagged stump.

"An ugly end for anything," I remarked.

"Walkin' sticks for townfolks," he snorted angrily, "to make 'em look better, people as lives in cooped un

## APPENDIX D: OCR RESULTS

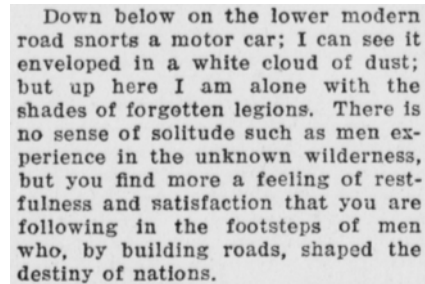
---

### D1. OCR SAMPLE 1

Source: \oclc\sn82015056\00000013.TIF

Crop coordinates:

Left	1050
Top	2200
Right	1710
Bottom	2640



#### OCR results from uncompressed image<sup>5</sup>

Down below on the lower modern road snorts a motor car; I can see It enveloped in a white cloud of dust; but up here I am alone with the shades of forgotten legions. There Is no sense of solitude such as men experience In the unknown wilderness, but you find more a feeling of restfulness and satisfaction that you are following In the footsteps of men who, by building roads, shaped the destiny of nations.

#### OCR results from 8:1 compressed image

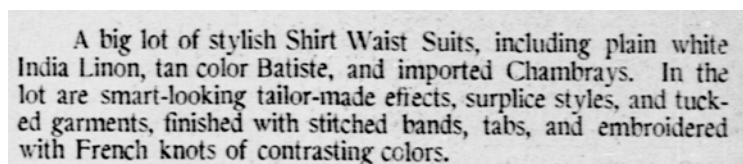
Down below on the lower modern road snorts a motor car; I can see It enveloped in a white cloud of dust; but up here I am alone with the shades of forgotten legions. There Is no sense of solitude such as men experience In the unknown wilderness, but you find more a feeling of restfulness and satisfaction that you are following In the footsteps of men who, by building roads, shaped the destiny of nations.

### D2. OCR SAMPLE 2

Source: iarchives\sn82015056\00000003.tif

Crop coordinates:

Left	4467
Top	5682
Right	6012
Bottom	6012



#### OCR results from uncompressed image

A big lot of stylish Shirt Vaist Suits, iivluding plain white India Linon, tan color Batiste, ani imported Chamhravs. In th lot are smart-looking tailor-made effects, surplice styles, and tucked garments. finished with stitched bands, tabs, and embroidered with French knots of contrasting colors.

#### OCR results from 8:1 compressed image

A big lot of stylish Shirt Waist Suits, including plain white India Linon, tan color Batiste, and imported Chamhravs. In th lot are smart-looking railor-iiade effects, surplice styles, and tucked garments, finished with stitched bands, tabs, and embroidered with French knots of contrasting colors.

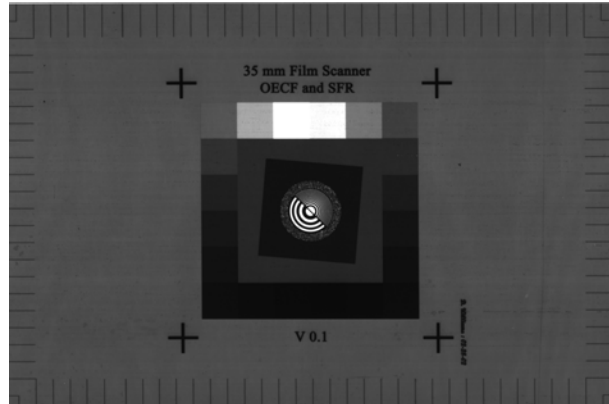
The OCR results were obtained using Microsoft® Office Document Imaging Version 11.0.1897.

---

<sup>5</sup> Text results that are different from those obtained by OCR of the other image and that are incorrect are shown underlined.

## APPENDIX E: TEST TARGET MEASUREMENTS

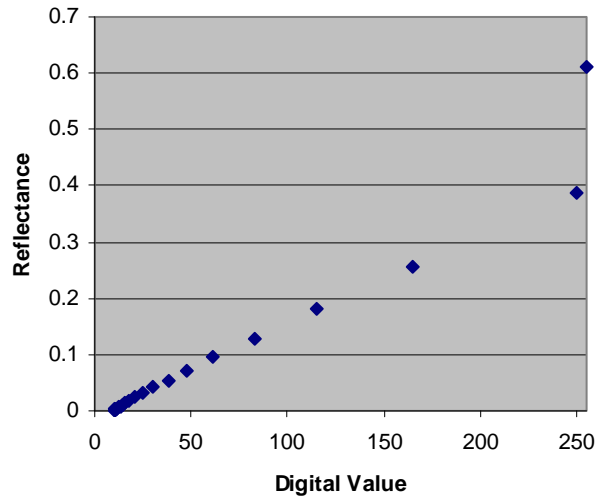
The test target image provided by the Library of Congress is shown below.



The following table shows the average digital values measured for the target's patches, the nominal Status A visual diffuse densities of the patches, and the corresponding reflectance values.

Digital Value	Density	Reflectance
255	0.214	0.611
250	0.411	0.388
165	0.591	0.256
115	0.74	0.182
83	0.888	0.129
61	1.023	0.0948
48	1.153	0.0703
38	1.265	0.0543
30	1.379	0.0418
25	1.506	0.0312
21	1.616	0.0242
18	1.734	0.0185
16	1.846	0.0143
14	1.961	0.0109
13	2.13	0.00741
12	2.24	0.00575
11	2.36	0.00437
10	2.535	0.00292
10	2.7	0.00200
10	2.87	0.00135

The following figure shows the plot of target reflectance against average digital value for the target's patches. The plot is a straight line up to a digital value of 250.



## APPENDIX F: SAMPLE METADATA

---

Following is the data field of a metadata box based on the template in Section 4 for the sample image of Appendix B.1.

```
<?xml version="1.0"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdfsyntax-ns#"
  <rdf:Description rdf:about="urn:library-
ofcongress:ndnp:mets:newspaper:page://sn82015056/1910-05-28/1/13"
  xmlns:dc="http://purl.org/dc/elements/1.1/">
  <dc:format>image/jp2</dc:format>
  <dc:title>
    <rdf:Alt>
      <rdf:li xml:lang="en">The National Forum.(Washington, D.C.) 1910-
05-28 [p 13]</rdf:li>
    </rdf:Alt>
  </dc:title>
  <dc:description>
    <rdf:Alt>
      <rdf:li xml:lang="en">Page from The National Forum (newspaper).
[See LCCN: sn82015056 for catalog record.]. Prepared on behalf of Library of
Congress </rdf:li>
    </rdf:Alt>
  </dc:description>
  <dc:date>
    <rdf:Seq>
      <rdf:li xml:lang="x-default">1910-05-28
      </rdf:li>
    </rdf:Seq>
  </dc:date>
  <dc:type>
    <rdf:Bag>
      <rdf:li xml:lang="en">text</rdf:li>
      <rdf:li xml:lang="en">newspaper</rdf:li>
    </rdf:Bag>
  </dc:type>
  <dc:identifier>
    <rdf:Alt>
      <rdf:li xml:lang="en">Reel number 00100493068. Sequence number
1</rdf:li>
    </rdf:Alt>
  </dc:identifier>
</rdf:Description>
</rdf:RDF>
```

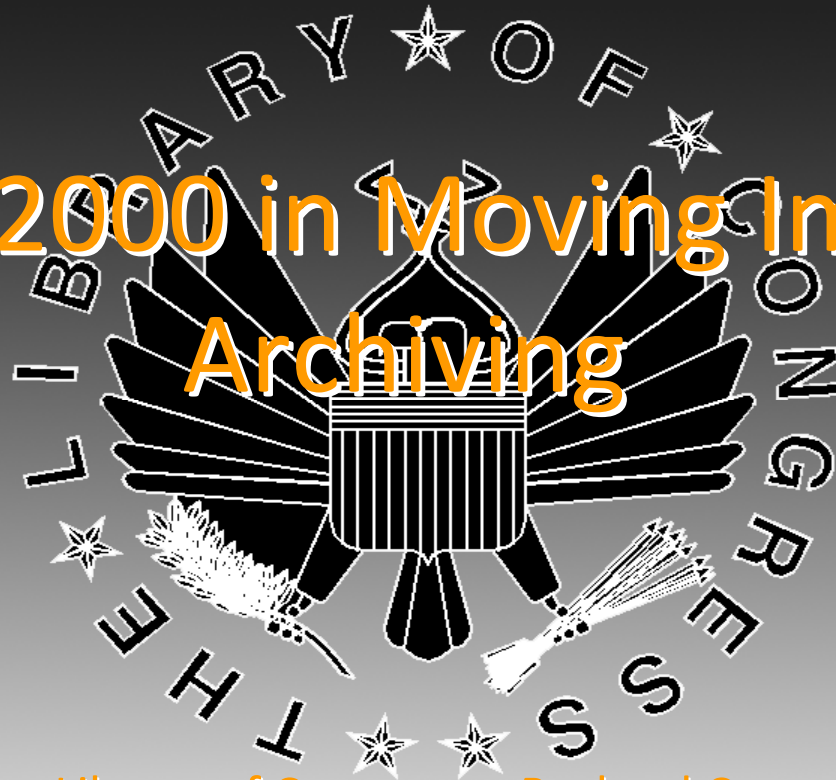
## BIBLIOGRAPHY

---

- [1] The National Digital Newspaper Program (NDNP) Technical Guidelines for Applicants, Phase I, July 2004  
*Note: This document was issued in connection with the Request for Proposals for the Phase I award competition, which is now closed. New guidelines are expected in Summer 2006.*
- [2] ISO/IEC 15444-1:2004, *Information technology -- JPEG 2000 image coding system: Core coding system*
- [3] International Color Consortium, Specification ICC.1:1998-09, File Format for Color Profiles
- [4] National Information Standards Organization and AIIM International, Draft NISO Z39.87-2006/AIIM 20-2006, Data Dictionary – Technical Metadata for Digital Still Images, January 2006
- [5] Library of Congress, Metadata for Images in XML Schema (MIX), <http://www.loc.gov/standards/mix/>, August 30, 2004



# JPEG2000 in Moving Image Archiving



The Library of Congress – Packard Campus  
National Audio Visual Conservation Center

James Snyder

Senior Systems Administrator, NAVCC

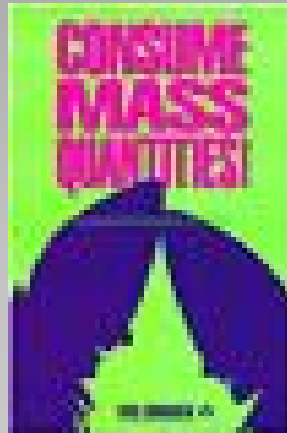
Culpeper, Virginia

# Our Mission: Digitize the Entire Collection

- Bring the 'dark archive' into the 'light'
- Make it more accessible to researchers and the public
  - Stream audio and moving image to the public
  - Without compromising the rights of the © holders
- Required the development of new technologies & processes
- Required new ways of thinking about preservation

# Digitizing Everything

Consume Mass Quantities!



# Overall Digitization Goals

- Archive is essentially a permanent data set
  - We have a different perspective on the word 'longevity'
- Files should be able to stand alone without references to external databases
  - Lots of metadata wrapped in the files!
- Digitize items in their original quality
  - Audio at highest bandwidth in the recording
  - Video at its native resolution
  - Film scanned at resolution equal to the highest available on film
    - 4096 x 3112 for 16mm; 8192 x 6224 for 35mm in development

# Overall Digitization Goals

- Use as much off-the-shelf products as feasible
- Invent items or write custom software only when a commercial product cannot fit the need
- Use industry standards as much as possible
  - Most video formats were created using standards anyway; why reinvent the wheel for preservation?

# Overall Digitization Goals

- Video:
  - JPEG2000 'lossless' (reversible 5x3) (ISO 15444)
  - MXF OP1a
- Film: (planned)
  - JPEG2000 lossless with MXF OP1a wrapper (goal)
  - Short-term expediency: DPX with BWF audio
  - 4k (4096 x 3112) JPEG2000 Lossless encoding now available from vendors
  - Working with vendors to extend to 8k (and beyond, if needed)
- Audio:
  - 96 kHz/24 bit BWF RF64 (Broadcast Wave Format)
  - Limited metadata capabilities: looking at MXF OP1a wrapper for more metadata in audio-only files as well



# Why JPEG2000 Lossless?

The first practical moving image compression standard that doesn't throw away picture content in any way.

# Why JPEG2000?

## JPEG2000

- An international standard (ISO 15444)
- The first, and currently only, standardized compression scheme that has a truly mathematically lossless mode
- No LA (Licensing Agreement) that has issues

## Other compression systems

- No mathematically lossless profiles in any other standardized compression scheme
- Other have licensing agreements that have legal or temporal issues
- Many are vendor specific and thus used at the whims of the manufacturer

# Why JPEG2000?

## JPEG2000

- Unlike MPEG, it's a standard not a toolkit
- Wavelet based
- Can be wrapped in a standardized file wrapper (MXF) which promotes interoperability

## Other compression systems

- MPEG is a toolkit, meaning different implementations exist for the same profile/level & bitrate
- MPEG, DV are DCT based, which results in unnatural boundaries that are very visible to the eye

# Why JPEG2000?

## JPEG2000

- Lossless means no concatenation artifacts created by encoding that aren't already there

## Other compression systems

- All lossy compression schemes suffer concatenation artifacts
- Particularly bad between codecs that compress using different toolkits:
  - MPEG>DV
  - DV>MPEG
  - MPEG-4/H.264 > low bitrate MPEG-2

# Why JPEG2000?

## JPEG2000

- Can accommodate multiple color spaces
  - YPbPr
  - RGB
  - XYZ
  - New ones being worked on
- Can accommodate multiple bit depths
  - Video: 8 & 10 bits/channel
  - Film: 10-16 bits/channel

## Other compression systems

- MPEG-2 is 8 bit ONLY
- MPEG-4: mostly 8 bit but one 10 bit profile for production
- YPbPr color space ONLY

# File Format

## MXF File Wrapper



# Why Not JPEG2000 File?

- Part 3 defines a .jp2 moving image file, but it can't handle the variety of sources required in archiving
- MXF file standard already in progress, with far more flexibility designed into the standard

# Solution:

Wrap JPEG2000 part 3 encoded  
moving image essence into the MXF  
file format

# Why MXF?

# Issue: Interoperability

- How to create files that will work across multiple platforms and vendors seamlessly?
- Most common production file formats today are both vendor specific:
  - .mov = Apple
  - .avi = Microsoft (original Windows video format)
- If the owner of the format decides to make a change or orphan the format: what then?

# Interoperability Solution

- File format standardized by the SMPTE (Society of Motion Picture & Television Engineers) & AMWA (Advanced Media Workflow Association)
- Allows different flavors of files to be created for specific production environments
- Can act as a wrapper for metadata & other types of associated data

# Interoperability Solution

- MXF: major categories are called “operational patterns” (OP)
- More focused subcategories are called “Application Specifications” (AS)
- Our version: OP1a AS-02
- Working on an archive-focused AS called AS-07 (aka AS-AP for Archiving and Preservation)



# How We Implemented

## SAMMA

# SAMMA

- The Library was a driving force in the creation of the first production model JPEG2000 lossless video encoder: the SAMMA Solo
  - Can only do 525i29.97 and 625i25
  - Produces proxy files, but only in post-encoding process
- 31 currently deployed at Culpeper
- SAMMA Sync 'TBC' not really a TBC (time base corrector): it's a frame sync
  - Can't correct the worst videotape problems
  - Sometimes (rarely) injects artifacts into the video!
  - We use the Leitch DPS-575 TBC to correct our analog video problems. Corrects virtually any tape that can be read.

# SAMMA

- The updated HD model premiered at NAB this past April.
  - Will encode both SD and HD and multiple frame rates including film frame rates
  - New SAMMA Sync still not a TBC, but MUCH better than the first version

# Vendor Diversity

- Omneon & Amberfin have teamed up to create a video server based solution where multiple encoders feed one server
  - Can handle SD, HD & 2k at multiple frame rates
  - We will be using this solution for Congressional video
- OpenCubeHD currently shipping an encoding, editing and file creation platform
- DVS premiered 4k (up to 4096 x 3112) JPEG2000 Lossless editing and encoding at NAB in April
  - Including 3D @ 4k

# Feature Diversity

- The entire production & distribution pieces are now in place:
  - Editing
  - Encoding & file creation
  - Metadata creation, editing & insertion
  - Proxy file creation at the same time
- Everything up to 3D

# Future Needs

- Real time 4k encoding and decoding
- Encoding beyond 4k
  - 2011 NAB vendors had UHDTV and 8k film scanning as proposed or shipping products on the show floor
- Encoding of the new color spaces being proposed
- Finalize metadata needs: creation, editing & insertion toolkits in MXF files



# We're Not the Only Ones

Digital Cinema standardized on MXF  
for the distribution of Digital Cinema  
Packages to theatres

QC?

The goal is to QC every file we  
produce

# Automated Software

- Interra Baton has a mature JPEG2000 Lossless automated QC package
  - Real time still a challenge; depends on computational power
  - We will be implementing this solution this year
- Tektronix Cerify is not quite as good as Baton, but getting better
- Digimetrix premiered their package at April's NAB and it shows promise

# Error Detection

How do we know the files are good throughout the system, or later on?

# SHA-1 Checksum

- Cryptographic Hash Checksums are designed to identify one bit flip in an entire file
- SHA-1 can accurately identify 1 bit flip in file sizes up to  $2^{61} - 1$  bits
- First year of production: 800 TiB in the archive: no bit flips

# Where Do We Store All This Material?

# Our Digital Repository

- 200 TiB SAN
  - Staging area for transmission to backup site and the tape library
  - Backup site has identical SAN & tape library
- Tape library
  - StorageTek SL-8500 robot with 9800 slots currently installed; 37,500 total slots planned by ~2015 (expansion depends on requirements)
  - Currently using T10000-B tapes with 1TiB/tape current capacity (9.8PiB available; 37.5PiB total as designed)
  - Moving to new T10000-C tapes @ 5TiB/tape (eventually 49PiB available; 187.5 PiB total as designed)
  - Upgrade path to ~48TiB/tape by ~2019



# The Digital Repository

- SAM-FS file system
- 1.35 PiB on tape as of Monday (5/9/2011)
- Increasing at approx. 20-40 TiB/week
  - 80-100 TiB/month
  - 60% of each month's production is JPEG2000 MXF files by data throughput
  - 20% of month's output is JPEG2000 MXF by file count
  - Total of 29,400 JPEG2000 files as of 5/9/2011
- First ExiB anticipated around or after 2020

Why T10000 tape?

Bit error rate matters!

# Digital Repository Requirements

- Data is effectively a permanent data set
  - This is America's cultural archive
- Archive contents must stand on their own (no external databases required to know all about a file)
- Must be file format agnostic
- Must be scalable to very large size (EiB+)

# Bit Error Rate Matters!

- When you get to the PiB level:
- $10^{-17}$  bit error rates is GiB of errors in your repository!
- T10k has best current error rate:  $10^{-19}$
- All other storage: currently the best is  $10^{-17}$ 
  - 2 orders of magnitude worse error rate!
- When you are migrating every 5-10 years your entire library, BIT ERROR RATE MATTERS!!!!

# Issues

- Most commercial IT equipment has bit error rates of  $10^{-14}$ , including Ethernet backbone equipment: what good is storage BER of  $10^{-17}$  when your system's best BER is  $10^{-14}$
- How often to check data integrity?
  - Continuous above a certain size
  - Reading the data can also damage it!
- How often to migrate?
  - Individual files: every 5-10 years (we think)
  - Subject to verification!

Future Challenges?

# Future Challenges

- New production systems coming online:
  - Congressional video archiving: 2-5 PiB/year?
    - 3300 hours x 720p59.97 HD/year
  - Born Digital file submissions: 2-5 PiB/year?
- HD video encoding now possible
- Live Capture system coming online



# Future Challenges

- Standards work continues on...
  - SMPTE AXF proposed standard for media-agnostic file definition
  - SMPTE/AMWA MXF Application Specification for media files with extra metadata & associated essences enabled
  - Update the MXF standard to properly define JPEG2000 interlace vs. progressive video cadence
  - Work with AES, SMPTE, AMWA, AMPAS & others on defining a complete set of metadata standards (or at least templates!)

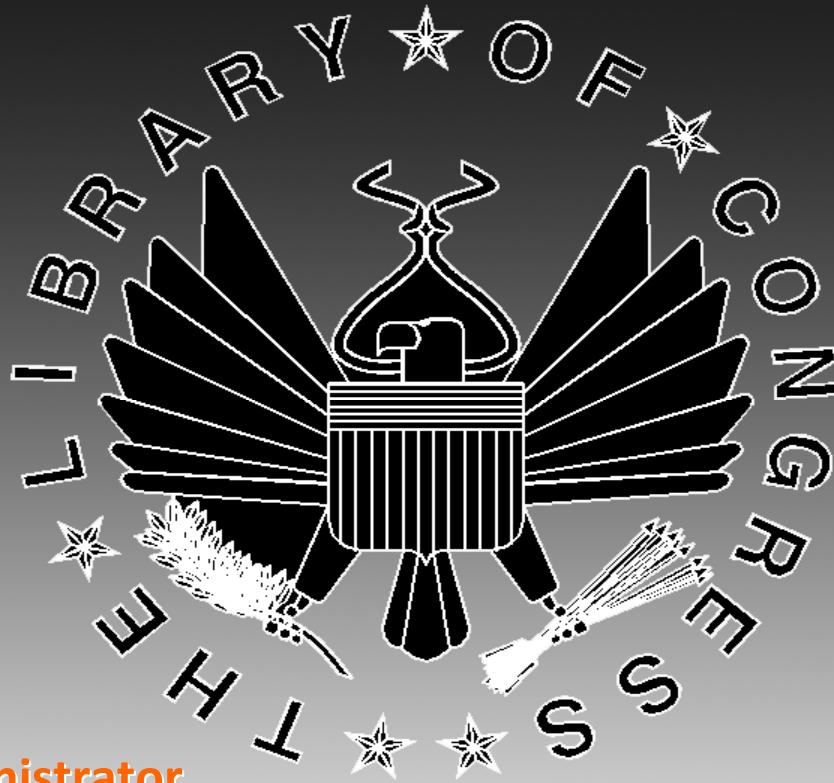
# Future Challenges

- Film scanning:
  - Real time 4k film scanners with non-bayered imagers
  - Test 8k film scanning for 35mm
  - Develop mass-migration capabilities for our 255 million feet of film

# Future Challenges

- Finding enough equipment to keep the migrations going
- Growing the Digital Repository into the exabyte realm...and beyond?
  - Not that far away!
- Developing the knowledge and training needed to make sure the 2-4 GENERATIONS of employees working on this project are adequately trained with proper documentation
- Getting the most bang for the bucks spent
- Funding (IE finding the bucks to spend)

# Thank You!



**James Snyder**  
**Senior Systems Administrator**  
**National Audio Visual Conservation Center**  
**Culpeper, Virginia**  
**[jsny@loc.gov](mailto:jsny@loc.gov) 202-707-7097**

# Sustainability of Digital Formats

## Planning for Library of Congress Collections

### Sustainability Factors

#### Table of Contents

- [Disclosure](#)
- [Adoption](#)
- [Transparency](#)
- [Self-documentation](#)
- [External dependencies](#)
- [Impact of patents](#)
- [Technical protection mechanisms](#)

#### Overview of factors

In considering the suitability of particular digital formats for the purposes of preserving digital information as an authentic resource for future generations, it is useful to articulate important factors that affect choices. The seven sustainability factors listed below apply across digital formats for all categories of information. These factors influence the likely feasibility and cost of preserving the information content in the face of future change in the technological environment in which users and archiving institutions operate. They are significant whatever strategy is adopted as the basis for future preservation actions: migration to new formats, emulation of current software on future computers, or a hybrid approach.

Additional factors will come into play relating to the ability to represent significant characteristics of the content. These factors reflect the quality and functionality that will be expected by future users. These factors will vary by genre or form of expression for content. For example, significant characteristics of sound are different from those of still pictures, whether digital or not, and not all digital formats for images are appropriate for all genres of still pictures. These factors are discussed in the sections of this Web site devoted to particular [Content Categories](#).

#### Disclosure

Disclosure refers to the degree to which complete specifications and tools for validating technical integrity exist and are accessible to those creating and sustaining digital content. Preservation of content in a given digital format over the long term is not feasible without an understanding of how the information is represented (encoded) as bits and bytes in digital files.

A spectrum of disclosure levels can be observed for digital formats. Non-proprietary, open standards are usually more fully documented and more likely to be supported by tools for validation than proprietary formats. However, what is most significant for this sustainability factor is not approval by a recognized standards body, but the existence of complete documentation, preferably subject to external expert evaluation. The existence of tools from various sources is valuable in its own right and as evidence that specifications are adequate. The existence and exploitation of underlying patents is not necessarily inconsistent with full disclosure but may inhibit the adoption of a format, as indicated below. In the future, deposit of full documentation in escrow with a trusted archive would provide some degree of disclosure to support the preservation of information in proprietary formats for which documentation is not publicly available. Availability, or deposit in escrow, of source code for associated rendering software, validation tools, and software development kits also contribute to disclosure.

[Back to top](#)

#### Adoption

Adoption refers to the degree to which the format is already used by the primary creators, disseminators, or users of information resources. This includes use as a master format, for delivery to end users, and as a means of interchange between systems. If a format is widely adopted, it is less likely to become obsolete rapidly, and tools for migration and emulation are more likely to emerge from industry without specific investment by archival institutions.

Evidence of wide adoption of a digital format includes bundling of tools with personal computers, native support in Web browsers or market-leading content creation tools, including those intended for professional use, and the existence of many competing products for creation, manipulation, or rendering of digital objects in the format. In some cases, the existence and exploitation of underlying patents may inhibit adoption, particularly if license terms

include royalties based on content usage. A format that has been reviewed by other archival institutions and accepted as a preferred or supported archival format also provides evidence of adoption.

[Back to top](#)

### **Transparency**

Transparency refers to the degree to which the digital representation is open to direct analysis with basic tools, including human readability using a text-only editor. Digital formats in which the underlying information is represented simply and directly will be easier to migrate to new formats and more susceptible to digital archaeology; development of rendering software for new technical environments or conversion software based on the "universal virtual computer" concept proposed by Raymond Lorie will be simpler.<sup>1</sup>

Transparency is enhanced if textual content (including metadata embedded in files for non-text content) is encoded in standard character encodings (e.g., UNICODE in the UTF-8 encoding) and stored in natural reading order. For preserving software programs, source code is much more transparent than compiled code. For non-textual information, standard or basic representations are more transparent than those optimized for more efficient processing, storage or bandwidth. Examples of direct forms of encoding include, for raster images, an uncompressed bit-map and for sound, pulse code modulation with linear quantization. For numeric data, standard representations exist for signed integers, decimal numbers, and binary floating point numbers of different precisions (e.g., IEEE 754-1985 and 854-1987, currently undergoing revision).

Many digital formats used for disseminating content employ encryption or compression. Encryption is incompatible with transparency; compression inhibits transparency. However, for practical reasons, some digital audio, images, and video may never be stored in an uncompressed form, even when created. Archival repositories must certainly accept content compressed using publicly disclosed and widely adopted algorithms that are either lossless or have a degree of lossy compression that is acceptable to the creator, publisher, or primary user as a master version.

The transparency factor relates to formats used for archival storage of content. Use of lossless compression or encryption for the express purpose of efficient and secure transmission of content objects to or from a repository is expected to be routine.

[Back to top](#)

### **Self-documentation**

Digital objects that are self-documenting are likely to be easier to sustain over the long term and less vulnerable to catastrophe than data objects that are stored separately from all the metadata needed to render the data as usable information or understand its context. A digital object that contains basic descriptive metadata (the analog to the title page of a book) and incorporates technical and administrative metadata relating to its creation and early stages of its life cycle will be easier to manage and monitor for integrity and usability and to transfer reliably from one archival system to its successor system. Such metadata will also allow scholars of the future to understand how what they observe relates to the object as seen and used in its original technical environment. The ability of a digital format to hold (in a transparent form) metadata beyond that needed for basic rendering of the content in today's technical environment is an advantage for purposes of preservation.

The value of richer capabilities for embedding metadata in digital formats has been recognized in the communities that create and exchange digital content. This is reflected in capabilities built in to newer formats and standards (e.g., TIFF/EP, JPEG2000, and the Extended Metadata Platform for PDF [XMP]) and also in the emergence of metadata standards and practices to support exchange of digital content in industries such as publishing, news, and entertainment. Archival institutions should take advantage of, and encourage, these developments. The Library of Congress will benefit if the digital object files it receives include metadata that identifies and describes the content, documents the creation of the digital object, and provides technical details to support rendering in future technical environments. For operational efficiency of a repository system used to manage and sustain digital content, some of the metadata elements are likely to be extracted into a separate metadata store. Some elements will also be extracted for use in the Library's catalog and other systems designed to help users find relevant resources.

Many of the metadata elements that will be required to sustain digital objects in the face of technological change are not typically recorded in library catalogs or records intended to support discovery. The OAIS Reference Model for an Open Archival Information System recognizes the need for supporting information (metadata) in several categories: representation (to allow the data to be rendered and used as information); reference (to identify and

describe the content); context (for example, to document the purpose for the content's creation); fixity (to permit checks on the integrity of the content data); and provenance (to document the chain of custody and any changes since the content was originally created). Digital formats in which such metadata can be embedded in a transparent form without affecting the content are likely to be superior for preservation purposes. Such formats will also allow metadata significant to preservation to be recorded at the most appropriate point, usually as early as possible in the content object's life cycle. For example, identifying that a digital photograph has been converted from the RGB colorspace, output by most cameras, to CMYK, the colorspace used by most printing processes, is most appropriately recorded automatically by the software application used for the transformation. By encouraging use of digital formats that are designed to hold relevant metadata, it is more likely that this information will be available to the Library of Congress when needed.

[Back to top](#)

### **External dependencies**

External dependencies refers to the degree to which a particular format depends on particular hardware, operating system, or software for rendering or use and the predicted complexity of dealing with those dependencies in future technical environments. Some forms of interactive digital content, although not tied to particular physical media, are designed for use with specific hardware, such as a microphone or a joystick. Scientific datasets built from sensor data may be useless without specialized software for analysis and visualization, software that may itself be very difficult to sustain, even with source code available.

This factor is primarily relevant for categories of digital content beyond those considered in more detail in this document, for which static media-independent formats exist. It is however worth including here, since dynamic content is likely to become commonplace as part of electronic publications. The challenge of sustaining dynamic content with such dependencies is more difficult than sustaining static content, and will therefore be much more costly.

[Back to top](#)

### **Impact of patents**

Patents related to a digital format may inhibit the ability of archival institutions to sustain content in that format. Although the costs for licenses to decode current formats are often low or nil, the existence of patents may slow the development of open source encoders and decoders and prices for commercial software for transcoding content in obsolescent formats may incorporate high license fees. When license terms include royalties based on use (e.g., a royalty fee when a file is encoded or each time it is used), costs could be high and unpredictable. It is not the existence of patents that is a potential problem, but the terms that patent-holders might choose to apply.

The core components of emerging ISO formats such as JPEG2000 and MPEG4 are associated with "pools" that offer licensing on behalf of a number of patent-holders. The license pools simplify licensing and reduce the likelihood that one patent associated with a format will be exploited more aggressively than others. However, there is a possibility that new patents will be added to a pool as the format specifications are extended, presenting the risk that the pool will continue far longer than the 20-year life of any particular patent it contains. Mitigating such risks is the fact that patents require a level of disclosure that should facilitate the development of tools once the relevant patents have expired.

The impact of patents may not be significant enough in itself to warrant treatment as an independent factor. Patents that are exploited with an eye to short-term cash flow rather than market development will be likely to inhibit *adoption*. Widespread adoption of a format may be a good indicator that there will be no adverse effect on the ability of archival institutions to sustain access to the content through migration, dynamic generation of service copies, or other techniques.

[Back to top](#)

### **Technical protection mechanisms**

To preserve digital content and provide service to users and designated communities decades hence, custodians must be able to replicate the content on new media, migrate and normalize it in the face of changing technology, and disseminate it to users at a resolution consistent with network bandwidth constraints. Content for which a trusted repository takes long-term responsibility must not be protected by technical mechanisms such as encryption, implemented in ways that prevent custodians from taking appropriate steps to preserve the digital content and make it accessible to future generations.



No digital format that is inextricably bound to a particular physical carrier is suitable as a format for long-term preservation; nor is an implementation of a digital format that constrains use to a particular device or prevents the establishment of backup procedures and disaster recovery operations expected of a trusted repository.

Some digital content formats have embedded capabilities to restrict use in order to protect the intellectual property. Use may be limited, for example, for a time period, to a particular computer or other hardware device, or require a password or active network connection. In most cases, exploitation of the technical protection mechanisms is optional. Hence this factor applies to the way a format is used in business contexts for particular bodies of content rather than to the format.

The embedding of information into a file that does not affect the use or quality of rendering of the work will not interfere with preservation, e.g., data that identifies rights-holders or the particular issuance of a work. The latter type of data indicates that this copy of this work was produced for an specific individual or other entity, and can be used to trace the movement of this copy if it is passed to another entity.

<sup>1</sup> For examples of Lorie's treatment of this subject, see his "Long Term Preservation of Digital Information" in E. Fox and C. Borgman, editors, Proceedings of the First ACM/IEEE Joint Conference on Digital Libraries (JCDL'01), pages 346-352, Roanoke, VA, June 24-28 2001, <http://doi.acm.org/10.1145/379437.379726>; and The UVC: a Method for Preserving Digital Documents: Proof of Concept (December 2002), [http://www.kb.nl/hrd/dd/dd\\_onderzoek/reports/4-uvc.pdf](http://www.kb.nl/hrd/dd/dd_onderzoek/reports/4-uvc.pdf).

[Back to top](#)

Last Updated: 02/29/2012



NATIONAL ENDOWMENT FOR THE  
**Humanities**



**LIBRARY OF  
CONGRESS**

JPEG2000 and the National Digital Newspaper Program

NATIONAL ENDOWMENT FOR THE HUMANITIES  
and LIBRARY OF CONGRESS

Deborah Thomas  
US Library of Congress

Fair tonight and tomorrow.

The

NUMBER 4166

WASH

**REFORM WINS A**

ELECTION SIGNALS  
START GREAT FIRE  
IN SAN FRANCISCO

Plant of the Daily Chronicle Almost Entirely Destroyed.

LIGHT IGNITED TOWER

Flames Spread Down into Building With Great Damage.

SAN FRANCISCO, Cal., Nov. 8.—The Chronicle tower, rising high above the clouds above the main building at the corner of the main building at Market and Kearney streets, caught fire at 11:30 last night in all probability as a result of the use of red fire in the printing plant signals during the evening, and the magnificent apparatus, with the four clocks that had become famous in all California, was entirely destroyed.

When the tower the roaring flames seemed downward and ate their way into the Chronicle main building and all their contents. At the top of the building and for a time it seemed that the entire structure would be destroyed.

Thought Lives in Danger.  
There were rumors of men out of their minds of escape with a work being to check the flames in the tower, and the fire department members laboring to bring the tower down.

SOME OF THE MEN WHO WON IN THE GREAT ELE



JOHN M. PALFREY ELECTED

**The Paducah Evening Sun.**

PADUCAH, KY., TUESDAY EVENING, JANUARY 19, 1909.

VOL. XXV NO. 16.

**DEMURRER FILED TO JURISDICTION OF CIRCUIT COURT**

Successful Candidates to Call for Democratic Primary Fight Back.

Think Contest Should Go Before Party Authority.

COMES UP AT TERM IN APRIL

**Fire at Lowes Destroys Property Valued at \$25,000—The Store of J. R. Lowe and Son is Burned**

Blaze Discovered Between 9 and 10 O'clock Last Night. Forces Uncontrollable—Both Telephones Are Lost.

Lowes, Ky., Jan. 18.—(Special.)—The store of J. R. Lowe and Son, located between 9 and 10 o'clock last night between 9 and 10 o'clock, in the fire the office of both telephones were destroyed.

The loss will amount to \$25,000 insurance, and the value of the building is estimated at \$5,000, with the amount of the fire loss unknown. The origin of the fire is not known. It started in the

**MADISONVILLE AGAIN.**

Madisonville, Ky., Jan. 18.—(Special.)—Madisonville voted dry yesterday by 128 majority. This is the second time within a year.

Fire early this morning destroyed the Williams building and all its contents, causing a loss of \$50,000. Last year the fire after the prohibition election fire occurred, raising \$100,000 loss.

**J. M. WHITE**

FORMER WELL KNOWN DRY GOODS MERCHANT

**SMALL PROPERTY OWNER PAYS FULL SHARE OF BURDEN**

Mayor Smith Desires to See Him Get Value Received on His Investment.

Advocates Increases in Real Estate Assessment.

NOBODY HUNT BY LICENSE

The house ordinance introduced last session to a plain demand that, occupation tax be reduced, and that property every line of business be taxed the same, whose license was increased, is represented by a demand for a reduction. The ordinance introduced last session to a plain demand that, occupation tax be reduced, and that property every line of business be taxed the same, whose license was increased, is represented by a demand for a reduction.

**Illinois Assembly Deadlocked With Hopkins Leading in Race for the United States Senatorship**

Chamberlain Seems to be Certain Winner in Oregon Contest—Several State Legislatures Cast Votes Today.

Springfield, Ill., Jan. 19.—The first ballot in senatorial politics of the houses of the legislature gave Hopkins the nominal vote of 75 in both houses in the senatorial race.

Chamberlain Named. (Chicago, N. H., Jan. 18.—Jacob H. Gallagher named senator today.

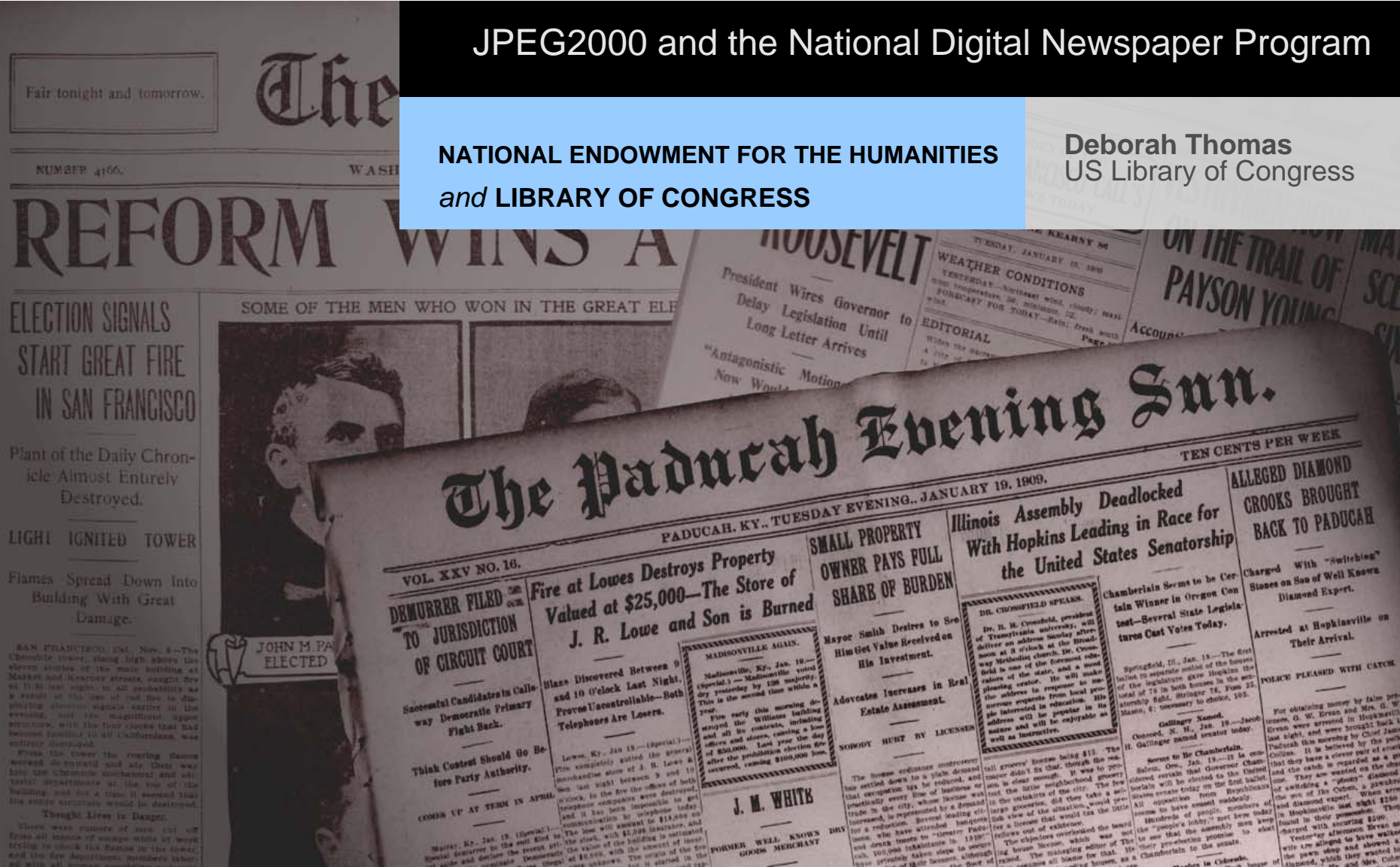
**ALLEGED DIAMOND CROOKS BROUGHT BACK TO PADUCAH**

Charged With "Switching" Stones on Son of Well Known Diamond Expert.

Arrested at Hopkinsville on Their Arrival.

POLICE PLEASED WITH CATCH

For obtaining money by false pretenses, G. W. Erwin and Mrs. G. W. Erwin were arrested in Hopkinsville last night, and were brought back to Paducah this morning by Chief James Collins. It is believed by the police that they have a clever pair of crooks. They are regarded as a good case. They were arrested on the charge of switching a "glorious" diamond on the son of the Coben, a pawnbroker and diamond expert. Erwin arrested in Hopkinsville last night 12:30 was found in their possession. They are charged with securing \$200.



# The National Digital Newspaper Program

## GOALS:

- To enhance access to historic American newspapers
- To develop best practices for the digitization of historic newspapers
- To apply emerging technologies to the products of USNP (United States Newspaper Program, 1984-2010)
  - 140,000 titles cataloged,
  - 900,000 holding records created,
  - more than 75 million pages filmed



**LIBRARY OF  
CONGRESS**

# The National Digital Newspaper Program



- NEH grants *2-year awards (up to \$400k) to state projects*, to select and digitize historic newspapers for full-text access (100,000 pages per award).



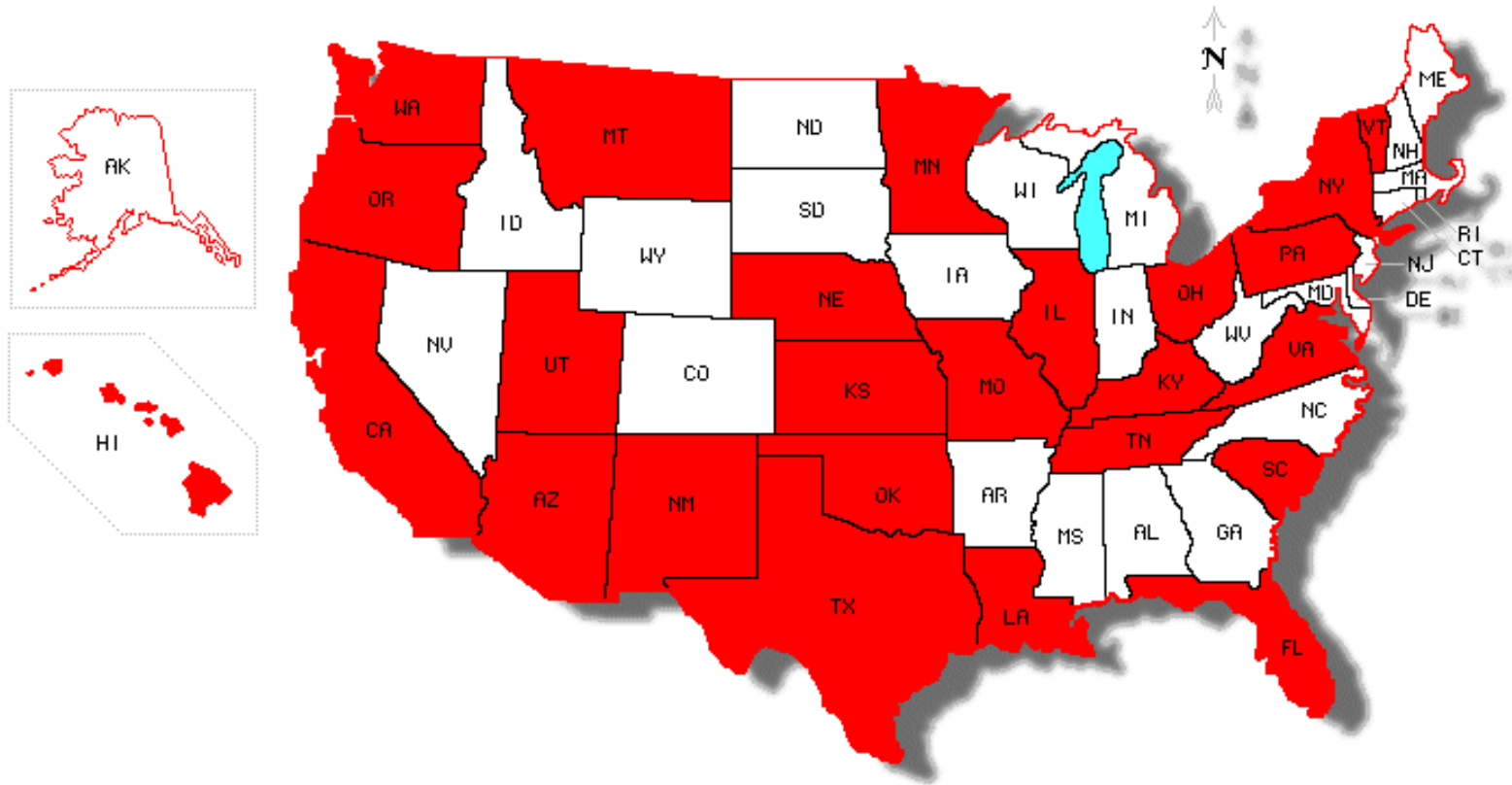
- LC creates and *hosts Chronicling America Web site* to provide freely accessible search and discovery for digitized papers and descriptive newspaper records.



- State projects *repurpose NDNP contributions for local purposes*, as desired.

## PARTNERS:

24 institutions | >4 million pages by 2012 | 1836-1922

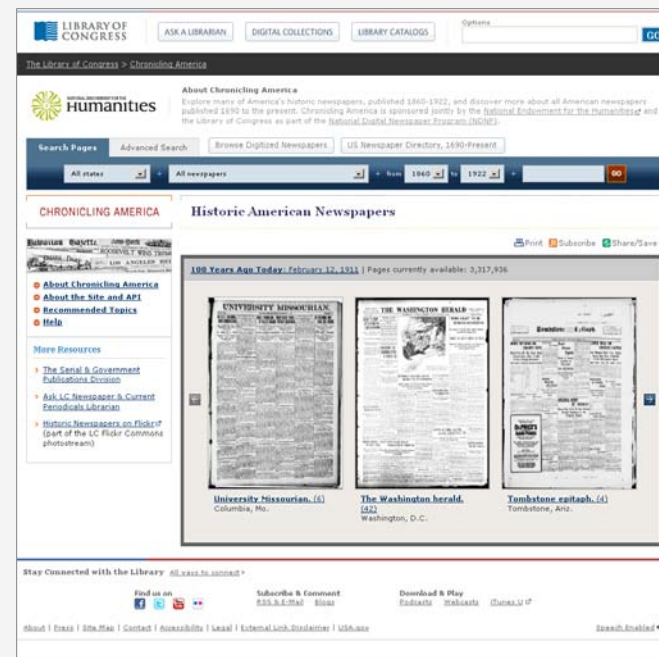


6-9-10



# Chronicling America: Historic American Newspapers

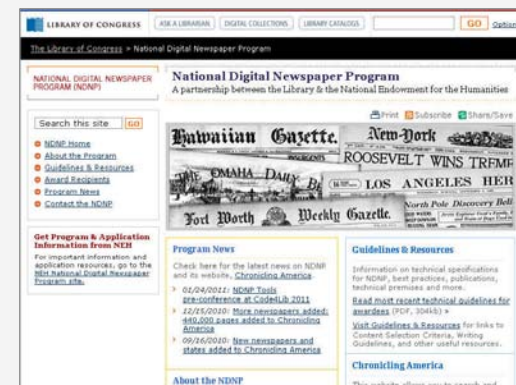
- >3.7 million pages
- 1859-1922
- >500 titles from 22 states and DC
- <http://chroniclingamerica.loc.gov/>
- Awards 2005-2010
  - *2005 awards* - CA, FL, KY, NY, UT, VA (1900-1910)
  - *2007 awards* - CA, KY, MN, NE, NY, TX, UT, VA (1880-1910)
  - *2008 awards* - AZ, HI, MO, OH, PA, WA (1880-1922)
  - *2009 awards* - IL, KS, LA, MT, OK, OR, SC (1860-1922)
  - *2010 awards* - AZ, HI, MO, NM, OH, PA, TN, VT, WA (1836-1922)
  - and onward! (next awards announced July 2011)
- Coming Soon:
  - Content from states added in 2010 (New Mexico, Tennessee, Vermont)
  - Newspapers from 1836-1859





# Beyond NDNP

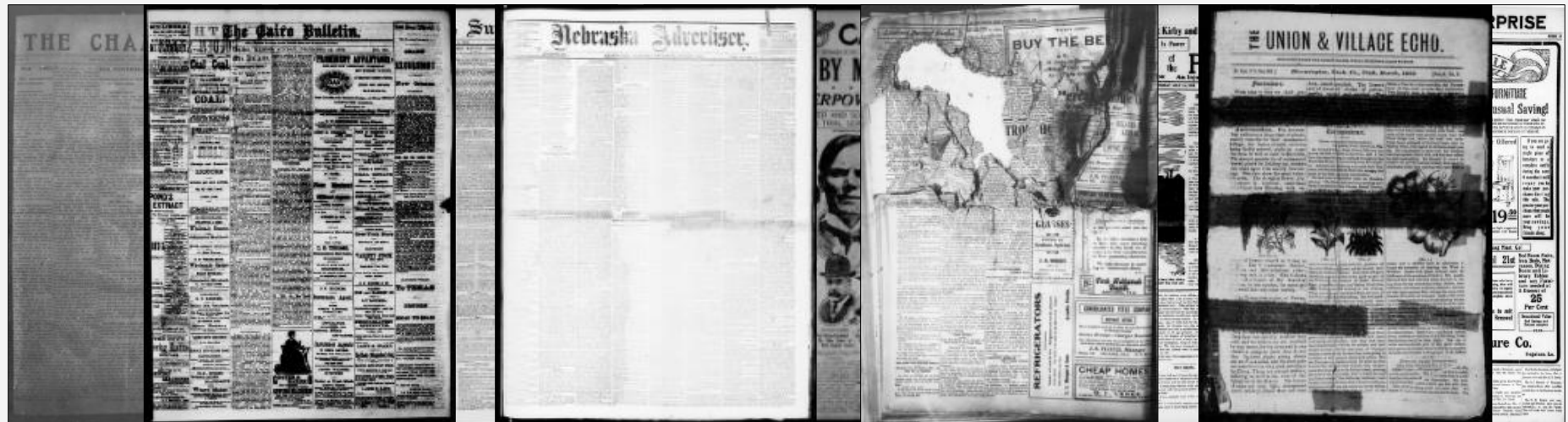
- Data specifications in use beyond NDNP
  - NDNP Guidelines - <http://www.loc.gov/ndnp/guidelines/>
  - Federal Agencies Digitization Guidelines Initiatives - <http://www.digitizationguidelines.gov/>
  - National Libraries - METS/ALTO used in
    - UK, France, Australia, New Zealand, Austria, Norway, Slovenia, Slovak Republic ...
- Open-Source Software Development –
  - LC Newspaper Viewer, available on Sourceforge.net - <http://sourceforge.net/projects/loc-ndnp/>
  - LC Newspaper Viewer in Action
    - e.g., Oregon Historical Newspapers - <http://oregonnews.uoregon.edu/>
    - Other awardees and interested parties working on software development collaboration through SourceForge





# Working with Historic Newspapers – Image Characteristics

- Scanned from microfilm 2n negatives
- Large format, little tiny type/varying type quality
- Changes in print technology over time – type, illustrations
- Varying quality: print (paper) and film (lighting, focus, process)
- Damage (acid paper, exposure, handling, etc.)
- Color space is grayscale rather than truly “high contrast” (bitonal)



# NDNP Data Specifications

- Should be as simple as is practical, producible with current technology
- Data to be created by multiple producers/vendors, and be aggregated into LC infrastructure
- Support desired research functions of the system
- Support enduring access

## DIGITAL OBJECT - Issue

- Archival Image: TIFF
- Production Image: JPEG 2000
- Printable Image: PDF
- ALTO XML for OCR
- METS with MODS/PREMIS/MIX metadata objects (issue/reel)

# JPEG2000 in NDNP

- Specification derived from "[JPEG 2000 Profile for the National Digital Newspaper Program](#)" Report, April 2006 (Prepared by: Robert Buckley and Roger Sam)
- Conforms with JPEG 2000, Part 1 (.jp2)
- Use 9-7 irreversible (lossy) filter
- Compressed to 1/8 of the TIFF or 1 bit/pixel
- Tiling, but no precincts
- Identifying RDF/Dublin Core metadata in XML box
- See NDNP JPEG2000 v2.7 profile - <http://www.loc.gov/ndnp/guidelines/archive/JPEG2kSpecs09.pdf>

# Benefits and Challenges of working with JPEG2000

## ■ BENEFITS

- Format is free to use
- Efficient compression (limited)
- Data transfer efficiency for access
- Supports tiling and efficient transformation supporting pan/zoom Web functions
- Used for production, reduces amount of storage needed on access servers

## ■ CHALLENGES

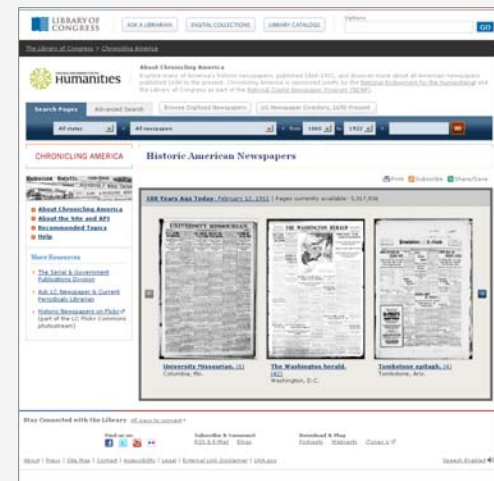
- Complex format, little forgiveness
- Complex specification, not available to the public
- Patent encumbered specification
- Commercial tool support – expensive and inconsistent
- Open-source tool support – limited in both conformance and performance

# Uses and Alternatives

- How JPEG2000 are used in NDNP:
  - Used in “production” role: used to export JPEG files to Web browser, supports “pan/zoom” behavior; available for download (compact file size)
  - Aware Imaging Library from Python (wrote code)
- Alternatives in use by NDNP Awardees:
  - Direct delivery of JPEG (browser native)
  - Pre-tiled single file in any format (PNG, JPEG, GIF)
  - Lossless compressed TIF (LZW)
  - Dynamic, cached delivery of derivatives (PNG, JPEG, GIF)

# Thank you!

- NDNP Public Web <http://www.loc.gov/ndnp/>
- NDNP Web Service  
*Chronicling America: Historic American Newspapers* <http://chroniclingamerica.loc.gov>
- Contact us at [ndnptech@loc.gov](mailto:ndnptech@loc.gov)
- Technical contact: [dbrun@loc.gov](mailto:dbrun@loc.gov)



## A Mobile Tele-Radiology Imaging System with JPEG2000 for an Emergency Care

Dong Keun Kim,<sup>1</sup> Eung Y. Kim,<sup>2</sup> Kun H. Yang,<sup>3</sup> Chung Ki Lee,<sup>4</sup> and Sun K. Yoo<sup>5</sup>

The aim of this study was to design a tele-radiology imaging system for rapid emergency care via mobile networks and to assess the diagnostic feasibility of the Joint Photographic Experts Group 2000 (JPEG2000) radiological imaging using portable devices. Rapid patient information and image exchange is helpful to make clinical decisions. We assessed the usefulness of the mobile tele-radiology system by measuring both a quantitative method, PNSR calculation, for image qualities, and its transmission time via mobile networks in different mobile networks, respectively; code division multiple access evolution-data optimized, wireless broadband, and high-speed downlink packet access; and the feasibility of the JPEG2000 computed tomography (CT) images by qualitatively assessing with the Alberta stroke program early CT score method with 12 CT image cases (seven normal and five abnormal cases). We found that the quality of the JPEG2000 radiological images was satisfied quantitatively and was judged as acceptable qualitatively at 5:1 and 10:1 compression levels for the mobile tele-radiology imaging system. The JPEG2000-format radiological images achieved a fast transmission while maintaining a diagnosis quality on a portable device via mobile networks. Unfortunately, a PDA device, having a limited screen resolution, posed difficulties in reviewing the JPEG2000 images regardless of the compression levels. An ultra mobile PC was preferable to study the medical image. The mobile tele-radiology imaging systems supporting JPEG2000 image transmission can be applied to actual emergency care services under mobile computing environments.

**KEY WORDS:** Mobile tele-radiology, JPEG2000, radiological CT image, emergency care

### INTRODUCTION

Emergency situations can unexpectedly occur anytime and anyplace. Making a rapid clinical decision is a crucial factor in emergency medical care. A mobile emergency tele-radiology system can be helpful to support rapid clinical decisions in emergency situations by specialists having difficulty

accessing a stationary system outside of hospitals.<sup>1-3</sup> Radiological images need to be interpreted by trained radiologists to achieve an accurate diagnosis. Immediate communications must be accomplished by rapid image transfer so that specialists can understand thoroughly the radiological results. Unfortunately, if the special radiologist is outside the hospitals or a resident is not able to interpret the images accurately, the emergency dispatch could be delayed. Such delays may result in emergency patient fatality because of deferred image interpretations. Therefore, we designed the mobile emergency patient information and imaging communication system that provide rapid access to the patient information and images either inside or outside of hospitals via a wireless mobile communication link such as code division multiple access evolution-data optimized (CDMA

<sup>1</sup>From the Division of Digital Media Technology, College of Software, Sangmyung University, Seoul, South Korea.

<sup>2</sup>From the Department of Radiology and Research Institute of Radiological Science, Yonsei University College of Medicine, Seoul, South Korea.

<sup>3</sup>From the Department of Radiology and Research Institute of Radiological Science Brain Korea 21 Project for Medical Science, Yonsei University College of Medicine, Seoul, South Korea.

<sup>4</sup>From the Graduate Programs of Biomedical Engineering, Yonsei University, Seoul, South Korea.

<sup>5</sup>From the Department of Medical Engineering and Brain Korea 21 Project for Medical Science, Yonsei University College of Medicine, 134 Shinchon-dong, Seodaemun-ku, Seoul, 120-752, South Korea.

Correspondence to: Sun K. Yoo, Department of Medical Engineering and Brain Korea 21 Project for Medical Science, Yonsei University College of Medicine, 134 Shinchon-dong, Seodaemun-ku, Seoul, 120-752, South Korea; tel: +82-2-22281919; fax: +82-2-3639923; e-mail: sunkyoo@yuhs.ac

Copyright © 2010 by Society for Imaging Informatics in Medicine

Online publication 8 September 2010

doi: 10.1007/s10278-010-9335-0



1x-EVDO), wireless broadband (WiBro), and high-speed downlink packet access (HSDPA) networks.

Many studies about radiological image quality evaluation for tele-radiology systems were assessed according to the Joint Photographic Experts Group (JPEG) compression ratios because preserving acceptable qualities of medical images is crucial. However, in order to decrease misinterpretation upon a diagnostically relevant loss of image quality with an efficient compression method and increase transmission performances varying wireless mobile networks to support the rapid reviewing and the demand of referring physicians to have quick access to the image data of their patients even from outside the hospital using a mobile device, evaluating influence of compression on radiological images is needed to be studied. Therefore, in order to overcome the data bandwidth limitation at wireless communication links and improve the efficiency of mobile tele-radiology systems for transferring massive Digital Imaging and Communications in Medicine (DICOM) medical images, we adopted the JPEG2000 coding method.<sup>4,5</sup>

In this study, we designed an integrated mobile tele-radiology imaging system. In emergency departments, emergency physicians can acquire radiological images and patient information and transmit those images via Web service operation, and remote clinicians can access those compressed images and patient information using portable devices through various mobile links. We assessed the feasibility of using this system for the rapid transmission of emergency patient information and images. Also, we investigated its usefulness for

making an effective medical decision in an emergency with JPEG2000 images displayed on a portable device.

## MOBILE TELE-RADIOLOGY SYSTEM

### System Configuration

The overall system configuration was described as shown in Figure 1. The designed system was composed of five components, representatively; (a) acquisition modality, (b) emergency DICOM viewer, (c) JPEG2000 compress module, (d) emergency mobile Web service, and (e) mobile browser as end-user mobile devices. The acquisition modality (a) is the CT modality (LightSpeed Plus; GE Healthcare, Milwaukee, WI, USA). The emergency DICOM viewer (b) was designed to allow review of patient information and images from the acquisition modality. The JPEG2000 compression module (c) was composed to convert DICOM images to JPEG 2000 format images and to record them on the mobile Web PACS system including patient information. The LEADTOOLS® (LEAD Technologies, Inc., NC, USA) JPEG2000 encoding library and Korean-DICOM (KDICOM) library were customized for this system. The mobile Web PACS system (d) was composed of the mobile Web service using a mobile internet toolkit with a Microsoft.NET platform in order to support the service in the mobile web browser regardless of the type of portable embedded system, mobile phone, and embedded operating system. We also included security using a

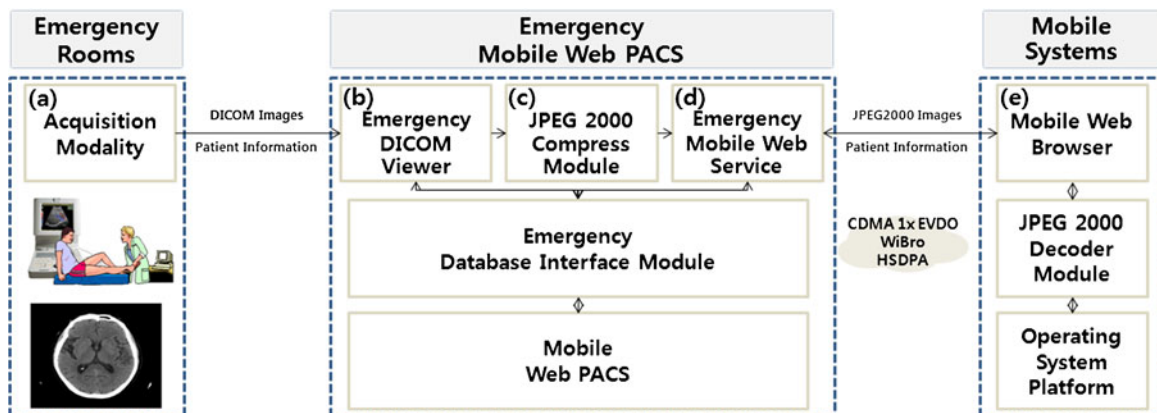


Fig. 1. The configuration of a mobile tele-radiology imaging system composed of an emergency mobile Web PACS and mobile system with five components; a acquisition modality, b emergency DICOM viewer, c JPEG2000 compress module, d emergency mobile Web service, and e mobile browser as end-user mobile devices.

log-in feature to provide an effective mobile Web service. The Web application software (e) in the end-user devices was developed using C# programming language in conjunction with the Windows Mobile 5 software development kit (SDK).

### System Operation

The designed mobile tele-radiology imaging system can support transmitting images and patient information via Web services operation and displayed on mobile devices in real time. This system was designed to transmit radiological images from a physician at an emergency room to a specialist located outside a hospital for emergency consultations. The radiological images were compressed with the JPEG2000 format in real time, and transmitted via various mobile networks such as CDMA 1x EV-DO, WIBRO, and HSDPA. The system operation and workflow were illustrated as a sequence diagram as shown in Figure 2.

The possible scenarios of emergency image transfer services are described as follows:

1. When an emergency patient arrived at an emergency department, the urgency of treatment and crucial symptom were determined by residents.
2. In spite of initial first aid and radiography, the residents could not achieve an accurate diagnosis on images through our designed Web PACS viewer because of an inability of the radiological image to convey the needed clinical information.
3. The acquired image from emergency patients were stored to the Web PACS system and simultaneously transmitted to a remote radiologist as JPEG2000 compressed format via mobile networks.
4. The remote radiologist studied the transmitted images along with detailed patient information and relevant images later by accessing the Web PACS system with a portable device.

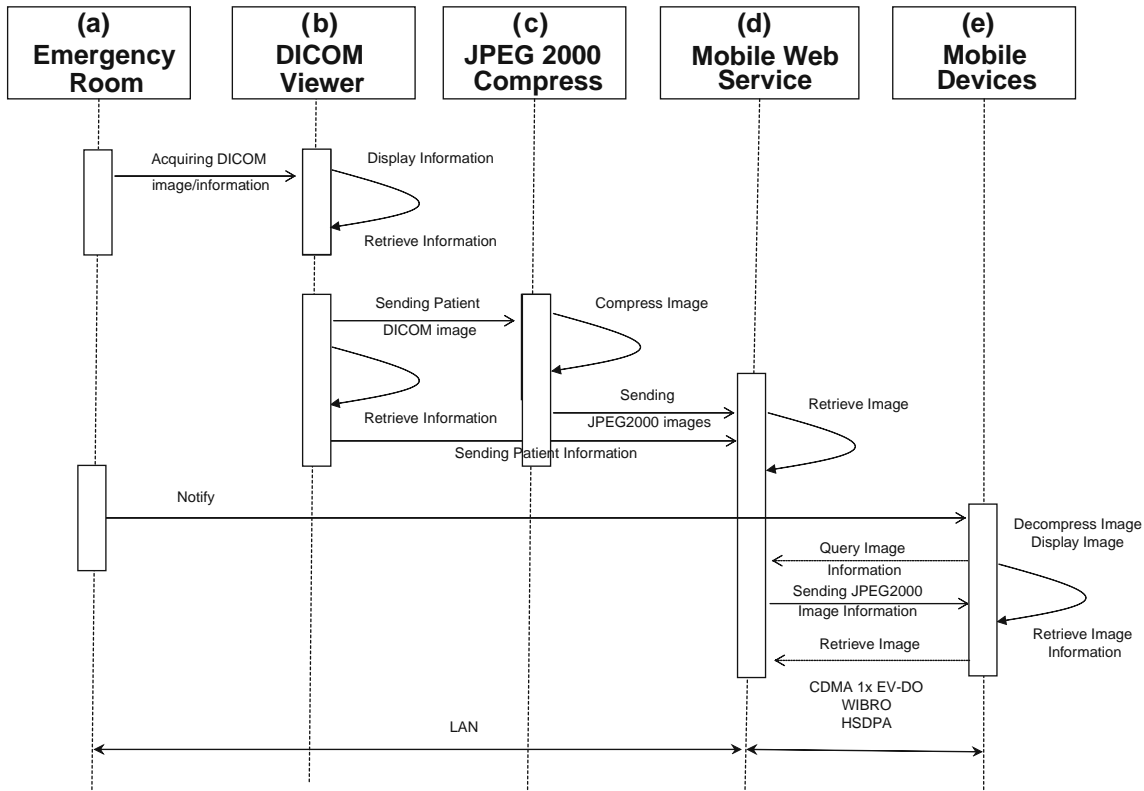


Fig. 2. The possible service scenario for mobile tele-radiology imaging system between emergency room and remote physicians via mobile networks.

## MATERIALS AND METHODS

### System Design

#### *JPEG2000*

The two-dimensional, wavelet-based, lossy image compression algorithm, JPEG2000, has been implemented in the DICOM standard since the year 2000.<sup>6</sup>

JPEG2000 provides high compression efficiency valuable in a medical imaging PACS environment. JPEG2000 can provide significantly higher compression ratios than the JPEG technique with less degradation and distortion. JPEG2000 adopts embedded block coding with optimal truncation as a source coder. In particular, a two-dimensional discrete wavelet transform is at the heart of JPEG2000. For image compression, a JPEG2000 codec (Aware JPEG; Aware, Bedford, MA) was used. The original images from the CT modality were of 16-bit depth and 512×512 pixels. The controlled compression levels of JPEG2000 codec were 5:1, 10:1, 20:1, 30:1, 40:1, 50:1, and 100:1, respectively.

#### *ASPECTS*

The Alberta stroke program early CT score (ASPECTS) was developed to offer the reliability and utility of a standard CT examination with a reproducible grading system to assess early ischemic

changes on pretreatment CT studies in patients with acute ischemic stroke of anterior circulation.<sup>7</sup> The score divides the middle cerebral artery (MCA) territory into ten regions of interest as shown in Figure 3.<sup>8</sup> The ASPECTS is a topographic scoring system applying a quantitative approach that does not ask physicians to estimate volumes from two-dimensional images. The amount of clinical discrepancy can be evaluated using scoring of a ten-point scale. A score of zero indicated diffuse ischemic involvement throughout the territory of middle cerebral artery in brain CT images, and contrary to ten evidenced normal status in brain CT images. In this study, for the assessment of image quality, we used APSECTS as a scoring system to compare between the original image and compressed image.

### System Evaluation

We investigated the usefulness of the designed mobile tele-radiology imaging system with two kinds of objective experiments.

1. Calculating the peak signal-to-noise ratio (PNSR) values in radiological images with respect to different compression ratios in JPEG 2000 format
2. Measuring the image transmission times at mobile networks, such as CDMA 1x EVDO, WIBRO, and HSDPA in terms of different compression ratios in JPEG2000 format

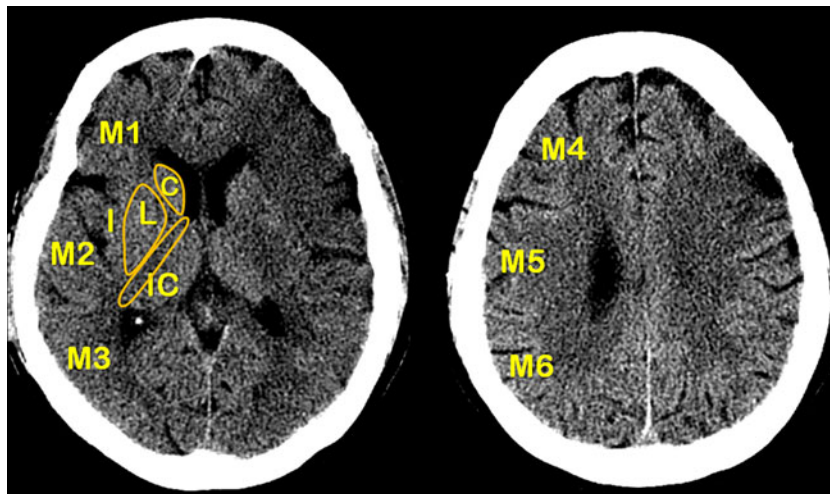


Fig. 3. The ten regions for APSECTS on a brain CT image; M1 anterior MCA cortex; M2 MCA cortex lateral to insular ribbon; M3 posterior MCA cortex; M4, M5, and M6 are anterior, lateral, and posterior MCA territories immediately superior to M1, M2, and M3, rostral to basal ganglia; I insular ribbon; L lentiform; C caudate, and IC internal capsule.

The PSNR method is the most widely used image quality metrics and was calculated from the expression;

$$PSNR = 20 \log_{10} \left\{ 255 / \sqrt{\frac{1}{MN} \sum_{y=1}^M \sum_{x=1}^N [I(x,y) - J(x,y)]^2} \right\}$$

where  $I(x, y)$  is the original image,  $J(x, y)$  is the decompressed image, and  $M$  and  $N$  are the dimensions of the image.

For evaluating the clinical diagnosis ability on portable devices subjectively, for the feasibility of medical images, considering an emergency involving an ischemia brain, we reviewed brain DICOM images of radiological lesions in 12 (seven normal and five abnormal) cases produced by the computerized tomography to compare between the original image and compressed image. The original image was displayed on a single monitor (CV812R, 18.1-in. CRT panel, TOTOKU Co., Japan). Compressed images were displayed on an ultra mobile PC (UMPC, VGN-UX58LN, 4.5-in. LCD panel, Sony Co., Japan). The monitor resolutions of both the single monitor and UMPC were  $1,280 \times 1,024$  pixels, and  $1,024 \times 600$  pixels, respectively. The image reviewing time was 10 s at each study.

RESULTS

Usefulness of Mobile Tele-Radiology System

The mobile system could query patient information and images on the mobile Web PACS instantly. The application user interface of the designed mobile system was handy to operate based on a Web browser application, as shown in Figure 4. The image transmission of mobile tele-radiology system was composed of four supported functional steps; (1) In a log-in process, only related or available doctors could review and query the emergency medical information through the user authentication process. When an incorrect identification or password was entered, guidelines for the correct process were suggested on the Web browser of mobile devices, (2) after the user authentication, the related patient’s name were listed up on a mobile device, (3) the patient information from a database in the hospital was displayed, (4) clicking the “image” button on a web browser, and the JPEG2000 radiological image was displayed.

The test results for a PSNR calculation of JPEG 2000 images in terms of compression ratios were tabulated in Table 1. Higher compression levels provided quantitatively lower image quality. Comparatively, the PSNR values of 5:1 and 10:1

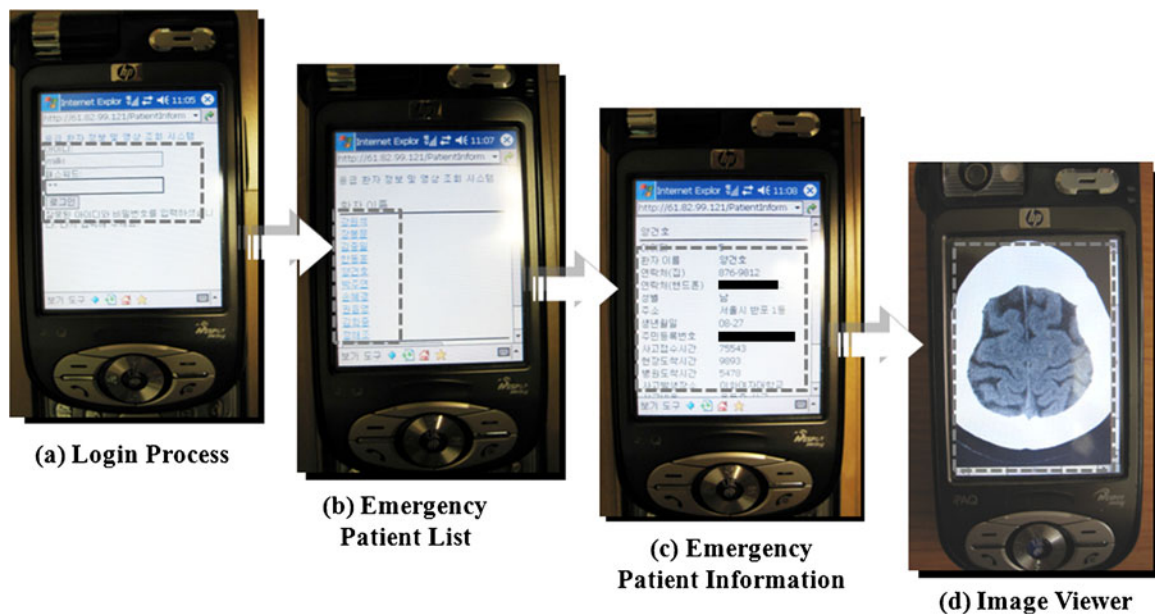


Fig. 4. The application user interface of the designed mobile system; a log-in process, b patients’ name list, c patient’s information, and d patient’s image display.

**Table 1. The result values of the PSNR test for various compression ratios**

Compression ratios	PSNR
5:1	52.64 ± 0.22
10:1	40.54 ± 0.17
20:1	31.12 ± 0.25
30:1	26.47 ± 0.24
40:1	24.18 ± 0.53
50:1	22.70 ± 0.32
100:1	19.08 ± 0.55

compression images having 52.64 dB and 40.54 dB, respectively, were higher than others. High PSNR values ( $\geq 40$  dB) for compression ratios up to 10:1 were clinically reasonable compression ratios according to the preliminary studies in this study.

Moreover, Table 2 shows the test results for transmission time of JPEG2000 images in terms of compression ratios at CDMA 1x EVDO, WIBRO, and HSDPA networks, respectively. The average transmission time was measured by 30 time repetitive transmission trials and calculated with five images randomly selected among 12 brain CT images. Higher compression ratios provided faster transmission speeds. A significant difference of transmission time (over 1 s) was realized at the 5:1 compression ratio, as compared to the original. Transmission time of JPEG2000 images at a 5:1 compression ratio was 1.23 s in CDMA 1x EVDO, 1.18 s in WIBRO, and 1.17 s in HSDPA networks, respectively. The transmission time of the HSDPA network was the fastest between both CDMA 1x EVDO and WIBRO networks, slightly. While considering the satisfaction of the PSNR test and the measured transmission time, the radiological CT images with the compression level of 5:1 was acceptable in this study.

### Feasibility of Image Quality

Allowing for the results of the measured transmission times and PSNR values, related compressed images with 5:1 and 10:1 ratios were included in the ASPECTS test. Tables 3, 4, and 5 show the test results of ASPECTS scoring experiment for the 12 original brain CT images with original images, 5:1 compression image, and 10:1 compression image, respectively. When the total score was 10 it was considered as a normal status,

**Table 2. Average transmission time using various compression ratios from hospital to portable mobile device—brain CT image (transmission trial number was 30)**

Compression ratios	File size (KB)	CDMA 1x EVDO		WIBRO		HSDPA	
		Theoretical transmission time (seconds)	Measured transmission time (seconds)	Theoretical transmission time (seconds)	Measured transmission time (seconds)	Theoretical transmission time (seconds)	Measured transmission time (seconds)
Original	524	1.74 (at 2.4 Mbps)	6.18	1 (at 4 Mbps)	5.91	0.3 (at 14 Mbps)	5.81
5:1	105	0.35 (at 2.4 Mbps)	1.23	0.2 (at 4 Mbps)	1.18	0.06 (at 14 Mbps)	1.17
10:1	53	0.17 (at 2.4 Mbps)	0.62	0.1 (at 4 Mbps)	0.60	0.03 (at 14 Mbps)	0.59
20:1	26.2	0.08 (at 2.4 Mbps)	0.30	0.05 (at 4 Mbps)	0.29	0.01 (at 14 Mbps)	0.29
30:1	17.5	0.05 (at 2.4 Mbps)	0.20	0.03 (at 4 Mbps)	0.19	<0.01 (at 14 Mbps)	0.19
40:1	13.2	0.04 (at 2.4 Mbps)	0.16	0.02 (at 4 Mbps)	0.15	<0.01 (at 14 Mbps)	0.15
50:1	10.6	0.03 (at 2.4 Mbps)	0.13	0.02 (at 4 Mbps)	0.12	<0.01 (at 14 Mbps)	0.12
100:1	5.4	0.01 (at 2.4 Mbps)	0.06	0.01 (at 4 Mbps)	0.06	<0.01 (at 14 Mbps)	0.06



**Table 3. The ASPECTS test results of the original image**

File no.	ASPECT scores according to ten different regions at brain CT images										Total	Normality
	M1	M2	M3	M4	M5	M6	C	L	IC	I		
1	1	0	0	0	0	0	0	0	1	0	2	N
2	1	1	1	1	1	1	1	0	1	1	9	Y
4	1	1	1	1	1	1	1	1	1	0	9	Y
5	1	1	0	1	1	0	1	0	1	0	6	N
6	0	0	0	0	0	0	0	0	1	0	1	N
7	1	1	1	1	1	1	1	1	1	0	9	Y
8	0	0	0	0	0	0	0	0	1	0	1	N
9	1	1	1	0	0	0	1	1	1	0	6	N
10	1	1	1	1	1	1	1	1	1	1	10	Y
11	0	1	1	1	1	1	1	1	1	0	8	N
12	1	1	1	1	1	1	1	0	1	0	8	Y

otherwise nearing to 0 as an abnormal status. The measured scores of ASPECTS in original image accorded with 5:1 and 10:1 compression image in terms of 12 brain CT images. As shown in Table 6, correspondence about the ASPECTS test was satisfied in terms of different compression ratios with brain CT images. There was no significant clinical discrepancy on according to the ASPECT scores of the brain CT images as shown in Table 6. Therefore, subjective of image quality did not differ significantly between original and both 5:1 and 10:1 compressed images in the APSECTS test.

**DISCUSSION**

When an emergency department cannot be staffed by specialized physicians, residents can only rely on direction provided to them by

specialists. In these emergency cases, where immediate clinical treatment is the important issue, mobile tele-radiology systems reduce the possibility of serious injuries. We conducted the present study to assess the diagnosis feasibility of a JPEG2000 radiological image viewed in portable devices with wireless transmission for rapid emergency care. Consequently, there was no significant difference between the JPEG 2000 compressed images at 5:1 and 10:1 compression ratios, quantitatively or qualitatively; diagnosis on portable mobile devices using those images was possible.

The designed mobile tele-radiology system was useful to access patient images and related patient information on the mobile devices. The use of mobile device systems with medical images has been reported.<sup>1,9,10</sup> Some systems are very user-friendly and fast to communicate with remote physicians because of an instant image trans-

**Table 4. The ASPECTS test results of the 5:1 compression image**

File no.	ASPECT scores according to ten different regions at brain CT images										Total	Normality
	M1	M2	M3	M4	M5	M6	C	L	IC	I		
1	1	0	0	0	0	0	0	0	1	0	2	N
2	1	1	1	1	1	1	1	0	1	0	8	Y
3	1	1	1	1	1	1	1	0	1	1	9	Y
4	1	1	1	1	1	1	1	1	1	0	9	Y
5	1	1	0	1	1	0	1	0	1	0	6	N
6	0	0	0	0	0	0	0	0	1	0	1	N
7	1	1	1	1	1	1	1	1	1	0	9	Y
8	0	0	0	0	0	0	0	0	1	0	1	N
9	1	1	1	0	0	0	1	1	1	0	6	N
10	1	1	1	1	1	1	1	1	1	1	10	Y
11	0	0	1	1	1	1	1	1	1	0	7	N
12	1	1	1	1	1	1	1	0	1	0	8	Y

**Table 5. The ASPECTS test results of the 10:1 compression image**

File no.	ASPECT scores according to ten different regions at brain CT images										Total	Normality
	M1	M2	M3	M4	M5	M6	C	L	IC	I		
1	1	0	0	0	0	0	0	0	1	0	2	N
2	1	1	1	1	1	1	1	0	1	0	8	Y
3	1	1	1	1	1	1	1	0	1	1	9	Y
4	1	1	1	1	1	1	1	1	1	0	9	Y
5	1	1	0	1	1	0	1	0	1	0	6	N
6	0	0	0	0	0	0	0	0	1	0	1	N
7	1	1	1	1	1	1	1	1	1	0	9	Y
8	0	0	0	0	0	0	0	0	1	0	1	N
9	1	1	1	0	0	0	1	1	1	0	6	N
10	1	1	1	1	1	1	1	1	1	1	10	Y
11	0	0	1	1	1	1	1	1	1	0	7	N
12	1	1	1	1	1	1	1	0	1	0	8	Y

mission as multimedia messaging services or a direct image capturing transmission; however, sending patient information corresponded to the hospital information system and preserving image quality was necessary to be considered for the point-of-expertise.

In this study, inquiry of both patient images and relevant information helped to make clinical decisions through the mobile Web PACS. Manipulation for identification access, image inquiry, and transmission was useful and not annoying. Functionalities for image manipulation, such as zoom-in, zoom-out, rotation, and magnification, were useful on a portable device.<sup>1-3</sup>

Regarding the transmission of medical images, there are essentially no theoretical bandwidth requirements, but the transmission time is crucial to the applicability of the mobile tele-radiology system in the emergency situations. It is associated with the bandwidth performance of mobile network. In this study, although the CDMA 1x EVDO data transmission speed was slightly slower than of other networks; in WIBRO and HSDPA, the transmission speed was acceptable for transmitting JPEG2000 compressed images in emergency situations. If the mobile tele-radiology system has connected with a CDMA 1x EVDO network, then the brain CT images with lossless

**Table 6. ASPECTS results in terms of different compression ratios with cases (C = correspondence observed in different brain CT images)**

File no.	Compression ratios					
	Original		5:1		10:1	
	ASPECTS total scores	Correspondence in ASPECTS	ASPECTS total scores	Agreement (%)	ASPECTS total scores	Agreement (%)
1	2	C	2	100	2	100
2	8	C	8	100	8	100
3	9	C	8	100	8	100
4	9	C	9	100	9	100
5	6	C	5	100	5	100
6	1	C	1	100	1	100
7	9	C	9	100	9	100
8	1	C	1	100	1	100
9	6	C	6	100	6	100
10	10	C	10	100	10	100
11	8	C	7	100	7	100
12	8	C	8	100	8	100



JPEG2000 compression (5:1) can be transmitted in approximately 1.5 s. Transmitting images using mobile devices requires not only data transmission time, but also associated with some static network setup time or log-in process time within 10 s. An actual image transmission service time with mobile tele-radiology systems can be considered more than its transmission time. If we consider the average number of images per examination to be assumed between 50 and 60 (one volume) for a CT brain case,<sup>11</sup> total transmission time using JPEG2000 compressed images at a compression level of 5:1 will be less than 1.5 min via CDMA 1x EVDO, WIBRO, and HSDPA networks. Besides timely rapid transmission of radiological images, preserving the fine quality of the images is also important for emergency care. It was reported that the JPEG2000 image referred to in the Web browser of the mobile devices was compressed into 15:1 in order to be diagnosed effectively by the radiologist using a personal digital assistant (PDA) device.<sup>12</sup> This study suggested a compression ratio of 10:1, as long as the original images were subsequently reviewed and a decrease in sensitivity at a ratio of 10:1 and above, but focused on the cerebral artery in the brain CT images. That study showed that compression ratios of up to 10:1 still provide diagnostically satisfactory image quality in the cerebral artery in brain CT images. Since the JPEG2000 format had efficient compression ratios, both fast transmission and acceptable quality was achievable.

As aforementioned, the use of an image viewer application in a mobile tele-radiology imaging system for multiple imaging processing would be quite useful. The user interface of a PDA application was handy to operate, and image studying and transmission time were not considerable, but reviewing image details, such as region of interest, on a PDA was difficult. Furthermore, it was reported that inspecting CT images can be medically sensible to review in detail using pocket-sized tele-radiology PDA terminals (2.8-in. LCD screen size and 320×240 pixels) for consultation purposes.<sup>10</sup> Normally, a CT image has a 512×512-pixel resolution; however, a PDA device has a limited screen resolution (240×320 pixels). The resolution of the CT images (512×512 pixels) did not correspond with the maximum display resolution of the PDA device (240×320 pixels). The disparity pixels (272×192 pixels) to the each side

of the CT images from 512 to 512 were discarded to match the maximum resolution when displaying the CT image images on the PDA screen. Physicians may require a more operational step to manipulate images such as zooming or rotating to supplement the limited resolution of PDA devices. In addition, performing various manipulations, such as scrolling, zooming, and tilting, with multiple CT images on a PDA device was aggravating. Clinically, minor discrepancies could occur because of low-resolution and ambient conditions such as reflected lights. On the other hand, the resolution of 1,024×600 pixels that was used with UMPCs may display original CT image resolutions advantageously so that the UMPC was appropriate to compensate for the insufficient screen resolutions of the PDA devices. Therefore, radiological image review was preferable on a UMPC device containing higher screen resolutions than PDAs. Finally, the major concern over a wireless network is its security, especially when personal information is involved. Although Web-based access control method was applied in this study, role-based access control methods for concrete user access controls and mobile virtual private network techniques for secure transmission can be involved in the designed system for proposing further practical systems.

## CONCLUSIONS

In conclusion, wireless transmission of JPEG2000 radiological images of emergency patients via mobile networks to remote specialists can help achieve proper first aid of emergency patients. We developed the mobile tele-radiology imaging system with the JPEG2000 for emergency care. This system is provided to remote physicians that require immediate access to patient's medical image and information from random locations. The results of the quantitative and qualitative evaluation about the designed mobile tele-radiology system showed that the application was useful for remote physicians due to fast image transmissions, and brain CT images with JPEG 2000 compression level 10:1 did not differ significantly associated with the original image. The performance of the system has been technically demonstrated as application of mobile tele-radiology system will help both physicians and

remote with sufficient image quality and rapid transmission rates for emergency cases.

#### ACKNOWLEDGEMENT

This research was financially supported by the Ministry of Knowledge Economy and Korea Industrial Technology Foundation through the Human Resource Training Project for Strategic Technology, and by the Basic Science Research Program through the National Research Foundation of Korea funded by the Ministry of Education, Science and Technology (2009-0074717)

#### REFERENCES

1. Kim DK, Yoo SK, Kim SH: Instant wireless transmission of radiological images using a personal digital assistant phone for emergency teleconsultation. *J Telemed Telecare* 11(2):58–61, 2005
2. Kim DK, Yoo SK, Park JJ, Kim SH: PDA-phone-based instant transmission of radiological images over a CDMA network by combining the PACS screen with a Bluetooth-interfaced local wireless link. *J Digit Imaging* 20:131–139, 2007
3. Jung SM, Yoo SK, Kim BS, Yun HY, Kim SR: Design of mobile emergency telemedicine system based on CDMA2000 1X-EVDO. *J Korean Soc Med Inform* 9:401–406, 2007
4. Sung MM, Kim HJ, Kim EK, Kwak JY, Yoo JK, Yoo HS: Clinical evaluation of JPEG2000 compression algorithm for digital mammography. *IEEE Trans Nucl Sci* 49:827–832, 2002
5. Sung MM, Kim HJ, Yoo SK, Choi BW, Nam JE, Kim HS, Lee JH, Yoo HS: Clinical evaluation of compression ratios using JPEG2000 on computed radiography chest images. *J Digit Image* 15:78–83, 2002
6. Ringl H, Scherthner RE, Bankier AA, Weber M, Herold PM, CJ PCS: JPEG2000 compression of thin-section CT images of the lung: effect of compression ratio on image quality. *Radiology* 240:869–877, 2006
7. Pexman JHW, Barber PA, Hill MD, Servick RJ, Demchuk AM, Hudon ME, Hu WY, Buchan AM: Use of the Alberta Stroke Program Early CT Score (ASPECTS) for assessing CT scans in patient with acute stroke. *Am J Neuroradiol* 22:1534–1542, 2001
8. Barber PA, Demchuk AM, Zhang J, Buchan AM: Validity and reliability of a quantitative computed tomography score in predicting outcome of hyperacute stroke before thrombolytic therapy. *Lancet* 355:1670–1674, 2000
9. Ng WH, Wang E, Ng G, Ng I: Multimedia Messaging Service teleradiology in the provision of emergency neurosurgery services. *Surg Neurol* 67:338–341, 2007
10. Reponen J, Niinimäki J, Kumpulainen T, Ilkko E, Karttunen A, Jartti P: Mobile teleradiology with smartphone terminals as a part of a multimedia electronic patient record. *Comput Assist Radiol Surg* 1281:916–921, 2005
11. Bankman In: Handbook of medical imaging processing and analysis. Academic Press, 2000
12. Yang KH, Cho HM, Jung HJ, Jang BM, Han DH, Lee CL, Kim HJ: Design of emergency medical imaging and informatics system for mobile computing environment. *World Congr Med Phys Biomed Eng* 14:4072–4076, 2006



## U.S. National Archives and Records Administration (NARA)

### Technical Guidelines for Digitizing Archival Materials for Electronic Access: Creation of Production Master Files – Raster Images

For the Following Record Types- Textual, Graphic Illustrations/Artwork/Originals, Maps, Plans, Oversized, Photographs, Aerial Photographs, and Objects/Artifacts

June 2004

Written by Steven Puglia, Jeffrey Reed, and Erin Rhodes

Digital Imaging Lab, Special Media Preservation Laboratory, Preservation Programs  
U.S. National Archives and Records Administration  
8601 Adelphi Road, Room B572, College Park, MD, 20740, USA  
Lab Phone: 301-837-3706  
Email: [preserve@nara.gov](mailto:preserve@nara.gov)

Acknowledgements: Thank you to Dr. Don Williams for target analyses, technical guidance based on his extensive experience, and assistance on the assessment of digital capture devices. Thank you to the following for reading drafts of these guidelines and providing comments: Stephen Chapman, Bill Comstock, Maggie Hale, and David Remington of Harvard University; Phil Michel and Kit Peterson of the Library of Congress; and Doris Hamburg, Kitty Nicholson, and Mary Lynn Ritzenthaler of the U.S. National Archives and Records Administration.

#### SCOPE:

The *NARA Technical Guidelines for Digitizing Archival Materials for Electronic Access* define approaches for creating digital surrogates for facilitating access and reproduction; they are not considered appropriate for preservation reformatting to create surrogates that will replace original records. The *Technical Guidelines* presented here are based on the procedures used by the Digital Imaging Lab of NARA's Special Media Preservation Laboratory for digitizing archival records and the creation of production master image files, and are a revision of the 1998 "*NARA Guidelines for Digitizing Archival Materials for Electronic Access*", which describes the imaging approach used for NARA's pilot Electronic Access Project.

The *Technical Guidelines* are intended to be informative, and not intended to be prescriptive. We hope to provide a technical foundation for digitization activities, but further research will be necessary to make informed decisions regarding all aspects of digitizing projects. These guidelines provide a range of options for various technical aspects of digitization, primarily relating to image capture, but do not recommend a single approach.

The intended audience for these guidelines includes those who will be planning, managing, and approving digitization projects, such as archivists, librarians, curators, managers, and others. Another primary audience includes those actually doing scanning and digital capture, such as technicians and photographers.

#### The following topics are addressed:

- Digital Image Capture – production master files, image parameters, digitization environment, color management, etc.
- Minimum Metadata – types, assessment, local implementation, etc. – we have included a discussion of metadata to ensure a minimum complement is collected/created so production master files are useable
- File Formats, Naming, and Storage – recommended formats, naming, directory structures, etc.
- Quality Control – image inspection, metadata QC, acceptance/rejection, etc.

#### The following aspects of digitization projects are not discussed in these guidelines:

- Project Scope – define goals and requirements, evaluate user needs, identification and evaluation of options, cost-benefit analysis, etc.
- Selection – criteria, process, approval, etc.

- Preparation – archival/curatorial assessment and prep, records description, preservation/conservation assessment and prep, etc.
- Descriptive systems – data standards, metadata schema, encoding schema, controlled vocabularies, etc.
- Project management – plan of work, budget, staffing, training, records handling guidelines, work done in-house vs. contractors, work space, oversight and coordination of all aspects, etc.
- Access to digital resources – web delivery system, migrating images and metadata to web, etc.
- Legal issues – access restrictions, copyright, rights management, etc.
- IT infrastructure – determine system performance requirements, hardware, software, database design, networking, data/disaster recovery, etc.
- Project Assessment – project evaluation, monitoring and evaluation of use of digital assets created, etc.
- Digital preservation – long-term management and maintenance of images and metadata, etc.

**In reviewing this document, please keep in mind the following:**

- The *Technical Guidelines* have been developed for internal NARA use, and for use by NARA with digitizing projects involving NARA holdings and other partner organizations. The *Technical Guidelines* support internal policy directive NARA 816 – *Digitization Activities for Enhanced Access*, at [http://www.nara-at-work.gov/nara\\_policies\\_and\\_guidance/directives/0800\\_series/nara816.html](http://www.nara-at-work.gov/nara_policies_and_guidance/directives/0800_series/nara816.html) (NARA internal link only). For digitization projects involving NARA holdings, all requirements in NARA 816 must be met or followed.
- The *Technical Guidelines* do not constitute, in any way, guidance to Federal agencies on records creation and management, or on the transfer of permanent records to the National Archives of the United States. For information on these topics, please see the Records Management section of the NARA website, at [http://www.archives.gov/records\\_management/index.html](http://www.archives.gov/records_management/index.html) and [http://www.archives.gov/records\\_management/initiatives/erm\\_overview.html](http://www.archives.gov/records_management/initiatives/erm_overview.html).
- As stated above, Federal agencies dealing with the transfer of scanned images of textual documents, of scanned images of photographs, and of digital photography image files as permanent records to NARA shall follow specific transfer guidance ([http://www.archives.gov/records\\_management/initiatives/scanned\\_textual.html](http://www.archives.gov/records_management/initiatives/scanned_textual.html) and [http://www.archives.gov/records\\_management/initiatives/digital\\_photo\\_records.html](http://www.archives.gov/records_management/initiatives/digital_photo_records.html)) and the regulations in 36 CFR 1228.270.
- The *Technical Guidelines* cover only the process of digitizing archival materials for on-line access and hardcopy reproduction. Other issues must be considered when conducting digital imaging projects, including the long-term management and preservation of digital images and associated metadata, which are not addressed here. For information on these topics, please see information about NARA's Electronic Records Archive project, at [http://www.archives.gov/electronic\\_records\\_archives/index.html](http://www.archives.gov/electronic_records_archives/index.html).
- The topics in these *Technical Guidelines* are inherently technical in nature. For those working on digital image capture and quality control for images, a basic foundation in photography and imaging is essential. Generally, without a good technical foundation and experience for production staff, there can be no claim about achieving the appropriate level of quality as defined in these guidelines.
- These guidelines reflect current NARA internal practices and we anticipate they will change over time. We plan on updating the *Technical Guidelines* on a regular basis. We welcome your comments and suggestions.

## TABLE OF CONTENTS:

Scope - 1

Introduction - 5

Metadata - 5

- o Common Metadata Types - 6
  - o Descriptive - 7
  - o Administrative - 8
  - o Rights - 8
  - o Technical - 9
  - o Structural - 10
  - o Behavior - 11
  - o Preservation - 11
  - o Image Quality Assessment - 12
  - o Records Management/Recordkeeping - 15
  - o Tracking - 15
  - o Meta-Metadata - 16
- o Assessment of Metadata Needs for Imaging Projects - 16
- o Local Implementation - 18
  - o Relationships - 20
  - o Batch Level Metadata - 20
  - o Permanent and Temporary Metadata - 21

Technical Overview - 21

- o Raster Image Characteristics:
  - o Spatial Resolution - 21
  - o Signal Resolution - 21
  - o Color Mode - 22
- o Digitization Environment - 23
  - o Viewing Conditions - 23
  - o Monitor Settings, Light Boxes, and Viewing Booths - 23
  - o The Room - 23
  - o Practical Experience - 24
  - o Monitor Calibration - 24
- o Quantifying Scanner/Digital Camera Performance - 24
  - o Test Frequency and Equipment Variability - 25
  - o Tests:
    - Opto-Electronic Conversion Function (OECF) - 26
    - Dynamic Range - 26
    - Spatial Frequency Response (SFR) - 26
    - Noise - 27
    - Channel Registration - 27
    - Uniformity - 27
    - Dimensional Accuracy - 28
    - Other Artifacts or Imaging Problems - 28
- o Reference Targets - 29
  - o Scale and Dimensional References - 29
  - o Targets for Tone and Color Reproduction - 30
    - Reflection Scanning - 30
    - Transmission Scanning – Positives - 30
    - Transmission Scanning – Negatives - 31

Imaging Workflow - 31

- o Adjusting Image Files - 31
  - o Overview - 32
- o Scanning Aimpoints - 32
  - o Aimpoints for Photographic Gray Scales - 35
  - o Alternative Aimpoints for Kodak Color Control Patches (color bars) - 36
  - o Aimpoint Variability - 36

- Minimum and Maximum Levels - 36
- Color Management Background - 36
  - ICC Color Management System - 37
  - Profiles - 37
  - Rendering Intents - 38
  - Color Management Modules - 38
- Image Processing - 38
  - Color Correction and Tonal Adjustments - 39
  - Sharpening - 39
- Sample Image Processing Workflow - 39
  - Scanning - 40
  - Post-Scan Adjustment/Correction - 40

#### Digitization Specifications for Record Types - 41

- Cleanliness of Work Area, Digitization Equipment, and Originals - 42
- Cropping - 42
- Backing Reflection Originals - 42
- Scanning Encapsulated or Sleeved Originals - 42
- Embossed Seals - 43
- Compensating for Minor Deficiencies - 43
- Scanning Text - 43
- Scanning Oversized - 44
- Scanning Photographs - 44
- Scanning Intermediates - 44
- Scanning Microfilm - 45
- Illustrations of Record Types - 46
- Requirements Tables:
  - Textual Documents, Graphic Illustrations/ Artwork/ Originals, Maps, Plans, and Oversized - 51
  - Photographs - Film/ Camera Originals - Black-and-White and Color - Transmission Scanning - 52
  - Photographs - Prints - Black-and-White, Monochrome, and Color - Reflection Scanning - 54
  - Aerial - Transmission Scanning - 56
  - Aerial - Reflection Scanning - 57
  - Objects and Artifacts - 58

#### Storage - 60

- File Formats - 60
- File Naming - 60
  - Directory Structure - 60
  - Versioning - 61
  - Naming Derivative Files - 61
- Storage Recommendations - 61
  - Digital Repositories and Long-Term Management of Files and Metadata - 61

#### Quality Control - 62

- Completeness - 62
- Inspection of Digital Image Files - 62
  - File Related - 62
  - Original/Document Related - 62
  - Metadata Related - 63
  - Image Quality Related - 63
- Quality Control of Metadata - 64
- Documentation - 65
- Testing Results and Acceptance/Rejection - 65

#### Appendices:

- A – Digitizing for Preservation vs. Production Masters - 66
- B – Derivative Files - 70
- C – Mapping of LCDRG Elements to Unqualified Dublin Core - 74
- D – File Format Comparison - 76
- E – Records Handling for Digitization - 79
- F – Resources - 82

## I. INTRODUCTION

These *Guidelines* define approaches for creating digital surrogates for facilitating access and reproduction. They are not considered appropriate for preservation reformatting to create surrogates that will replace original records. For further discussion of the differences between these two approaches, see Appendix A, Digitization for Preservation vs. Production Masters.

These guidelines provide technical benchmarks for the creation of “production master” raster image (pixel-based) files. Production masters are files used for the creation of additional derivative files for distribution and/or display via a monitor and for reproduction purposes via hardcopy output at a range of sizes using a variety of printing devices (see Appendix B, Derivative Files, for more information). Our aim is to use the production master files in an automated fashion to facilitate affordable reprocessing. Many of the technical approaches discussed in these guidelines are intended for this purpose.

Production master image files have the following attributes-

- The primary objective is to produce digital images that look like the original records (textual, photograph, map, plan, etc.) and are a “reasonable reproduction” without enhancement. The *Technical Guidelines* take into account the challenges involved in achieving this and will describe best practices or methods for doing so.
- Production master files document the image at the time of scanning, not what it may once have looked like if restored to its original condition. Additional versions of the images can be produced for other purposes with different reproduction renderings. For example, sometimes the reproduction rendering intent for exhibition (both physical and on-line exhibits) and for publication allows basic enhancement. Any techniques that can be done in a traditional darkroom (contrast and brightness adjustments, dodging, burning, spotting, etc.) may be allowed on the digital images.
- Digitization should be done in a “use-neutral” manner, not for a specific output. Image quality parameters have been selected to satisfy most types of output.

If digitization is done to meet the recommended image parameters and all other requirements as described in these *Technical Guidelines*, we believe the production master image files produced should be usable for a wide variety of applications and meet over 95% of reproduction requests. If digitization is done to meet the alternative minimum image parameters and all other requirements, the production master image files should be usable for many access applications, particularly for web usage and reproduction requests for 8”x10” or 8.5”x11” photographic quality prints.

If your intended usage for production master image files is different and you do not need all the potential capabilities of images produced to meet the recommended image parameters, then you should select appropriate image parameters for your project. In other words, your approach to digitization may differ and should be tailored to the specific requirements of the project.

Generally, given the high costs and effort for digitization projects, we do not recommend digitizing to anything less than our alternative minimum image parameters. This assumes availability of suitable high-quality digitization equipment that meets the assessment criteria described below (see Quantifying Scanner/Digital Camera Performance) and produces image files that meet the minimum quality described in the *Technical Guidelines*. If digitization equipment fails any of the assessment criteria or is unable to produce image files of minimum quality, then it may be desirable to invest in better equipment or to contract with a vendor for digitization services.

## II. METADATA

**NOTE:** All digitization projects undertaken at NARA and covered by NARA 816 *Digitizing Activities for Enhanced Access*, including those involving partnerships with outside organizations, must ensure that descriptive information is prepared in accordance with NARA 1301 Life Cycle Data Standards and Lifecycle Authority Control, at [http://www.nara-at-work.gov/nara\\_policies\\_and\\_guidance/directives/1300\\_series/nara1301.html](http://www.nara-at-work.gov/nara_policies_and_guidance/directives/1300_series/nara1301.html) (NARA internal link only), and its associated Lifecycle Data Requirements Guide, and added to NARA's Archival Research Catalog (ARC) at a time mutually agreed-upon with NARA.



Although there are many technical parameters discussed in these *Guidelines* that define a high-quality production master image file, we do not consider an image to be of high quality unless metadata is associated with the file. Metadata makes possible several key functions—the identification, management, access, use, and preservation of a digital resource—and is therefore directly associated with most of the steps in a digital imaging project workflow: file naming, capture, processing, quality control, production tracking, search and retrieval design, storage, and long-term management. Although it can be costly and time-consuming to produce, metadata adds value to production master image files: images without sufficient metadata are at greater risk of being lost.

No single metadata element set or standard will be suitable for all projects or all collections. Likewise, different original source formats (text, image, audio, video, etc.) and different digital file formats may require varying metadata sets and depths of description. Element sets should be adapted to fit requirements for particular materials, business processes and system capabilities.

Because no single element set will be optimal for all projects, implementations of metadata in digital projects are beginning to reflect the use of “application profiles,” defined as metadata sets that consist of data elements drawn from different metadata schemes, which are combined, customized and optimized for a particular local application or project. This “mixing and matching” of elements from different schemas allows for more useful metadata to be implemented at the local level while adherence to standard data values and structures is still maintained. Locally-created elements may be added as extensions to the profile, data elements from existing schemas might be modified for specific interpretations or purposes, or existing elements may be mapped to terminology used locally.

Because of the likelihood that heterogeneous metadata element sets, data values, encoding schemes, and content information (different source and file formats) will need to be managed within a digital project, it is good practice to put all of these pieces into a broader context at the outset of any project in the form of a data or information model. A model can help to define the types of objects involved and how and at what level they will be described (i.e., are descriptions hierarchical in nature, will digital objects be described at the file or item level as well as at a higher aggregate level, how are objects and files related, what kinds of metadata will be needed for the system, for retrieval and use, for management, etc.), as well as document the rationale behind the different types of metadata sets and encodings used. A data model informs the choice of metadata element sets, which determine the content values, which are then encoded in a specific way (in relational database tables or an XML document, for example).

Although there is benefit to recording metadata on the item level to facilitate more precise retrieval of images within and across collections, we realize that this level of description is not always practical. Different projects and collections may warrant more in-depth metadata capture than others; a deep level of description at the item level, however, is not usually accommodated by traditional archival descriptive practices. The functional purpose of metadata often determines the amount of metadata that is needed. Identification and retrieval of digital images may be accomplished on a very small amount of metadata; however, management of and preservation services performed on digital images will require more finely detailed metadata—particularly at the technical level, in order to render the file, and at the structural level, in order to describe the relationships among different files and versions of files.

Metadata creation requires careful analysis of the resource at hand. Although there are current initiatives aimed at automatically capturing a given set of values, we believe that metadata input is still largely a manual process and will require human intervention at many points in the object’s lifecycle to assess the quality and relevance of metadata associated with it.

This section of the *Guidelines* serves as a general discussion of metadata rather than a recommendation of specific metadata element sets; although several elements for production master image files are suggested as minimum-level information useful for basic file management. We are currently investigating how we will implement and formalize technical and structural metadata schemes into our workflow and anticipate that this section will be updated on a regular basis.

### **Common Metadata Types:**

Several categories of metadata are associated with the creation and management of production master image files. The following metadata types are the ones most commonly implemented in imaging projects. Although these categories are defined separately below, there is not always an obvious distinction between them, since each type contains elements that are both descriptive and administrative in nature. These types are commonly broken down by what functions the metadata supports. In general, the types of metadata listed below, except for descriptive, are usually found “behind the scenes” in databases rather than in public access systems. As a result, these types of metadata tend to be less standardized and more aligned with local requirements.

## Descriptive-

Descriptive metadata refers to information that supports discovery and identification of a resource (the who, what, when and where of a resource). It describes the content of the resource, associates various access points, and describes how the resource is related to other resources intellectually or within a hierarchy. In addition to bibliographic information, it may also describe physical attributes of the resource such as media type, dimension, and condition. Descriptive metadata is usually highly structured and often conforms to one or more standardized, published schemes, such as Dublin Core or MARC. Controlled vocabularies, thesauri, or authority files are commonly used to maintain consistency across the assignment of access points. Descriptive information is usually stored outside of the image file, often in separate catalogs or databases from technical information about the image file.

Although descriptive metadata may be stored elsewhere, it is recommended that some basic descriptive metadata (such as a caption or title) accompany the structural and technical metadata captured during production. The inclusion of this metadata can be useful for identification of files or groups of related files during quality review and other parts of the workflow, or for tracing the image back to the original.

Descriptive metadata is not specified in detail in this document; however, we recommend the use of the Dublin Core Metadata Element<sup>1</sup> set to capture minimal descriptive metadata information where metadata in another formal data standard does not exist. Metadata should be collected directly in Dublin Core; if it is not used for direct data collection, a mapping to Dublin Core elements is recommended. A mapping to Dublin Core from a richer, local metadata scheme already in use may also prove helpful for data exchange across other projects utilizing Dublin Core. Not all Dublin Core elements are required in order to create a valid Dublin Core record. However, we suggest that production master images be accompanied by the following elements at the very minimum:

### Minimum descriptive elements

Identifier	Primary identifier should be unique to the digital resource (at both object and file levels) Secondary identifiers might include identifiers related to the original (such as Still Picture ID) or Record Group number (for accessioned records)
Title/Caption	A descriptive name given to the original or the digital resource, or information that describes the content of the original or digital resource
Creator	(If available) Describes the person or organization responsible for the creation of the intellectual content of the resource
Publisher	Agency or agency acronym; Description of responsible agency or agent

These selected elements serve the purpose of basic identification of a file. Additionally, the Dublin Core elements “Format” (describes data types) and “Type” (describes limited record types) may be useful in certain database applications where sorting or filtering search results across many record genres or data types may be desirable. Any local fields that are important within the context of a particular project should also be captured to supplement Dublin Core fields so that valuable information is not lost. We anticipate that selection of metadata elements will come from more than one preexisting element set—elements can always be tailored to specific formats or local needs. Projects should support a modular approach to designing metadata to fit the specific requirements of the project. Standardizing on Dublin Core supplies baseline metadata that provides access to files, but this should not exclude richer metadata that extends beyond the Dublin Core set, if available.

For large-scale digitization projects, only minimal metadata may be affordable to record during capture, and is likely to consist of linking image identifiers to page numbers and indicating major structural divisions or anomalies of the resource (if applicable) for text documents. For photographs, capturing caption information (and Still Photo identifier) is ideal. For other non-textual materials, such as posters and maps, descriptive information taken directly from the item being scanned as well as a local identifier should be captured. If keying of captions into a database is prohibitive, if possible scan captions as part of the image itself. Although this information will not be searchable, it will serve to provide some basis of identification for the subject matter of the photograph. Recording of identifiers is important for uniquely identifying resources and is necessary for locating and managing them. It is likely that digital images will be associated with more than one identifier—for the image itself, for metadata or database records that describe the image, and for reference back to the original.

For images to be entered into NARA’s Archival Research Catalog (ARC), a more detailed complement of metadata is required. For a more detailed discussion of descriptive metadata requirements for digitization projects at NARA,

<sup>1</sup> Dublin Core Metadata Initiative, (<http://dublincore.org/usage/terms/dc/current-elements/>). The Dublin Core element set is characterized by simplicity in creation of records, flexibility, and extensibility. It facilitates description of all types of resources and is intended to be used in conjunction with other standards that may offer fuller descriptions in their respective domains.

we refer readers to NARA's Lifecycle Data Requirements Guide (LCDRG), at: [http://www.archives.gov/research\\_room/arc/arc\\_info/lifecycle\\_data\\_requirements.doc](http://www.archives.gov/research_room/arc/arc_info/lifecycle_data_requirements.doc) (June 2004), and NARA internal link- [http://www.nara-at-work.gov/archives\\_and\\_records\\_mgmt/archives\\_and\\_activities/accessioning\\_processing\\_description/lifecycle/index.html](http://www.nara-at-work.gov/archives_and_records_mgmt/archives_and_activities/accessioning_processing_description/lifecycle/index.html) (January 2002), which contains data elements developed for the archival description portion of the records lifecycle, and associates these elements with many different hierarchical levels of archival materials from record groups to items. The LCDRG also specifies rules for data entry. The LCDRG also requires a minimum set of other metadata to be recorded for raster image files at the file level, including technical metadata that enables images to display properly in the ARC interface.

Additionally, enough compatibility exists between Dublin Core and the data requirements that NARA has developed for archival description to provide a useful mapping between data elements, if a digital project requires that metadata also be managed locally (outside of ARC), perhaps in a local database or digital asset management system that supports data in Dublin Core. Please see Appendix C for a listing of mandatory elements identified in the Lifecycle Data Requirements Guide at the record group, series, file unit and item level, with Dublin Core equivalents.

Because ARC will be used as the primary source for descriptive information about the holdings of permanent records at NARA, we refer readers to the LCDRG framework rather than discuss Encoded Archival Description (EAD) of finding aids. NARA has developed its own hierarchical descriptive structure that relates to Federal records in particular, and therefore has not implemented EAD locally. However, because of the prevalence of the use of EAD in the wider archival and digitization communities, we have included a reference here. For more information on EAD, see the official EAD site at the Library of Congress at <http://lcweb.loc.gov/ead/>; as well as the Research Library Group's Best Practices Guidelines for EAD at <http://www.rlg.org/rlgead/eadguides.html>.

#### **Administrative-**

The Dublin Core set does not provide for administrative, technical, or highly structured metadata about different document types. Administrative metadata comprises both technical and preservation metadata, and is generally used for internal management of digital resources. Administrative metadata may include information about rights and reproduction or other access requirements, selection criteria or archiving policy for digital content, audit trails or logs created by a digital asset management system, persistent identifiers, methodology or documentation of the imaging process, or information about the source materials being scanned. In general, administrative metadata is informed by the local needs of the project or institution and is defined by project-specific workflows. Administrative metadata may also encompass repository-like information, such as billing information or contractual agreements for deposit of digitized resources into a repository.

For additional information, see Harvard University Library's Digital Repository Services (DRS) User Manual for Data Loading, Version 2.04 at [http://hul.harvard.edu/ois/systems/drs/drs\\_load\\_manual.pdf](http://hul.harvard.edu/ois/systems/drs/drs_load_manual.pdf), particularly Section 5.0, "DTD Element Descriptions" for application of administrative metadata in a repository setting; Making of America 2 (MOA2) Digital Object Standard: Metadata, Content, and Encoding at <http://www.cdlib.org/about/publications/CDLObjectStd-2001.pdf>; the Dublin Core also has an initiative for administrative metadata at <http://metadata.net/admin/draft-iannella-admin-01.txt> in draft form as it relates to descriptive metadata. The Library of Congress has defined a data dictionary for various formats in the context of METS, Data Dictionary for Administrative Metadata for Audio, Image, Text, and Video Content to Support the Revision of Extension Schemas for METS, available at <http://lcweb.loc.gov/rr/mopic/avprot/extension2.html>.

#### **Rights-**

Although metadata regarding rights management information is briefly mentioned above, it encompasses an important piece of administrative metadata that deserves further discussion. Rights information plays a key role in the context of digital imaging projects and will become more and more prominent in the context of preservation repositories, as strategies to act upon digital resources in order to preserve them may involve changing their structure, format, and properties. Rights metadata will be used both by humans to identify rights holders and legal status of a resource, and also by systems that implement rights management functions in terms of access and usage restrictions.

□

Because rights management and copyright are complex legal topics, the General Counsel's office (or a lawyer) should be consulted for specific guidance and assistance. The following discussion is provided for informational purposes only and should not be considered specific legal advice.

Generally, records created by employees of the Federal government as part of their routine duties, works for hire created under contract to the Federal government, and publications produced by the Federal government are all in

the public domain. However, it is not enough to assume that if NARA has physical custody of a record that it also owns the intellectual property in that record. NARA also has custody of other records, where copyright may not be so straightforward – such as personal letters written by private individuals, personal papers from private individuals, commercially published materials of all types, etc.—which are subject to certain intellectual property and privacy rights and may require additional permissions from rights holders. After transfer or donation of records to NARA from other federal agencies or other entities, NARA may either: own both the physical record and the intellectual property in the record; own the physical record but not the intellectual property; or the record is in the public domain. It is important to establish who owns or controls both the physical record and the copyright at the beginning of an imaging project, as this affects reproduction, distribution, and access to digital images created from these records. □

Metadata element sets for intellectual property and rights information are still in development, but they will be much more detailed than statements that define reproduction and distribution policies. At a minimum, rights-related metadata should include: the legal status of the record; a statement on who owns the physical and intellectual aspects of the record; contact information for these rights holders; as well as any restrictions associated with the copying, use, and distribution of the record. To facilitate bringing digital copies into future repositories, it is desirable to collect appropriate rights management metadata at the time of creation of the digital copies. At the very least, digital versions should be identified with a designation of copyright status, such as: “public domain;” “copyrighted” (and whether clearance/permissions from rights holder has been secured); “unknown;” “donor agreement/contract;” etc.

Preservation metadata dealing with rights management in the context of digital repositories will likely include detailed information on the types of actions that can be performed on data objects for preservation purposes and information on the agents or rights holders that authorize such actions or events.

For an example of rights metadata in the context of libraries and archives, a rights extension schema has recently been added to the Metadata Encoding and Transmission Standard (METS), which documents metadata about the intellectual rights associated with a digital object. □ This extension schema contains three components: a rights declaration statement; detailed information about rights holders; and context information, which is defined as “who has what permissions and constraints within a specific set of circumstances.” The schema is available at: <http://www.loc.gov/standards/rights/METSRights.xsd>

For additional information on rights management, see: Peter B. Hirtle, “Archives or Assets?” at <http://techreports.library.cornell.edu:8081/Dienst/UI/1.0/Display/cul.lib/2003-2>; June M. Besek, Copyright Issues Relevant to the Creation of a Digital Archive: A Preliminary Assessment, January 2003 at <http://www.clir.org/pubs/reports/pub112/contents.html>; Adrienne Muir, “Copyright and Licensing for Digital Preservation,” at <http://www.cilip.org.uk/update/issues/jun03/article2june.html>; Karen Coyle, Rights Expression Languages, A Report to the Library of Congress, February 2004, available at [http://www.loc.gov/standards/Coylereport\\_final1single.pdf](http://www.loc.gov/standards/Coylereport_final1single.pdf); MPEG-21 Overview v.5 contains a discussion on intellectual property and rights at <http://www.chiariglione.org/mpeg/standards/mpeg-21/mpeg-21.htm>; for tables that reference when works pass into the public domain, see Peter Hirtle, “When Works Pass Into the Public Domain in the United States: Copyright Term for Archivists and Librarians,” at [http://www.copyright.cornell.edu/training/Hirtle\\_Public\\_Domain.htm](http://www.copyright.cornell.edu/training/Hirtle_Public_Domain.htm) and Mary Minow, “Library Digitization Projects: Copyrighted Works that have Expired into the Public Domain” at <http://www.librarylaw.com/DigitizationTable.htm>; and for a comprehensive discussion on libraries and copyright, see: Mary Minow, *Library Digitization Projects and Copyright* at <http://www.llrx.com/features/digitization.htm>.

### Technical-

Technical metadata refers to information that describes attributes of the digital image (not the analog source of the image) and helps to ensure that images will be rendered accurately. It supports content preservation by providing information needed by applications to use the file and to successfully control the transformation or migration of images across or between file formats. Technical metadata also describes the image capture process and technical environment, such as hardware and software used to scan images, as well as file format-specific information, image quality, and information about the source object being scanned, which may influence scanning decisions. Technical metadata helps to ensure consistency across a large number of files by enforcing standards for their creation. At a minimum, technical metadata should capture the information necessary to render, display, and use the resource.

Technical metadata is characterized by information that is both objective and subjective—attributes of image quality that can be measured using objective tests as well as information that may be used in a subjective assessment of an image’s value. Although tools for automatic creation and capture of many objective components



are badly needed, it is important to determine what metadata should be highly structured and useful to machines, as opposed to what metadata would be better served in an unstructured, free-text note format. The more subjective data is intended to assist researchers in the analysis of digital resource or imaging specialists and preservation administrators in determining long-term value of a resource.

In addition to the digital image, technical metadata will also need to be supplied for the metadata record itself if the metadata is formatted as a text file or XML document or METS document, for example. In this sense, technical metadata is highly recursive, but necessary for keeping both images and metadata understandable over time.

Requirements for technical metadata will differ for various media formats. For digital still images, we refer to the NISO Data Dictionary - Technical Metadata for Digital Still Images at [http://www.niso.org/standards/resources/Z39\\_87\\_trial\\_use.pdf](http://www.niso.org/standards/resources/Z39_87_trial_use.pdf). It is a comprehensive technical metadata set based on the Tagged Image File Format specification, and makes use of the data that is already captured in file headers. It also contains metadata elements important to the management of image files that are not present in header information, but that could potentially be automated from scanner/camera software applications. An XML schema for the NISO technical metadata has been developed at the Library of Congress called MIX (Metadata in XML), which is available at <http://www.loc.gov/standards/mix/>.

See also the TIFF 6.0 Specification at <http://partners.adobe.com/asn/developer/pdfs/tn/TIFF6.pdf> as well as the Digital Imaging Group's DIG 35 metadata element set at [http://www.i3a.org/i\\_dig35.html](http://www.i3a.org/i_dig35.html); and Harvard University Library's Administrative Metadata for Digital Still Images data dictionary at [http://hul.harvard.edu/ldi/resources/ImageMetadata\\_v2.pdf](http://hul.harvard.edu/ldi/resources/ImageMetadata_v2.pdf).

A new initiative led by the Research Libraries Group called "*Automatic Exposure: Capturing Technical Metadata for Digital Still Images*" is investigating ways to automate the capture of technical metadata specified in the NISO Z39.87 draft standard. The initiative seeks to build automated capture functionality into scanner and digital camera hardware and software in order to make this metadata readily available for transfer into repositories and digital asset management systems, as well as to make metadata capture more economically viable by reducing the amount of manual entry that is required. This implies a level of trust that the metadata that is automatically captured and internal to the file is inherently correct.

See <http://www.rlg.org/longterm/autotechmetadata.html> for further discussion of this initiative, as well as the discussion on Image Quality Assessment, below.

Initiatives such as the Global Digital Format Registry (<http://hul.harvard.edu/gdfr/>) could potentially help in reducing the number of metadata elements that need to be recorded about a file or group of files regarding file format information necessary for preservation functions. Information maintained in the Registry could be pointed to instead of recorded for each file or batch of files.

### **Structural-**

Structural metadata describes the relationships between different components of a digital resource. It ties the various parts of a digital resource together in order to make a useable, understandable whole. One of the primary functions of structural metadata is to enable display and navigation, usually via a page-turning application, by indicating the sequence of page images or the presence of multiple views of a multi-part item. In this sense, structural metadata is closely related to the intended behaviors of an object. Structural metadata is very much informed by how the images will be delivered to the user as well as how they will be stored in a repository system in terms of how relationships among objects are expressed.

Structural metadata often describes the significant intellectual divisions of an item (such as chapter, issue, illustration, etc.) and correlates these divisions to specific image files. These explicitly labeled access points help to represent the organization of the original object in digital form. This does not imply, however, that the digital must always imitate the organization of the original—especially for non-linear items, such as folded pamphlets. Structural metadata also associates different representations of the same resource together, such as production master files with their derivatives, or different sizes, views, or formats of the resource.

Example structural metadata might include whether the resource is simple or complex (multi-page, multi-volume, has discrete parts, contains multiple views); what the major intellectual divisions of a resource are (table of contents, chapter, musical movement); identification of different views (double-page spread, cover, detail); the extent (in files, pages, or views) of a resource and the proper sequence of files, pages and views; as well as different technical (file formats, size), visual (pre- or post-conservation treatment), intellectual (part of a larger collection or work), and use (all instances of a resource in different formats--TIFF files for display, PDF files for printing, OCR file for full text searching) versions.

File names and organization of files in system directories comprise structural metadata in its barest form. Since meaningful structural metadata can be embedded in file and directory names, consideration of where and how structural metadata is recorded should be done upfront. See Section V. Storage for further discussion on this topic.

No widely adopted standards for structural metadata exist since most implementations of structural metadata are at the local level and are very dependent on the object being scanned and the desired functionality in using the object. Most structural metadata is implemented in file naming schemes and/or in databases that record the order and hierarchy of the parts of an object so that they can be identified and reassembled back into their original form.

The Metadata Encoding and Transmission Standard (METS) is often discussed in the context of structural metadata, although it is inclusive of other types of metadata as well. METS provides a way to associate metadata with the digital files they describe and to encode the metadata and the files in a standardized manner, using XML. METS requires structural information about the location and organization of related digital files to be included in the METS document. Relationships between different representations of an object as well as relationships between different hierarchical parts of an object can be expressed. METS brings together a variety of metadata about an object all into one place by allowing the encoding of descriptive, administrative, and structural metadata. Metadata and content information can either be wrapped together within the METS document, or pointed to from the METS document if they exist in externally disparate systems. METS also supports extension schemas for descriptive and administrative metadata to accommodate a wide range of metadata implementations. Beyond associating metadata with digital files, METS can be used as a data transfer syntax so objects can easily be shared; as a Submission Information Package, an Archival Information Package, and a Dissemination Information Package in an OAIS-compliant repository (see below); and also as a driver for applications, such as a page turner, by associating certain behaviors with digital files so that they can be viewed, navigated, and used. Because METS is primarily concerned with structure, it works best with “library-like” objects in establishing relationships among multi-page or multi-part objects, but it does not apply as well to hierarchical relationships that exist in collections within an archival context.

See <http://www.loc.gov/standards/mets/> for more information on METS.

#### **Behavior-**

Behavior metadata is often referred to in the context of a METS object. It associates executable behaviors with content information that define how a resource should be utilized or presented. Specific behaviors might be associated with different genres of materials (books, photographs, Powerpoint presentations) as well as with different file formats. Behavior metadata contains a component that abstractly defines a set of behaviors associated with a resource as well as a “mechanism” component that points to executable code (software applications) that then performs a service according to the defined behavior. The ability to associate behaviors or services with digital resources is one of the attributes of a METS object and is also part of the “digital object architecture” of the Fedora digital repository system. See <http://www.fedora.info/documents/master-spec-12.20.02.pdf> for a discussion of Fedora and digital object behaviors.

#### **Preservation-**

Preservation metadata encompasses all information necessary to manage and preserve digital assets over time. Preservation metadata is usually defined in the context of the OAIS reference model (Open Archival Information System, <http://ssdoo.gsfc.nasa.gov/nost/isoas/overview.html>), and is often linked to the functions and activities of a repository. It differs from technical metadata in that it documents processes performed over time (events or actions taken to preserve data and the outcomes of these events) as opposed to explicitly describing provenance (how a digital resource was created) or file format characteristics, but it does encompass all types of the metadata mentioned above, including rights information. Although preservation metadata draws on information recorded earlier (technical and structural metadata would be necessary to render and reassemble the resource into an understandable whole), it is most often associated with analysis of and actions performed on a resource after submission to a repository. Preservation metadata might include a record of changes to the resource, such as transformations or conversions from format to format, or indicate the nature of relationships among different resources.

Preservation metadata is information that will assist in preservation decision-making regarding the long-term value of a digital resource and the cost of maintaining access to it, and will help to both facilitate archiving strategies for digital images as well as support and document these strategies over time. Preservation metadata is commonly linked with digital preservation strategies such as migration and emulation, as well as more “routine” system-level actions such as copying, backup, or other automated processes carried out on large numbers of objects. These strategies will rely on all types of pre-existing metadata and will also generate and record new

metadata about the object. It is likely that this metadata will be both machine-processable and “human-readable” at different levels to support repository functions as well as preservation policy decisions related to these objects.

In its close link to repository functionality, preservation metadata may reflect or even embody the policy decisions of a repository; but these are not necessarily the same policies that apply to preservation and reformatting in a traditional context. The extent of metadata recorded about a resource will likely have an impact on future preservation options to maintain it. Current implementations of preservation metadata are repository- or institution-specific. We anticipate that a digital asset management system may provide some basic starter functionality for low-level preservation metadata implementation, but not to the level of a repository modeled on the OAIS.

See also A Metadata Framework to Support the Preservation of Digital Objects at [http://www.oclc.org/research/projects/pmwg/pm\\_framework.pdf](http://www.oclc.org/research/projects/pmwg/pm_framework.pdf) and Preservation Metadata for Digital Objects: A Review of the State of the Art at [http://www.oclc.org/research/projects/pmwg/presmeta\\_wp.pdf](http://www.oclc.org/research/projects/pmwg/presmeta_wp.pdf), both by the OCLC/RLG Working Group on Preservation Metadata, for excellent discussions of preservation metadata in the context of the OAIS model. A new working group, “Preservation Metadata: Implementation Strategies,” is working on developing best practices for implementing preservation metadata and on the development of a recommended core set of preservation metadata. Their work can be followed at <http://www.oclc.org/research/projects/pmwg/>.

For some examples of implementations of preservation metadata element sets at specific institutions, see: OCLC Digital Archive Metadata, at [http://www.oclc.org/support/documentation/pdf/da\\_metadata\\_elements.pdf](http://www.oclc.org/support/documentation/pdf/da_metadata_elements.pdf); Florida Center for Library Automation Preservation Metadata, at [http://www.fcla.edu/digitalArchive/pdfs/Archive\\_data\\_dictionary20030703.pdf](http://www.fcla.edu/digitalArchive/pdfs/Archive_data_dictionary20030703.pdf); Technical Metadata for the Long-Term Management of Digital Materials, at [http://dvl.dtic.mil/metadata\\_guidelines/TechMetadata\\_26Mar02\\_1400.pdf](http://dvl.dtic.mil/metadata_guidelines/TechMetadata_26Mar02_1400.pdf); and The National Library of New Zealand, Metadata Standard Framework, Preservation Metadata, at [http://www.natlib.govt.nz/files/4initiatives metaschema\\_revised.pdf](http://www.natlib.govt.nz/files/4initiatives metaschema_revised.pdf).

#### **Image quality assessment (NARA-NWTS Digital Imaging Lab proposed metadata requirement)-**

The technical metadata specified in the *NISO Data Dictionary - Technical Metadata for Digital Still Images* contains many metadata fields necessary for the long-term viability of the image file. However, we are not convinced that it goes far enough in providing information necessary to make informed preservation decisions regarding the value and quality of a digital still raster image. Judgments about the quality of an image require a visual inspection of the image, a process that cannot be automated. Quality is influenced by many factors—such as the source material from which the image was scanned, the devices used to create the image, any subsequent processing done to the image, compression, and the overall intended use of the image. Although the data dictionary includes information regarding the analog source material and the scanning environment in which the image was created, we are uncertain whether this information is detailed enough to be of use to administrators, curators, and others who will need to make decisions regarding the value and potential use of digital still images. The value of metadata correlates directly with the future use of the metadata. It seems that most technical metadata specified in the NISO data dictionary is meant to be automatically captured from imaging devices and software and intended to be used by systems to render and process the file, not necessarily used by humans to make decisions regarding the value of the file. The metadata can make no guarantee about the quality of the data. Even if files appear to have a full complement of metadata and meet the recommended technical specifications as outlined in these *Technical Guidelines*, there may still be problems with the image file that cannot be assessed without some kind of visual inspection.

The notion of an image quality assessment was partly inspired by the National Library of Medicine Permanence Ratings (see <http://www.nlm.nih.gov/pubs/reports/permanence.pdf> and <http://www.rlg.org/events/pres-2000/byrnes.html>), a rating for resource permanence or whether the content of a resource is anticipated to change over time. However, we focused instead on evaluating image quality and this led to the development of a simplified rating system that would: indicate a quality level for the suitability of the image as a production master file (its suitability for multiple uses or outputs), and serve as a potential metric that could be used in making preservation decisions about whether an image is worth maintaining over time. If multiple digital versions of a single record exist, then the image quality assessment rating may be helpful for deciding which version(s) to keep.

The rating is linked to image defects introduced in the creation of intermediates and/or introduced during digitization and image processing, and to the nature and severity of the defects based on evaluating the digital



images on-screen at different magnifications. In essence, a “good” rating for image files implies an appropriate level of image quality that warrants the effort to maintain them over time.

The image quality assessment takes into account the attributes that influence specifications for scanning a production master image file: format, size, intended use, significant characteristics of the original that should be maintained in the scan, and the quality and characteristics of the source material being scanned. This rating system could later be expanded to take into account other qualities such as object completeness (are all pages or only parts of the resource scanned?); the source of the scan (created in-house or externally provided?); temporal inconsistencies (scanned at different times, scanned on different scanners, scan of object is pre- or post-conservation treatment?), and enhancements applied to the image for specific purposes (for exhibits, cosmetic changes among others).

This rating is not meant to be a full technical assessment of the image, but rather an easy way to provide information that supplements existing metadata about the format, intent, and use of the image, all of which could help determine preservation services that could be guaranteed and associated risks based on the properties of the image. We anticipate a preservation assessment will to be carried out later in the object’s lifecycle based on many factors, including the image quality assessment.

Image quality rating metadata is meant to be captured at the time of scanning, during processing, and even at the time of ingest into a repository. When bringing batches or groups of multiple image files into a repository that do not have individual image quality assessment ratings, we recommend visually evaluating a random sample of images and applying the corresponding rating to all files in appropriate groups of files (such as all images produced on the same model scanner or all images for a specific project).

Record whether the image quality assessment rating was applied as an individual rating or as a batch rating. If a batch rating, then record how the files were grouped.

**Image Quality Assessment Ratings**

Rating	Description	Use	Defect Identification
2	<ul style="list-style-type: none"> <li>No obvious visible defects in image when evaluating the histogram and when viewed on-screen, including individual color channels, at:</li> </ul> <p>100% or 1:1 pixel display (micro) and actual size (1"=1") and full image (global)</p>	Generally, image suitable as production master file.	
1	<ul style="list-style-type: none"> <li>No obvious visible defects in image when evaluating the histogram and when viewed on-screen, including individual color channels, at:</li> </ul> <p>actual size (1"=1") and full image (global)</p> <ul style="list-style-type: none"> <li>Minor defects visible at:</li> </ul> <p>100% or 1:1 pixel display (micro)</p>	Image suitable for less critical applications (e.g., suitable for output on typical inkjet and photo printers) or for specific intents (e.g., for access images, uses where these defects will not be critical).	<p>Identify and record the defects relating to intermediates and the digital images – illustrative examples:</p> <p>Intermediates-</p> <ul style="list-style-type: none"> <li>out of focus copy negative</li> <li>scratched microfilm</li> <li>surface dirt</li> <li>etc.</li> </ul> <p>Digital images-</p> <ul style="list-style-type: none"> <li>oversharpened image</li> <li>excessive noise</li> <li>posterization and quantization artifacts</li> <li>compression artifacts</li> <li>color channel misregistration</li> <li>color fringing around text</li> <li>etc.</li> </ul>
0	<ul style="list-style-type: none"> <li>Obvious visible defects when evaluating the histogram and when viewed on-screen, including individual color channels, at:</li> </ul> <p>100% or 1:1 pixel display (micro) and/or actual size (1"=1") and/or full image (global)</p>	<p>Image unsuitable for most applications.</p> <p>In some cases, despite the low rating, image may warrant long-term retention if-</p> <ul style="list-style-type: none"> <li>image is the "best copy available"</li> <li>known to have been produced for a very specific output</li> </ul>	<p>Identify and record the defects relating to intermediates and the digital images - illustrative examples:</p> <p>Intermediates-</p> <ul style="list-style-type: none"> <li>all defects listed above</li> <li>uneven illumination during photography</li> <li>under- or over-exposed copy transparencies</li> <li>reflections in encapsulation</li> <li>etc.</li> </ul> <p>Digital images-</p> <ul style="list-style-type: none"> <li>all defects listed above</li> <li>clipped highlight and/or clipped shadow detail</li> <li>uneven illumination during scanning</li> <li>reflections in encapsulation</li> <li>image cropped</li> <li>etc.</li> </ul>

As stated earlier, image quality assessment rating is applied to the digital image but is also linked to information regarding the source material from which it was scanned. Metadata about the image files includes a placeholder for information regarding source material, which includes a description of whether the analog source is the original or an intermediate, and if so, what kind of intermediate (copy, dupe, microfilm, photocopy, etc.) as well as the source format. Knowledge of deficiencies in the source material (beyond identifying the record type and format) helps to inform image quality assessment as well.

The practicality of implementing this kind of assessment has not yet been tested, especially since it necessitates a review of images at the file level. Until this conceptual approach gains broader acceptance and consistent implementation within the community, quality assessment metadata may only be useful for local preservation decisions. As the assessment is inherently technical in nature, a basic foundation in photography and imaging is

helpful in order to accurately evaluate technical aspects of the file, as well as to provide a degree of trustworthiness in the reviewer and in the rating that is applied.

### **Records management/recordkeeping-**

Another type of metadata, relevant to the digitization of federal records in particular, is records management metadata. Records management metadata is aligned with administrative-type metadata in that its function is to assist in the management of records over time; this information typically includes descriptive (and, more recently, preservation) metadata as a subset of the information necessary to both find and manage records. Records management metadata is usually discussed in the context of the systems or domains in which it is created and maintained, such as Records Management Application (RMA) systems. □ This includes metadata about the records as well as the organizations, activities, and systems that create them. □ The most influential standard in the United States on records management metadata is the Department of Defense's Design Criteria Standard for Electronic Records Management Software Applications (DOD 5015.2) at <http://www.dtic.mil/whs/directives/corres/html/50152std.htm>. This standard focuses on minimum metadata elements a RMA should capture and maintain, defines a set of metadata elements at the file plan, folder, and record levels, and generally discusses the functionality that an RMA should have as well as the management, tracking, and integration of metadata that is held in RMAs. □

Records Management metadata should document whether digital images are designated as permanent records, new records, temporary records, reference copies, or are accorded a status such as "indefinite retention." A determination of the status of digital images in a records management context should be made upfront at the point of creation of the image, as this may have an effect on the level and detail of metadata that will be gathered for a digital object to maintain its significant properties and functionality over the long term. Official designation of the status of the digital images will be an important piece of metadata to have as digital assets are brought into a managed system, such as NARA's Electronic Records Archive (ERA), which will have extensive records management capabilities. □

In addition to a permanent or temporary designation, records management metadata should also include documentation on any access and/or usage restrictions for the image files. Metadata documenting restrictions that apply to the images could become essential if both unrestricted and restricted materials and their metadata are stored and managed together in the same system, as these files will possess different maintenance, use and access requirements. Even if restricted files are stored on a physically separate system for security purposes, metadata about these files may not be segregated and should therefore include information on restrictions.

For digitization projects done under NARA 816 guidance, we assume classified, privacy restricted, and any records with other restrictions will not be selected for digitization. However, records management metadata should still include documentation on access and usage restrictions - even unrestricted records should be identified as "unrestricted." This may be important metadata to express at the system level as well, as controls over access to and use of digital resources might be built directly into a delivery or access system. □

In the future, documentation on access and use restrictions relevant to NARA holdings might include information such as: "classified" (which should be qualified by level of classification); "unclassified" or "unrestricted;" "declassified;" and "restricted," (which should be qualified by a description of the restrictions, i.e., specific donor-imposed restrictions), for example. Classification designation will have an impact on factors such as physical storage (files may be physically or virtually stored separately), who has access to these resources, and different maintenance strategies.

Basic records management metadata about the image files will facilitate bringing them into a formal system and will inform functions such as scheduling retention timeframes, how the files are managed within a system, what types or levels of preservation services can be performed, or how they are distributed and used by researchers, for example.

### **Tracking-**

Tracking metadata is used to control or facilitate the particular workflow of an imaging project during different stages of production. Elements might reflect the status of digital images as they go through different stages of the workflow (batch information and automation processes, capture, processing parameters, quality control, archiving, identification of where/media on which files are stored); this is primarily internally-defined metadata that serves as documentation of the project and may also serve also serve as a statistical source of information to track and report on progress of image files. Tracking metadata may exist in a database or via a directory/folder system.

**Meta-metadata-**

Although this information is difficult to codify, it usually refers to metadata that describes the metadata record itself, rather than the object it is describing, or to high-level information about metadata “policy” and procedures, most often on the project level. Meta-metadata documents information such as who records the metadata, when and how it gets recorded, where it is located, what standards are followed, and who is responsible for modification of metadata and under what circumstances.

It is important to note that metadata files yield “master” records as well. These non-image assets are subject to the same rigor of quality control and storage as master image files. Provisions should be made for the appropriate storage and management of the metadata files over the long term.

**Assessment of Metadata Needs for Imaging Projects:**

Before beginning any scanning, it is important to conduct an assessment both of existing metadata and metadata that will be needed in order to develop data sets that fit the needs of the project. The following questions frame some of the issues to consider:

- *Does metadata already exist in other systems (database, finding aid, on item itself) or structured formats (Dublin Core, local database)?*

If metadata already exists, can it be automatically derived from these systems, pointed to from new metadata gathered during scanning, or does it require manual input? Efforts to incorporate existing metadata should be pursued. It is also extremely beneficial if existing metadata in other systems can be exported to populate a production database prior to scanning. This can be used as base information needed in production tracking, or to link item level information collected at the time of scanning to metadata describing the content of the resource. An evaluation of the completeness and quality of existing metadata may need to be made to make it useful (e.g., what are the characteristics of the data content, how is it structured, can it be easily transformed?)

It is likely that different data sets with different functions will be developed, and these sets will exist in different systems. However, efforts to link together metadata in disparate systems should be made so that it can be reassembled into something like a METS document, an Archival XML file for preservation, or a Presentation XML file for display, depending on what is needed. Metadata about digital images should be integrated into peer systems that already contain metadata about both digital and analog materials. By their nature, digital collections should not be viewed as something separate from non-digital collections. Access should be promoted across existing systems rather than building a separate stand-alone system.

- *Who will capture metadata?*

Metadata is captured by systems or by humans and is intended for system or for human use. For example, certain preservation metadata might be generated by system-level activities such as data backup or copying. Certain technical metadata is used by applications to accurately render an image. In determining the function of metadata elements, it is important to establish whether this information is important for use by machines or by people. If it is information that is used and/or generated by systems, is it necessary to explicitly record it as metadata? What form of metadata is most useful for people? Most metadata element sets include less structured, note or comment-type fields that are intended for use by administrators and curators as data necessary for assessment of the provenance, risk of obsolescence, and value inherent to a particular class of objects. Any data, whether generated by systems or people, that is necessary to understand a digital object, should be considered as metadata that may be necessary to formally record. But because of the high costs of manually generating metadata and tracking system-level information, the use and function of metadata elements should be carefully considered. Although some metadata can be automatically captured, there is no guarantee that this data will be valuable over the long term.

- *How will metadata be captured?*

Metadata capture will likely involve a mix of manual and automated entry. Descriptive and structural metadata creation is largely manual; some may be automatically generated through OCR processes to create indexes or full text; some technical metadata may be captured automatically from imaging software and devices; more sophisticated technical metadata, such as image quality assessment metadata used to inform preservation decisions, will require visual analysis and manual input.

An easy-to-use and customizable database or asset management system with a graphical and intuitive front end, preferably structured to mimic a project’s particular metadata workflow, is desirable and will make for more efficient metadata creation.

○ *When will metadata be collected?*

Metadata is usually collected incrementally during the scanning process and will likely be modified over time. At least, start with a minimal element set that is known to be needed and add additional elements later, if necessary.

Assignment of unique identifier or naming scheme should occur upfront. We also recommend that descriptive metadata be gathered prior to capture to help streamline the scanning process. It is usually much more difficult to add new metadata later on, without consultation of the originals. The unique file identifier can then be associated with a descriptive record identifier, if necessary.

A determination of what structural metadata elements to record should also occur prior to capture, preferably during the preparation of materials for capture or during collation of individual items. Information about the hierarchy of the collection, the object types, and the physical structure of the objects should be recorded in a production database prior to scanning. The structural parts of the object can be linked to actual content files during capture. Most technical metadata is gathered at the time of scanning. Preservation metadata is likely to be recorded later on, upon ingest into a repository.

○ *Where will the metadata be stored?*

Metadata can be embedded within the resource (such as an image header or file name) or can reside in a system external to the resource (such as a database) or both. Metadata can be also encapsulated with the file itself, such as with the Metadata Encoded Transmission Standard (METS). The choice of location of metadata should encourage optimal functionality and long-term management of the data.

Header data consists of information necessary to decode the image, and has somewhat limited flexibility in terms of data values that can be put into the fields. Header information accommodates more technical than descriptive metadata (but richer sets of header data can be defined depending on the image file format). The advantage is that metadata remains with the file, which may result in more streamlined management of content and metadata over time. Several tags are saved automatically as part of the header during processing, such as dimensions, date, and color profile information, which can serve as base-level technical metadata requirements. However, methods for storing information in file format headers are very format-specific and data may be lost in conversions from one format to another. Also, not all applications may be able to read the data in headers. Information in headers should be manually checked to see if data has transferred correctly or has not been overwritten during processing. Just because data exists in headers does not guarantee that it has not been altered or has been used as intended. Information in headers should be evaluated to determine if it has value. Data from image headers can be extracted and imported into a database; a relationship between the metadata and the image must then be established and maintained.

Storing metadata externally to the image in a database provides more flexibility in managing, using, and transforming it and also supports multi-user access to the data, advanced indexing, sorting, filtering, and querying. It can better accommodate hierarchical descriptive information and structural information about multi-page or complex objects, as well as importing, exporting, and harvesting of data to external systems or other formats, such as XML. Because metadata records are resources that need to be managed in their own right, there is certainly benefit to maintaining metadata separately from file content in a managed system. Usually a unique identifier or the image file name is used to link metadata in an external system to image files in a directory.

We recommend that metadata be stored both in image headers as well as in an external database to facilitate migration and repurposing of the metadata. References between the metadata and the image files can be maintained via persistent identifiers. A procedure for synchronization of changes to metadata in both locations is also recommended, especially for any duplicated fields. This approach allows for metadata redundancy in different locations and at different levels of the digital object for ease of use (image file would not have to be accessed to get information; most header information would be extracted and added into an external system). Not all metadata should be duplicated in both places (internal and external to the file). Specific metadata is required in the header so that applications can interpret and render the file; additionally, minimal descriptive metadata such as a unique identifier or short description of the content of the file should be embedded in header information in case the file becomes disassociated from the tracking system or repository. Some applications and file formats offer a means to store metadata within the file in an intellectually structured manner, or allow the referencing of standardized schemes, such as Adobe XMP or the XML metadata boxes in the JPEG 2000 format. Otherwise, most metadata will reside in external databases, systems, or registries.

○ *How will the metadata be stored?*

Metadata schemes and data dictionaries define the content rules for metadata creation, but not the format in which metadata should be stored. Format may partially be determined by where the metadata is stored (file headers,

relational databases, spreadsheets) as well as the intended use of the metadata—does it need to be human-readable, or indexed, searched, shared, and managed by machines? How the metadata is stored or encoded is usually a local decision. Metadata might be stored in a relational database or encoded in XML, such as in a METS document, for example. Guidelines for implementing Dublin Core in XML are also available at: <http://dublincore.org/documents/2002/09/09/dc-xml-guidelines/>.

Adobe's Extensible Metadata Platform (XMP) is another emerging, standardized format for describing where metadata can be stored and how it can be encoded, thus facilitating exchange of metadata across applications. The XMP specification provides both a data model and a storage model. Metadata can be embedded in the file in header information or stored in XML "packets" (these describe how the metadata is embedded in the file). XMP supports the capture of (primarily technical) metadata during content creation and modification and embeds this information in the file, which can then be extracted later into a digital asset management system or database or as an XML file. If an application is XMP enabled or aware (most Adobe products are), this information can be retained across multiple applications and workflows. XMP supports customization of metadata to allow for local field implementation using their Custom File Info Panels application. XMP supports a number of internal schemas, such as Dublin Core and EXIF (a metadata standard used for image files, particularly by digital cameras), as well as a number of external extension schemas. The RLG initiative, "*Automatic Exposure: Capturing Technical Metadata for Digital Still Images*," mentioned earlier is considering the use of XMP to embed technical metadata in image files during capture and is developing a Custom File Info Panel for NISO Z39.87 technical metadata. XMP does not guarantee the automatic entry of all necessary metadata (several fields will still require manual entry, especially local fields), but allows for more complete customized, and accessible metadata about the file.

See <http://www.adobe.com/products/xmp/main.html> for more detailed information on the XMP specification and other related documents.

o *Will the metadata need to interact or be exchanged with other systems?*

This requirement reinforces the need for standardized ways of recording metadata so that it will meet the requirements of other systems. Mapping from an element in one scheme to an analogous element in another scheme will require that the meaning and structure of the data is shareable between the two schemes, in order to ensure usability of the converted metadata. Metadata will also have to be stored in or assembled into a document format, such as XML, that promotes easy exchange of data. METS-compliant digital objects, for example, promote interoperability by virtue of their standardized, "packaged" format.

o *At what level of granularity will the metadata be recorded?*

Will metadata be collected at the collection level, the series level, the imaging project level, the item (object) level, or file level? Although the need for more precise description of digital resources exists so that they can be searched and identified, for many large-scale digitization projects, this is not realistic. Most collections at NARA are neither organized around nor described at the individual item level, and cannot be without significant investment of time and cost. Detailed description of records materials is often limited by the amount of information known about each item, which may require significant research into identification of subject matter of a photograph, for example, or even what generation of media format is selected for scanning. Metadata will likely be derived from and exist on a variety of levels, both logical and file, although not all levels will be relevant for all materials. Certain information required for preservation management of the files will be necessary at the individual file level. An element indicating level of aggregation (e.g., item, file, series, collection) at which metadata applies can be incorporated, or the relational design of the database may reflect the hierarchical structure of the materials being described.

o *Adherence to agreed-upon conventions and terminology?*

We recommend that standards, if they exist and apply, be followed for the use of data elements, data values, and data encoding. Attention should be paid to how data is entered into fields and whether controlled vocabularies have been used, in case transformation is necessary to normalize the data.

## Local Implementation:

Because most of what we scan comes to the Imaging Lab on an item-by-item basis, we are capturing minimal descriptive and technical metadata at the item level only during the image capture and processing stage. Until a structure into which we can record hierarchical information both about the objects being scanned and their higher-level collection information is in place, we are entering basic metadata in files using Adobe Photoshop. Information about the file is added to the IPTC (International Press Telecommunications Council) fields in Photoshop in anticipation of mapping these values to an external database. The IPTC fields are used as placeholder fields only. This information is embedded in the file using Adobe XMP (Extensible Metadata Platform: <http://www.adobe.com/products/xmp/main.html>). Primary identifier is automatically imported into the "File

Info” function in Photoshop from our scanning software. We anticipate implementing the Custom Panel Description File Format feature available in XMP to define our own metadata set and then exporting this data into an asset management system, since the data will be stored in easily migratable XML packets.

The following tables outline minimal descriptive, technical, and structural metadata that we are currently capturing at the file level (table indicates the elements that logically apply at the object level):

#### Descriptive/Structural Placeholder Fields – Logical and/or File Attributes

Element Name		Note	Level (Object, File) of Metadata
Primary Identifier		Unique identifier (numerical string) of the digital image. This identifier also serves as the identifier for an associated descriptive metadata record in an external database. May be derived from an existing scheme. This identifier is currently “manually” assigned. We anticipate a “machine” assigned unique identifier to be associated with each image as it is ingested into a local repository system; this will be more like a “persistent identifier.” Since multiple identifiers are associated with one file, it is likely that this persistent identifier will be the cardinal identifier for the image.	Object, File
Secondary Identifier(s)		Other unique identifier(s) associated with the original	Object, File
Title		Title [informal or assigned] or caption associated with the resource	Object
Record Group ID		Record Group Identifier (if known)	Object
Record Group Descriptor		Title of Record Group (if known)	Object
Series		Title of Series (if known)	Object
Box or Location		Box Number or Location (if known)	Object
Structural view or page (sequence)		Description of view, page number, or file number	File
Publisher		Owner or Producer of image. Default is “U.S. National Archives”	Object
Source*	Text	Generation   Media	Object
	Film	Generation   Format   Color Mode   Media   Creation Date	
	Photo Print	Color Mode   Media	
	Digital Photo	Not yet determined; may include Generation; Dimensions; Capture Mode/Settings; Quality Level; Compression Level, etc.	

\*Describes physical attributes of the source material that may assist in interpretation of image quality; describes capture and processing decisions; or indicates known problems with the original media that may affect the quality of the scan. A controlled vocabulary is used for these fields. We feel that it is important to record source object information in technical metadata. Knowledge of the source material will inform image quality assessment and future preservation decisions. For images derived from another digital image, source information will be described in a relationship field, most likely from a set of typed relationships (e.g., “derived from”).



Technical metadata is currently entered into an external project database to describe specific derivative files. We anticipate that this information will map up to attributes of the production master files. The following table describes suggested minimum technical metadata fields for production masters.

Example technical metadata – File Attributes (some generated by file header) – All elements apply at file level -

Element Name	Note
Copy	“Role,” “function,” or “class” of the image (e.g., production master, delivery, or print-optimized derivative). Currently this functional designation is also embedded in the file identifier. This element may serve to indicate level of preservation service required.
File format type/Version	(e.g., TIFF, JPEG)
Location	Pointer to local file directory where image is stored
Image creation date	YYYY-MM-DD format
Photographer/Operator	Producer of image (name of scanner operator)
Compression Type/Level	Type and Level of compression applied (Adobe Photoshop-specific setting)
Color Mode	(e.g., RGB, Grayscale)
Gamma Correction	Default value is 2.2
Color Calibration	ICC Profile. Default value is AdobeRGB 1998 for RGB images and Grayscale 2.2 for grayscale images.
Pixel Array	Pixel width x height
Spatial Resolution	Expressed in ppi (e.g., 300)
Image quality*	Uses controlled values from authority table. Documents image quality characteristics that may influence future decisions on image value.
File Name	Primary identifier (uniqueID_scanyear_componentpart_imagerole)
Source Information	Describes characteristics of the immediate analog source (original or intermediary) from which the digital image was made (see “Source” in table above)

\*See “Image Quality Assessment” discussion above.

Structural metadata is currently embedded into the file name in a sequential numbering scheme for multi-part items and is reflected in working file directory structures. We anticipate that the file name, which follows the scheme: unique ID\_scan year\_component part\_image role.format extension, can be parsed so that component parts of a digital resource can be logically related together. We also record minimal structural metadata in the header information, such as “front” and “back” for double-sided items or “cover,” “page 1,” “page 2,” “double-page spread” etc. for multi-page items or multi-views. “Component part” is strictly a file sequence number and does not reflect actual page numbers. This metadata is currently recorded as text since the data is not intended to feed into any kind of display or navigation application at the moment.

### Relationships-

Currently there is no utility to record basic relationships among multi-page or multi-part image files beyond documenting relationships in file names. Until a digital asset management system is in place, our practice is to capture as much metadata as possible in the surrounding file structure (names, directories, headers). However, we consider that simple labels or names for file identifiers coupled with more sophisticated metadata describing relationships across files are the preferred way forward to link files together. This metadata would include file identifiers and metadata record identifiers and a codified or typed set of relationships that would help define the associations between image files and between different representations of the same resource. (Relationships between the digital object and the analog source object or the place of the digital object in a larger collection hierarchy would be documented elsewhere in descriptive metadata). Possible relationship types include identification of principal or authoritative version (for production master file); derivation relationships indicating what files come from what files; whether the images were created in the lab or come from another source; structural relationships (for multi-page or –part objects); sibling relationships (images of the same intellectual resource, but perhaps scanned from different source formats). We intend to further refine our work on relationships in the coming months, and start to define metadata that is specific to aggregations of files.

### Batch level metadata-

Currently, data common to all files produced in the Imaging Lab (such as byte order, file format, etc.) is not recorded at the logical level at this time, but we anticipate integrating this kind of information into the construction of a digital asset management system. We are continuing discussions on how to formalize “Lab common

knowledge,” such as details about the hardware and software configurations used to scan and process digital images, target information, and capture and image processing methodologies into our technical metadata specifications.

#### **Permanent and temporary metadata-**

When planning for a digital imaging project, it may not be necessary to save all metadata created and used during the digitization phase of the project. For example, some tracking data may not be needed once all quality control and redo work has been completed. It may not be desirable, or necessary, to bring all metadata into a digital repository. For NARA’s pilot Electronic Access Project, metadata fields that were calculated from other fields, such as square area of a document (used during the pre-scan planning phase to determine scanning resolution and size of access file derivatives), were not saved in the final database since they could be recalculated in the future. Also, it may not be desirable or necessary to provide access to all metadata that is maintained within a system to all users. Most administrative and technical metadata will need to be accessible to administrative users to facilitate managing the digital assets, but does not need to be made available to general users searching the digital collections.

### **III. TECHNICAL OVERVIEW**

#### **Raster Image Characteristics:**

##### **Spatial Resolution-**

Spatial resolution determines the amount of information in a raster image file in terms of the number of picture elements or pixels per unit measurement, but it does not define or guarantee the quality of the information. Spatial resolution defines how finely or widely spaced the individual pixels are from each other. The higher the spatial resolution the more finely spaced and the larger number of pixels overall. The lower the spatial resolution the more widely spaced the pixels and the fewer number of pixels overall.

Spatial resolution is measured as pixels per inch or PPI, also pixels per millimeter or pixels per centimeter are used. Resolution is often referred to as dots per inch or DPI, in common usage the terms PPI and DPI are used interchangeably. Since raster image files are composed of pixels, technically PPI is a more accurate term and is used in this document (one example in support of using the PPI term is that Adobe Photoshop software uses the pixels per inch terminology). DPI is the appropriate term for describing printer resolution (actual dots vs. pixels); however, DPI is used often in scanning and image processing software to refer to spatial resolution and this usage is an understandable convention.

The spatial resolution and the image dimensions determine the total number of pixels in the image; an 8”x10” photograph scanned at 100 ppi produces an image that has 800 pixels by 1000 pixels or a total of 800,000 pixels. The numbers of rows and columns of pixels, or the height and width of the image in pixels as described in the previous sentence, is known as the pixel array. When specifying a desired file size, it is always necessary to provide both the resolution and the image dimensions; ex. 300 ppi at 8”x10” or even 300 ppi at original size.

The image file size, in terms of data storage, is proportional to the spatial resolution (the higher the resolution, the larger the file size for a set document size) and to the size of the document being scanned (the larger the document, the larger the file size for a set spatial resolution). Increasing resolution increases the total number of pixels resulting in a larger image file. Scanning larger documents produces more pixels resulting in larger image files.

Higher spatial resolution provides more pixels, and generally will render more fine detail of the original in the digital image, but not always. The actual rendition of fine detail is more dependent on the spatial frequency response of the scanner or digital camera (see Quantifying Scanner/Digital Camera Performance below), the image processing applied, and the characteristics of the item being scanned. Also, depending on the intended usage of the production master files, there may be a practical limit to how much fine detail is actually needed.

##### **Signal Resolution-**

Bit-depth or signal resolution, sometimes called tonal resolution, defines the maximum number of shades and/or colors in a digital image file, but does not define or guarantee the quality of the information.

In a 1-bit file each pixel is represented by a single binary digit (either a 0 or 1), so the pixel can be either black or white. There are only two possible combinations or  $2^1 = 2$ .

The common standard for grayscale and color images is to use 8-bits (eight binary digits representing each pixel) of data per channel and this provides a maximum of 256 shades per channel ranging from black to white;  $2^8 = 256$  possible combinations of zeroes and ones.

High-bit or 16-bits (16 binary digits representing each pixel) per channel images can have a greater number of shades compared to 8-bit per channel images, a maximum of over 65,000 shades vs. 256 shades;  $2^{16} = 65,536$  possible combinations of zeroes and ones.

Well done 8-bits per channel imaging will meet most needs; with a limited ability for major corrections, transformations, and re-purposing because gross corrections of 8-bit per channel images may cause shades to drop out of the image, creating a posterization effect, due to the limited number of shades.

High-bit images can match the effective shading and density range of photographic originals (assuming the scanner is actually able to capture the information), and, due to the greater shading (compared to 8-bits per channel), may be beneficial when re-purposing images and when working with images that need major or excessive adjustments to the tone distribution and/or color balance. However, at this time, monitors for viewing images and output devices for printing images all render high-bit images at 8-bits per pixel, so there is limited practical benefit to saving high-bit images and no way to verify the accuracy and quality of high-bit images. Also, it is best to do a good job during digitization to ensure accurate tone and color reproduction, rather than relying on post-scan correction of high-bit images. Poorly done high-bit imaging has no benefit.

### **Color Mode-**

Grayscale image files consist of a single channel, commonly either 8-bits (256 levels) or 16-bits (65,536 levels) per pixel with the tonal values ranging from black to white. Color images consist of three or more grayscale channels that represent color and brightness information, common color modes include RGB (red, green, blue), CMYK (cyan, magenta, yellow, black), and LAB (lightness, red-green, blue-yellow). The channels in color files may be either 8-bits (256 levels) or 16-bits (65,536 levels). Display and output devices mathematically combine the numeric values from the multiple channels to form full color pixels, ranging from black to white and to full colors.

RGB represents an additive color process- red, green and blue light are combined to form white light. This is the approach commonly used by computer monitors and televisions, film recorders that image onto photographic film, and digital printers/enlargers that print to photographic paper. RGB files have three color channels: 3 channels x 8-bits = 24-bit color file or 3 channels x 16-bits = 48-bit color. All scanners and digital cameras create RGB files, by sampling for each pixel the amount of light passing through red, green and blue filters that is reflected or transmitted by the item or scene being digitized. Black is represented by combined RGB levels of 0-0-0, and white is represented by combined RGB levels of 255-255-255. This is based on 8-bit imaging and 256 levels from 0 to 255; this convention is used for 16-bit imaging as well, despite the greater number of shades. All neutral colors have equal levels in all three color channels. A pure red color is represented by levels of 255-0-0, pure green by 0-255-0, and pure blue by 0-0-255.

CMYK files are an electronic representation of a subtractive process- cyan (C), magenta (M) and yellow (Y) are combined to form black. CMYK mode files are used for prepress work and include a fourth channel representing black ink (K). The subtractive color approach is used in printing presses (four color printing), color inkjet and laser printers (four color inks, many photo inkjet printers now have more colors), and almost all traditional color photographic processes (red, green and blue sensitive layers that form cyan, magenta and yellow dyes).

LAB color mode is a device independent color space that is matched to human perception- three channels representing lightness (L, equivalent to a grayscale version of the image), red and green information (A), and blue and yellow information (B). LAB mode benefits would include the matching to human perception and they do not require color profiles (see section on color management), disadvantages include the potential loss of information in the conversion from the RGB mode files from scanners and digital cameras, need to have high-bit data, and few applications and file formats support the mode.

Avoid saving files in CMYK mode, CMYK files have a significantly reduced color gamut (see section on color management) and are not suitable for production master image files for digital imaging projects involving holdings/collections in cultural institutions. While theoretically LAB may have benefits, at this time we feel that RGB files produced to the color and tone reproduction described in these guidelines and saved with an Adobe RGB 1998 color profile are the most practical option for production master files and are relatively device independent. We acknowledge our workflow to produce RGB production master files may incur some level of loss of data, however we believe the benefits of using RGB files brought to a common rendering outweigh the minor loss.

## **Digitization Environment:**

Our recommendations and the ISO standards referred to below are based on using CRT monitors. Most LCD monitors we have tested do not compare in quality to the better CRTs in rendering fine detail and smooth gradients. Also, LCD monitors may have artifacts that make it difficult to distinguish image quality problems in the image files, and the appearance of colors and monitor brightness shift with the viewing angle of the LCD panel. This is changing rapidly and the image quality of current high-end LCD monitors is very close to the quality of better CRT monitors. If used, LCD monitors should meet the criteria specified below.

### **Viewing conditions-**

A variety of factors will affect the appearance of images, whether displayed or printed on reflective, transmissive or emissive devices or media. Those factors that can be quantified must be controlled to assure proper representation of an image.

We recommend following the guidance in the following standards-

- o ISO 3664 *Viewing Conditions- For Graphic Technology and Photography*

Provides specifications governing viewing images on reflective and transmissive media, as well as images displayed on a computer monitor without direct comparison to any form of the originals.

- o ISO 12646 *Graphic Technology – Displays for Colour Proofing – Characteristics and Viewing Conditions* (currently a draft international standard or DIS)

Provides specific requirements for monitors and their surrounds for direct comparison of images on a computer monitor with originals (known as softproofing).

**NOTE-** The following are common parameters controlled by users, however refer to the standards for complete requirements and test methods. In particular, ISO 12646 specifies additional hardware requirements for monitors to ensure a reasonable quality level necessary for comparison to hardcopy.

### **Monitor settings, light boxes, and viewing booths-**

We assume the assessment of many digital images will be made in comparison to the originals that have been digitized, therefore ISO 12646 should be followed where it supplements or differs from ISO 3664.

We recommend digital images be viewed on a computer monitor set to 24 bits (millions of colors) or greater, and calibrated to a gamma of 2.2.

ISO 12646 recommends the color temperature of the monitor also be set to 5000K (D50 illuminant) to match the white point of the illumination used for viewing the originals.

Monitor luminance level must be at least 85 cd/m<sup>2</sup>, and should be 120 cd/m<sup>2</sup> or higher.

The computer/monitor desktop should be set to a neutral gray background (avoid images, patterns, and/or strong colors), preferably no more than 10% of the maximum luminance of the screen.

For viewing originals, we recommend using color correct light boxes or viewing booths that have a color temperature of 5000K (D50 illuminant), as specified in ISO 3664.

ISO 3664 provides two luminance levels for viewing originals, ISO 12646 recommends using the lower levels (P2 and T2) when comparing to the image on screen.

The actual illumination level on originals should be adjusted so the perceived brightness of white in the originals matches the brightness of white on the monitor.

### **The room-**

The viewing environment should be painted/decorated a neutral, matte gray with a 60% reflectance or less to minimize flare and perceptual biases.

Monitors should be positioned to avoid reflections and direct illumination on the screen.

ISO 12646 requires the room illumination be less than 32 lux when measured anywhere between the monitor and the observer, and the light a color temperature of approximately 5000K.

#### Practical experience-

In practice, we have found a tolerable range of deviation from the measurements required in the ISO standards. When the ambient room lighting is kept below the limit set in ISO 12646, its color temperature can be lower than 5000K, as long as it is less than the monitor color temperature.

To compensate for environments that may not meet the ISO standards, as well as difficulties comparing analog originals to images on a monitor, the color temperature may need to be set higher than 5000K so that the range of grays from white to black appears neutral when viewed in the actual working environment. The higher color temperature may also be necessary for older monitors to reach an appropriate brightness, as long as neutrals don't appear too blue when compared to neutral hardcopy under the specified illumination.

#### Monitor calibration-

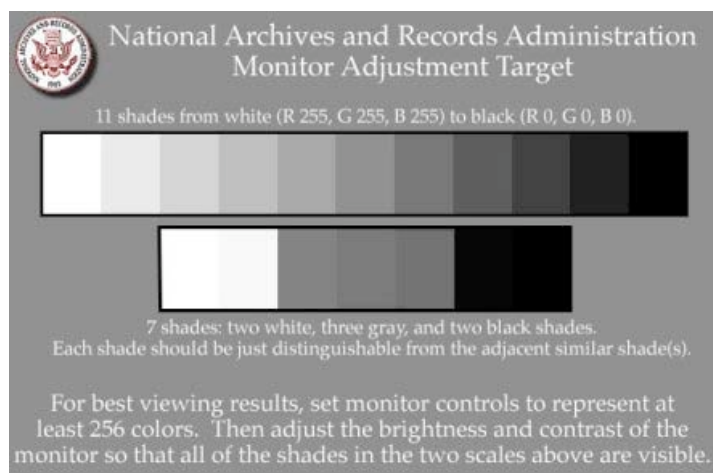
In order to meet and maintain the monitor settings summarized above, we recommend using CRT monitors designed for the graphic arts, photography, or multimedia markets.

A photosensor-based color calibrator and appropriate software (either bundled with the monitor or a third party application) should be used to calibrate the monitor to the aims discussed above. This is to ensure desired color temperature, luminance level, neutral color balance, and linearity of the red, green, and blue representation on the monitor are achieved.

If using an ICC color managed workflow (see section on color management), an ICC profile should be created after monitor calibration for correct rendering of images.

The monitor should be checked regularly and recalibrated when necessary.

Using a photosensor-based monitor calibrator, however, does not always ensure monitors are calibrated well. Ten years of practical experience has shown calibrators and calibration software may not work accurately or consistently. After calibration, it is important to assess the monitor visually, to make sure the monitor is adjusted appropriately. Assess overall contrast, brightness, and color neutrality of the gray desktop. Also, evaluate both color neutrality and detail rendering in white and black areas. This can be done using an image target of neutral patches ranging from black to white and saved in LAB color mode (since LAB does not require an ICC profile and can be viewed independently of the color managed process). In addition, it may be helpful to evaluate sample images or scans of targets – such as the *NARA Monitor Adjustment Target* (shown below) and/or a known image such as a scan of a Kodak grayscale adjusted to the aimpoints (8-8-8 / 105-105-105 / 247-247-247) described below.



When the monitor is adjusted and calibrated appropriately, the *NARA Monitor Adjustment Target* (shown at left) and/or an adjusted image of a Kodak gray scale will look reasonably accurate. Images with ICC color profiles will display accurately within color managed applications, and sRGB profiled images should display reasonably accurately outside color managed applications as well. The *NARA Monitor Adjustment Target* and the gray scale aimpoints are based on an empirical evaluation of a large number of monitors, on both Windows and Macintosh computers, and represent the average of the group. Over the last six years calibrating and adjusting monitors in this manner, we have found the onscreen representation to be very good on a wide variety of monitors and computers.

#### Quantifying Scanner/Digital Camera Performance:

Much effort has gone into quantifying the performance of scanners and digital cameras in an objective manner. The following tests are used to check the capabilities of digitization equipment, and provide information on how to best use the equipment.

Even when digitization equipment is assessed as described below, it is still necessary to have knowledgeable and experienced staff to evaluate images visually. At this time, it is not possible to rely entirely on the objective test measurements to ensure optimum image quality. It is still necessary to have staff with the visual literacy and technical expertise to do a good job with digitization and to perform quality control for digital images. This is true for the digitization of all types of archival records, but very critical for the digitization of photographic images.

Also, these tests are useful when evaluating and comparing scanners and digital cameras prior to purchase. Ask manufacturers and vendors for actual test results, rather than relying on the specifications provided in product literature, some performance claims in product literature are often overstated. If test results are not available, then try to scan test targets during a demonstration and consider having the analyses performed by a contract service.

During digitization projects, tests should be performed on a routine basis to ensure scanners and digital cameras/copy systems are performing optimally. Again, if it is not possible to analyze the tests in-house, then consider having a service perform the analyses on the resulting image files.

The following standards either are available or are in development, these test methods can be used for objective assessment of scanner or digital camera/copy system performance-

Terminology	ISO 12231
Opto-electronic Conversion Function	ISO 14524
Resolution: Still Picture Cameras	ISO 12233
Resolution: Print Scanners	ISO 16067-1
Resolution: Film Scanners	ISO 16067-2
Noise: Still Picture Cameras	ISO 15739
Dynamic Range: Film Scanners	ISO 21550

These standards can be purchased from ISO at <http://www.iso.ch> or from IHS Global at <http://global.ihs.com>. At this time, test methods and standards do not exist for all testing and device combinations. However, many tests are applicable across the range of capture device types and are cited in the existing standards as normative references.

Other test methods may be used to quantify scanner/digital camera performance. We anticipate there will be additional standards and improved test methods developed by the group working on the above standards. Unfortunately, at this time image analysis software is expensive and complex making it difficult to perform all the tests needed to properly quantify scanner/digital camera performance. Also, there is a range of test targets needed for these tests and they can be expensive to purchase.

The following requirements for performance criteria are based on measurements of the variety of actual scanners and digital cameras used in the NWTs Digital Imaging Lab. Where limits are specified, the limits are based on the performance of equipment we consider subjectively acceptable. This subjective acceptability is based on many years combined staff experience in the fields of photography, of photographic reformatting and duplication of a variety of archival records, and of digital imaging and digitization of a variety of archival records.

No digitization equipment or system is perfect, they all have trade-offs in regards to image quality, speed, and cost. The engineering of scanners and digital cameras represents a compromise, and for many markets image quality is sacrificed for higher speed and lower cost of equipment. Many document and book scanners, office scanners (particularly inexpensive ones), and high-speed scanners (all types) may not meet the limits specified, particularly for properties like image noise. Also, many office and document scanners are set at the default to force the paper of the original document to pure white in the image, clipping all the texture and detail in the paper (not desirable for most originals in collections of cultural institutions). These scanners will not be able to meet the desired tone reproduction without recalibration (which may not be possible), without changing the scanner settings (which may not overcome the problem), or without modification of the scanner and/or software (not easily done).

### **Test Frequency and Equipment Variability:**

After equipment installation and familiarization with the hardware and software, an initial performance capability evaluation should be conducted to establish a baseline for each specific digitization device. At a minimum, this benchmark assessment would include for example-

- resolution performance for common sampling rates (e.g. 300, 400, 600, and 800 ppi for reflection scanning)
- OECF and noise characterization for different gamma settings
- lighting and image uniformity

Many scanners can be used both with the software/device drivers provided by the manufacturer and with third-party software/device drivers, characterize the device using the specific software/device drivers to be used for production digitization. Also, performance can change dramatically (and not always for the better) when software/device drivers are updated, characterize the device after every update.

A full suite of tests should be conducted to quantify the performance of digitization systems. Some tests probably only need to be redone on an infrequent basis, while others will need to be done on a routine basis. Depending on the performance consistency of equipment, consider performing tests using production settings on a weekly basis or for each batch of originals, whichever comes first. You may want to perform appropriate tests at the beginning of each batch and at the end of each batch to confirm digitization was consistent for the entire batch.

Scanner/digital camera performance will vary based on actual operational settings. Tests can be used to optimize scanner/camera settings. The performance of individual scanners and digital cameras will vary over time (see test frequency above). Also, the performance of different units of the same model scanner/camera will vary. Test every individual scanner/camera with the specific software/device driver combination(s) used for production. Perform appropriate test(s) any time there is an indication of a problem. Compare these results to past performance through a cumulative database. If large variability is noted from one session to the next for given scanner/camera settings, attempt to rule out operator error first.

## Tests:

### **Opto-electronic conversion function (OECF) – for grayscale and color imaging–**

- Follow ISO 14524.
- Perform OECF analysis for both grayscale and color imaging.
- Perform separate tests and analyses for both reflection and transmission scanning/digitization.
- Run tests at the manufacturer's standard/default settings and at actual production settings.
- Guidance – If these technical guidelines are followed, the actual or final OECF for the production master files is defined by our aimpoints.
- Variability – Limits for acceptable variability are unknown at this time.

### **Dynamic range – for grayscale and color imaging–**

- Follow ISO 14524 and ISO 21550.
- Perform dynamic range analysis for both grayscale and color imaging.
- Perform separate tests and analyses for both reflection and transmission scanning/digitization.
- Guidance – Use of dynamic range analysis –
  - Do not rely on manufacturers' claims regarding the ability of scanners/digital cameras to capture large density ranges as a guide for what originals can be scanned with a particular scanner/camera. Most claims are only based on the sampling bit-depth and not on actual measured performance. Also, the performance of different units of the same model scanner/camera will vary, as well as the performance of individual units will vary over time. Performance will vary based on actual operational settings as well.
  - Do not scan originals that have larger density ranges than the measured dynamic range for a particular scanner/camera and mode (reflection vs. transmission). So, if the measured dynamic range for transmission scanning is 3.2, do not that scanner to scan a color transparency with a density range of greater than 3.2.
- Variability – Limits for acceptable variability are unknown at this time.

### **Spatial frequency response (SFR) – for grayscale and color imaging–**

- Follow ISO 12233, ISO 16067-1, and ISO 16067-2.
- Perform SFR analysis for both grayscale and color imaging.
- Perform separate tests and analyses for both reflection and transmission scanning/digitization.
- Slant edge or sinusoidal targets and corresponding analyses should be used. Generally, do not use character based or line-pair based targets for SFR or resolution analysis.
- For reflection tests – scan targets at a resolution of at least 50% increase from desired resolution (if you plan to save 400 ppi files, then use at least 600 ppi for this test;  $400 \text{ ppi} \times 1.5 = 600 \text{ ppi}$ ), test scans at 100% increase from desired resolution are preferable (if you plan to save 400 ppi, then use at least 800 ppi for testing;  $400 \text{ ppi} \times 2 = 800 \text{ ppi}$ ). Alternative – scan targets at the maximum optical resolution cited by the manufacturer, be aware that depending on the scanner and given the size of the target, this can produce very large test files.
- For transmission tests – scan targets at the maximum resolution cited by the manufacturer, generally it is not necessary to scan at higher interpolated resolutions.
- Guidance – Use of MTF (modulation transfer function) analysis for SFR-
  - Do not rely on manufacturers' claims regarding the resolution of scanners/digital cameras, even optical resolution specifications are not a guarantee the appropriate level of image detail will be captured. Most



claims are over-rated in regards to resolution, and resolution is not the best measure of spatial frequency response (modulation transfer function is the best measurement).

- Evaluation of the MTF curve will provide the maximum resolution a scanner/digital camera system is actually achieving. Use this measured performance (perhaps an appropriate term would be system limited resolution) as a guide. If your scan resolution requirement exceeds the measured performance (system limited resolution), then generally we would not recommend using the scanner for that digitization work.
- The following formula can be used to assist with determining when it is appropriate to use a scanner/digital camera-
  - Scan resolution = desired output resolution x magnification factor.
  - For all items scanned at original size, the magnification factor is one and the scan resolution is the same as your desired output resolution.
  - For images that need to be enlarged, such as scanning a copy transparency or negative and scanning the image to original size, then multiply the desired output resolution by the magnification factor to determine the actual scan resolution – as an example, the desired output resolution is 400 ppi while scanning an image on a 4"x5" copy negative that needs to be enlarged 300% in the scanning software to match original size, the actual scan resolution is  $400 \text{ ppi} \times 3 = 1,200 \text{ ppi}$ .
- Variability – Limits for acceptable variability are unknown at this time.

#### **Noise** – for grayscale and color imaging–

- Follow ISO 15739.
- Perform noise measurement for both grayscale and color imaging.
- Perform separate tests and analyses for both reflection and transmission scanning/digitization.
- Limits –
  - For textual documents and other non-photographic originals with low maximum densities, less than 2.0 visual density-
    - Not to exceed 1.0 counts, out of 255
    - Lower is better
  - For photographs and originals with higher maximum densities, higher than 2.0 visual density
    - Not to exceed 0.7 counts, out of 255
    - Lower is better
- Variability – Limits for acceptable variability are unknown at this time.

#### **Channel registration** – for color imaging–

- Perform color channel registration measurement for color imaging.
- Perform separate tests and analyses for both reflection and transmission scanning/digitization.
- Limits –
  - For all types of originals-
    - Not to exceed 0.5 pixel misregistration.
- Guidance – Lower is better. Good channel registration is particularly important when digitizing textual documents and other line based originals in color; misregistration is very obvious as color halos around monochrome text and lines.
- Variability – Limits for acceptable variability are unknown at this time.

#### **Uniformity – illumination, color, lens coverage, etc.** – for grayscale and color imaging–

- Evaluate uniformity for both grayscale and color imaging.
- Perform separate tests and analyses for both reflection and transmission scanning/digitization.
- The following provides a simple visual method of evaluating brightness and color uniformity, and assists with identifying areas of unevenness and the severity of unevenness-
  - Scan the entire platen or copy board using typical settings for production work. For the reflection test - scan a Kodak photographic gray scale in the middle of the platen/copy board backed with an opaque sheet of white paper that covers the entire platen/copy board; for scanners ensure good contact between the paper and the entire surface of the platen. For the transmission test – scan a Kodak black-and-white film step tablet in the middle of the platen and ensure the rest of the platen is clear. The gray scale and step tablet are included in the scan to ensure auto ranging functions work properly. Scan the gray scale and step tablet, each image should show the scale centered in an entirely white image.
  - For image brightness variability - Evaluate the images using the “Threshold” tool in Adobe Photoshop. Observe the histogram in the Threshold dialog box and look for any clipping of the highlight tones of the image. Move the Threshold slider to higher threshold values and observe when the first portion of the white background turns black and note the numeric level. Continue adjusting the threshold higher until almost the entire image turns black (leaving small areas of white is OK) and note the numeric level (if the highlights have been clipped the background will not turn entirely black even at the highest threshold level of 255 – if

this is the case, use 255 for the numeric value and note the clipping). Subtract the lower threshold value from the higher threshold value. The difference represents the range of brightness levels of the areas of non-uniformity. With a perfectly uniform image, the threshold would turn the entire image black within a range of only 1 to 2 levels. Observe the areas that initially turn black as the threshold is adjusted, if possible avoid these areas of the platen/copy board when scanning. These most frequently occur near the edge of the platen or field of view.

- For color variability - Evaluate the images using the "Levels" tool in Adobe Photoshop. Move the white-point slider lower while holding the Option (Macs)/Alternate (Windows) key to see the clipping point. Observe when the first pixels in the highlight areas turn from black to any color and note the numeric level for the white-point. Continue shifting the white-point lower until almost the entire image turns white (leaving small areas of black or color is OK) and note the numeric level. Subtract the lower white-point value from the higher white-point value, the difference represents the range of color levels of the areas of non-uniformity. With a perfectly uniform image the threshold would turn the entire image white within a range of only 1 to 2 levels.
- Guidance - Make every effort to produce uniform images and to minimize variation. Avoid placing originals being scanned on areas of the platen/copy board that exhibit significant unevenness. Brightness and color variability ranges of 8 levels or less for RGB (3% or less for grayscale) are preferred. Achieving complete field uniformity may be difficult. Some scanners/digital cameras perform a normalization function to compensate for non-uniformity, many do not. It is possible, but very time consuming, to manually compensate for non-uniformity. Conceptually, this uses a low-resolution (50 ppi) grayscale image record of the uniformity performance along with the OECF conditions. In the future effective automated image processing functions may exist to compensate for unevenness in images, this should be done as part of the immediate post-capture image processing workflow.
- Variability - Limits for acceptable variability are unknown at this time.

#### **Dimensional accuracy** - for 1-bit, grayscale, and color imaging-

- For all types of imaging, including 1-bit, grayscale and color.
- Perform separate tests and analyses for reflection and transmission scanning/digitization.
- The following provides a simple approach to assessing dimensional accuracy and consistency-
  - Overall dimensional accuracy-
    - For reflection scanning- scan an Applied Image QA-2 (280mm x 356mm or 11"x14") or IEEE Std 167A-1995 facsimile test target (216mm x 280mm or 8.5"x11") at the resolution to be used for originals. Use target closest in size to the platen size or copy board size.
    - For transmission scanning- Consider scanning thin, clear plastic drafting scales/rulers. If these are too thick, create a ruler in a drafting/drawing application (black lines only on a white background) and print the ruler onto overhead transparency film on a laser printer using the highest possible resolution setting of the printer (1,200 ppi minimum). Compare printed scales to an accurate engineering ruler or tape measure to verify accuracy. Size the scales to match the originals being scanned, shorter and smaller scales for smaller originals. Scan scales at the resolution to be used for originals.
  - Dimensional consistency - for reflection and transmission scanning- scan a measured grid of equally spaced black lines creating 1" squares (2.54 cm) at the resolution that originals are to be scanned. Grids can be produced using a drafting/drawing application and printed on an accurate printer (tabloid or 11"x17" laser printer is preferred, but a good quality inkjet printer can be used and will have to be for larger grids). Reflection grids should be printed on a heavy-weight, dimensionally stable, opaque, white paper. Transmission grids should be printed onto overhead transparency film. Measure grid, both overall and individual squares, with an accurate engineering ruler or tape measure to ensure it is accurate prior to using as a target.
  - Determine the overall dimensional accuracy (as measured when viewed at 200% or 2:1 pixel ratio) for both horizontal and vertical dimensions, and determine dimensional consistency (on the grid each square is 1" or 2.54 cm) across both directions over the full scan area.
- Guidance -
  - Images should be consistent dimensionally in both the horizontal and vertical directions. Overall dimensions of scans should be accurate on the order of 1/10<sup>th</sup> of an inch or 2.45 mm, accuracy of 1/100<sup>th</sup> of an inch or 0.245 mm is preferred. Grids should not vary in square size across both directions of the entire platen or scan area compared to the grid that was scanned.
  - Aerial photography, engineering plans, and other similar originals may require a greater degree of accuracy.
- Variability - Limits for acceptable variability are unknown at this time.

#### **Other artifacts or imaging problems-**

- Note any other problems that are identified while performing all the above assessments.
  - Examples - streaking in blue channel, blur in fast direction.

- Unusual noise or grain patterns that vary spatially across the field.
- One dimensional streaks, and single or clustered pixel dropouts – sometimes these are best detected by visual inspection of individual color channels.
- Color misregistration that changes with position – this is frequently observed along high contrast slant edges.

### Reference Targets:

We recommend including reference targets in each image of originals being scanned, including, at a minimum, a photographic gray scale as a tone and color reference and an accurate dimensional scale. If a target is included in each image, you may want to consider making access derivatives from the production masters that have the reference target(s) cropped out. This will reduce file size for the access files and present an uncluttered appearance to the images presented.

In a high production environment, it may be more efficient to scan targets separately and do it once for each batch of originals. The one target per batch approach is acceptable as long as all settings and operation of the equipment remains consistent for the entire batch and any image processing is applied consistently to all the images. For scanners and digital cameras that have an “auto range” function, the single target per batch approach may not work because the tone and color settings will vary due to the auto range function, depending on the density and color of each original.

All targets should be positioned close to but clearly separated from the originals being scanned. There should be enough separation to allow easy cropping of the image of the original to remove the target(s) if desired, but not so much separation between the original and target(s) that it dramatically increases the file size. If it fits, orient the target(s) along the short dimension of originals, this will produce smaller file sizes compared to having the target(s) along the long dimension (for the same document, a more rectangular shaped image file is smaller than a squarer image). Smaller versions of targets can be created by cutting down the full-size targets. Do not make the tone and color targets so small that it is difficult to see and use the target during scanning (this is particularly important when viewing and working with low resolution image previews within scanning software).

Make sure the illumination on the targets is uniform in comparison to the lighting of the item being scanned (avoid hot spots and/or shadows on the targets). Position targets to avoid reflections.

If the originals are digitized under glass, place the tone and color reference targets under the glass as well. If originals are encapsulated or sleeved with polyester film, place the tone and color reference targets into a polyester sleeve.

For digital copy photography set-ups using digital cameras, when digitizing items that have depth, it is important to make sure all reference targets are on the same level as the image plane – for example, when digitizing a page in a thick book, make sure the reference targets are at the same height/level as the page being scanned.

All types of tone and color targets will probably need to be replaced on a routine basis. As the targets are used they will accumulate dirt, scratches, and other surface marks that reduce their usability. It is best to replace the targets sooner, rather than using old targets for a longer period of time.

### Scale and dimensional references-

Use an accurate dimensional scale as a reference for the size of original documents.

For reflection scanning, scales printed on photographic paper are very practical given the thinness of the paper and the dimensional accuracy that can be achieved during printing. Consider purchasing IEEE Std 167A-1995 facsimile test targets and using the ruler portion of the target along the left-hand edge. Due to the relatively small platen size of most scanners, you may need to trim the ruler off the rest of the target. Different length scales can be created to match the size of various originals. The Kodak Q-13 (8” long) or Q-14 (14” long) color bars have a ruler along the top edge and can be used as a dimensional reference; however, while these are commonly used, they are not very accurate.

For transmission scanning, consider using thin, clear plastic drafting scales/rulers. If these are too thick, create a ruler in a drafting/drawing application (black lines only on a white background) and print the ruler onto overhead transparency film on a laser printer using the highest possible resolution setting of the printer (600 ppi minimum). Compare printed scales to an accurate engineering ruler or tape measure to verify accuracy prior to using as a target. Again, different length scales can be created to match the size of various originals.

### **Targets for tone and color reproduction-**

Reference targets can be used to assist with adjusting scanners and image files to achieve objectively “good images” in terms of tone and color reproduction. This is particularly true with reflection scanning. Copy negatives and copy transparencies should be produced with targets, gray scales and color bars, so they can be printed or scanned to match the original. Unfortunately, scanning original negatives is much more subjective, and this is also the case for copy negatives and copy transparencies that do not contain targets.

### **Reflection scanning-**

We recommend including a Kodak Q-13 (8” long) or Q-14 (14” long) Gray Scale (20 steps, 0.10 density increments, and density range from approximately 0.05 to 1.95) within the area scanned. The Kodak gray scales are made of black-and-white photographic paper and have proven to work well as a reference target, including:

- Good consistency from gray scale to gray scale
- Good color neutrality
- Reasonably high visual density of approximately 1.95
- Provide the ability to quantify color and tone for the full range of values from black-point up to white-point
- The spectral response of the photographic paper has been a reasonable match for a wide variety of originals being scanned on a wide variety of scanners/digital cameras, few problems with metamerism
- The semi-matte surface tends to minimize problems with reflections and is less susceptible to scratching

The Kodak Color Control Patches (commonly referred to as color bars) from the Q-13 and Q-14 should only be used as a supplement to the gray scale, and never as the only target. The color bars are produced on a printing press and are not consistent. Also, the color bars do not provide the ability to assess color and tone reproduction for the full range of values from black-point to white-point.

Other gray scales produced on black-and-white photographic papers could be used. However, many have a glossy surface that will tend to scratch easily and cause more problems with reflections. Also, while being monochrome, some gray scales are not neutral enough to be used as a target.

IT8 color input targets (ex. Kodak Q-60) should not be used as scanning reference targets. IT8 targets are used for producing custom color profiles for scanning specific photographic papers, and therefore are produced on modern color photographic paper. Often, the neutral patches on IT8 targets are not neutral and the spectral response of the color photographic paper is not likely to match the response of most materials being scanned, therefore IT8 targets will not work well as a scanning reference. Also, there is little consistency from one IT8 target to another, even when printed on the same color photo paper.

Consider using a calibrated densitometer or colorimeter to measure the actual visual density or L\*A\*B\* values of each step of the gray scales used as reference targets. Then use a laser printer to print the actual densities and/or L\*A\*B\* values (small font, white text on a gray background) and tape the information above the gray scale so the corresponding values are above each step; for the Kodak gray scales you may need to reprint the identifying numbers and letters for each step. This provides a quick visual reference within the digital image to the actual densities.

### **Transmission scanning – positives–**

Generally, when scanning transmissive positives, like original color transparencies and color slides, a tone and color reference target is usually not necessary. Most scanners are reasonably well calibrated for scanning color transparencies and slides (usually they are not so well calibrated for scanning negatives).

Transparencies and slides have the highest density range of photographic materials routinely scanned. You may need to include within the scan area both a maximum density area of the transparency (typically an unexposed border) and a portion of empty platen to ensure proper auto ranging. Mounted slides can present problems, it is easy to include a portion of the mount as a maximum density area, but since it may not be easy to include a clear area in the scan, you should check highlight levels in the digital image to ensure no detail was clipped.

Ideally, copy transparencies and slides were produced with a gray scale and color bars in the image along with the original. The gray scale in the image should be used for making tone and color adjustments. Caution, carefully evaluate using the gray scales in copy transparencies and slides to make sure that the illumination was even, there are no reflections on the gray scale, and the film was properly processed with no color cross-overs (the highlights and shadows have very different color casts). If any problems exist, you may have problems using the gray scale in the image, as tone and color adjustments will have to be done without relying on the gray scale.

For the best results with transmission scanning, it is necessary to control extraneous light known as flare. It may be necessary to mask the scanner platen or light box down to the just the area of the item being scanned or digitized.

Generally, photographic step tablets on black-and-white film (see discussion on scanning negatives below) are not good as tone and color reference targets for color scanning.

#### **Transmission scanning – negatives–**

We recommend including an uncalibrated Kodak Photographic Step Tablet (21 steps, 0.15 density increments, and density range of approximately 0.05 to 3.05), No. 2 (5" long) or No. 3 (10" long), within the scan area. The standard density range of a step tablet exceeds the density range of most originals that would be scanned, and the scanner can auto-range on the step tablet minimizing loss of detail in the highlight and/or shadow areas of the image.

For production masters, we recommend the brightness range be optimized or matched to the density range of the originals. It may be necessary to have several step tablets, each with a different density range to approximately match the density range of the items being scanned; it is preferable the density range of the step tablet just exceeds the density range of the original. These adjusted step tablets can be produced by cutting off the higher density steps of standard step tablets. If originals have a very short or limited density range compared to the reference targets, this may result in quantization errors or unwanted posterization effects when the brightness range of the digital image is adjusted; this is particularly true for images from low-bit or 8-bit per channel scanners compared to high-bit scanners/cameras.

Ideally, copy negatives were produced with a gray scale and/or color bars in the image along with the original. The gray scale in the image should be used for making tone and/or color adjustments. Caution, carefully evaluate using the gray scales in copy negatives to make sure that the illumination was even, there are no reflections on the gray scale, and for color film the film was properly processed with no color cross-overs (the highlights and shadows have very different color casts). If any problems exist with the quality of the copy negatives, you may have problems using the gray scale in the image, as tone and/or color adjustments will have to be done without relying on the gray scale.

For the best results with transmission scanning, it is necessary to control extraneous light known as flare. It may be necessary to mask the scanner platen or light box down to the just the area of the item being scanned or digitized. This is also true for step tablets being scanned as reference targets. Also, due to the progressive nature of the step tablet, with the densities increasing along the length, it may be desirable to cut the step tablet into shorter sections and mount them out of sequence in an opaque mask; this will minimize flare from the low density areas influencing the high density areas.

Consider using a calibrated densitometer to measure the actual visual and color density of each step of the step tablets used as reference targets. Use a laser printer to print the density values as gray letters against a black background and print onto overhead transparency film, size and space the characters to fit adjacent to the step tablet. Consider mounting the step tablet (or a smaller portion of the step tablet) into an opaque mask with the printed density values aligned with the corresponding steps. This provides a quick visual reference within the digital image to the actual densities.

## **IV. IMAGING WORKFLOW**

### **Adjusting Image Files:**

There is a common misconception that image files saved directly from a scanner or digital camera are pristine or unmolested in terms of the image processing. For almost all image files this is simply untrue. Only "raw" files from scanners or digital cameras are unadjusted, all other digital image files have a range of image processing applied during scanning and prior to saving in order to produce digital images with good image quality.

Because of this misconception, many people argue you should not perform any post-scan or post-capture adjustments on image files because the image quality might be degraded. We disagree. The only time we would recommend saving unadjusted files is if they meet the exact tone and color reproduction, sharpness, and other image quality parameters that you require. Otherwise, we recommend doing minor post-scan adjustment to optimize image quality and bring all images to a common rendition. Adjusting production master files to a common rendition provides significant benefits in terms of being able to batch process and treat all images in the same manner. Well designed and calibrated scanners and digital cameras can produce image files that require little

or no adjustment, however, based on our practical experience, there are very few scanners/cameras that are this well designed and calibrated.

Also, some people suggest it is best to save raw image files, because no “bad” image processing has been applied. This assumes you can do a better job adjusting for the deficiencies of a scanner or digital camera than the manufacturer, and that you have a lot of time to adjust each image. Raw image files will not look good on screen, nor will they match the appearance of originals. Raw image files cannot be used easily; this is true for inaccurate unadjusted files as well. Every image, or batch of images, will have to be evaluated and adjusted individually. This level of effort will be significant, making both raw files and inaccurate unadjusted files inappropriate for production master files.

We believe the benefits of adjusting images to produce the most accurate visual representation of the original outweigh the insignificant data loss (when processed appropriately), and this avoids leaving images in a raw unedited state. If an unadjusted/raw scan is saved, future image processing can be hindered by unavailability of the original for comparison. If more than one version is saved (unadjusted/raw and adjusted), storage costs may be prohibitive for some organizations, and additional metadata elements would be needed. In the future, unadjusted or raw images will need to be processed to be used and to achieve an accurate representation of the originals and this will be difficult to do.

### **Overview:**

We recommend using the scanner/camera controls to produce the most accurate digital images possible for a specific scanner or digital camera. Minor post-scan/post-capture adjustments are acceptable using an appropriate image processing workflow that will not significantly degrade image quality.

We feel the following goals and tools are listed in priority order of importance-

- . 1. Accurate imaging - use scanner controls and reference targets to create grayscale and color images that are:
  - . i. Reasonably accurate in terms of tone and color reproduction, if possible without relying on color management.
  - . ii. Consistent in terms of tone and color reproduction, both image to image consistency and batch to batch consistency.
  - . iii. Reasonably matched to an appropriate use-neutral common rendering for all images.
- . 2. Color management – as a supplement to accurate imaging, use color management to compensate for differences between devices and color spaces:
  - . i. If needed to achieve best accuracy in terms of tone, color, and saturation - use custom profiles for capture devices and convert images to a common wide-gamut color space to be used as the working space for final image adjustment.
  - . ii. Color transformation can be performed at time of digitization or as a post scan/digitization adjustment.
- . 3. Post scan/digitization adjustment - use appropriate image processing tools to:
  - . i. Achieve final color balance and eliminate color biases (color images).
  - . ii. Achieve desired tone distribution (grayscale and color images).
  - . iii. Sharpen images to match appearance of the originals, compensate for variations in originals and the digitization process (grayscale and color images).

The following sections address various types of image adjustments that we feel are often needed and are appropriate. The amount of adjustment needed to bring images to a common rendition will vary depending on the original, on the scanner/digital camera used, and on the image processing applied during digitization (the specific scanner or camera settings).

### **Scanning aimpoints-**

One approach for ensuring accurate tone reproduction (the appropriate distribution of the tones) for digital images is to place selected densities on a gray scale reference target at specific digital levels or aimpoints. Also, for color images it is possible to improve overall color accuracy of the image by neutralizing or eliminating color biases of the same steps of the gray scale used for the tone reproduction aimpoints.

This approach is based on working in a gray-balanced color space, independent of whether it is an ICC color managed workflow or not.

In a digital image, the white point is the lightest spot (highest RGB levels for color files and lowest % black for grayscale files) within the image, the black point is the darkest spot (lowest RGB levels for color files and highest % black for grayscale files), and a mid-point refers to a spot with RGB levels or % black in the middle of the range.

Generally, but not always, the three aimpoints correspond to the white-point, a mid-point, and the black-point within a digital image, and they correspond to the lightest patch, a mid-density patch, and the darkest patch on the reference gray scale within the digital image. This assumes the photographic gray scale has a larger density range than the original being scanned. In addition to adjusting the distribution of the tones, the three aimpoints can be used for a three point neutralization of the image to eliminate color biases in the white-point, a mid-point, and the black-point.

The aimpoints cited in this section are guidelines only. Often it is necessary to vary from the guidelines and use different values to prevent clipping of image detail or to provide accurate tone and color reproduction.



The above images illustrate the importance of controlling the image tones so detail or information is not clipped or lost. The top images have been carefully adjusted using aimpoints and a reference target so all image detail is visible and distinct. The bottom images have been adjusted so the highlight detail in the photograph on the left and the light shades (for the Red channel) in the document on the right have been clipped or rendered at the maximum brightness value (measured as percent gray for the grayscale image and RGB levels for the color image). Clipping or loss of image detail can happen in the shadow detail or dark shades as well, with the pixels being rendered at the lowest brightness value or black. The loss of detail and texture is obvious in the magnified close-up details. Looking at the overall images, the difference in appearance is subtle, but the loss of information is apparent. [photograph on left- President Harry S. Truman, 7/3/1947, NARA – Harry S. Truman Library; document on right- 11<sup>th</sup> Amendment, RG 11 General Records of the United States Government, NARA Old Military and Civil LICON]

Since the aimpoints rely on a photographic gray scale target, they are only applicable when a gray scale is used as a reference. If no gray scale is available (either scanned with the original or in a copy transparency/negative), the Kodak Color Control Patches (color bars) can be used and alternative aimpoints for the color bars are provided. We recommend using a photographic gray scale and not relying on the color bars as the sole target.

Many image processing applications have automatic and manual “place white-point” and “place black-point” controls that adjust the selected areas to be the lightest and darkest portions of the image, and that will neutralize the color in these areas as well as. Also, most have a “neutralize mid-point” control, but usually the tonal adjustment for brightness has to be done separately with a “curves”, “levels”, “tone curve”, etc, control. The better applications will let you set the specific RGB or % black levels for the default operation of the place white-point, place black-point, and neutralize mid-point controls.

Typically, both the brightness placement (for tone reproduction) and color neutralization to adjust the color balance (for color reproduction) should be done in the scanning step and/or as a post-scan adjustment using image processing software. A typical manual workflow in Adobe Photoshop is black-point placement and neutralization



(done as a single step, control set to desired neutral level prior to use), white-point placement and neutralization (done as a single step, control set to desired neutral level prior to use), mid-point neutralization (control set to neutral value prior to use), and a gamma correction to adjust the brightness of the mid point (using levels or curves). For grayscale images the mid-point neutralization step is not needed. The tools in scanner software and other image processing software should allow for a similar approach, the sequence of steps may need to be varied to achieve best results.

The three point tone adjustment and color neutralization approach does not guarantee accurate tone and color reproduction. It works best with most scanners with reflection scanning, but it can be difficult to achieve good tone and color balance when scanning copy negatives/transparencies. It can be very difficult to produce an accurate digital image reproduction from color copy negatives/transparencies that exhibit color cross-over or other defects such as under/over exposure or a strong color cast. The three point neutralization approach will help minimize these biases, but may not eliminate the problems entirely.

If the overall color balance of an image is accurate, using the three point neutralization to adjust the color reproduction may cause the color balance of the shades lighter and darker than the mid-point to shift away from being neutral. For accurate color images that need to have just the tone distribution adjusted, apply levels or curves adjustments to the luminosity information only, otherwise the overall color balance is likely to shift.

When scanning photographic prints it is important to be careful about placing the black point, in some cases the print being scanned will have a higher density than the darkest step of the photographic gray scale. In these cases, you should use a lighter aimpoint for the darkest step of the gray scale so the darkest portion of the image area is placed at the normal aimpoint value (for RGB scans, the shadow area on the print may not be neutral in color and the darkest channel should be placed at the normal aimpoint).

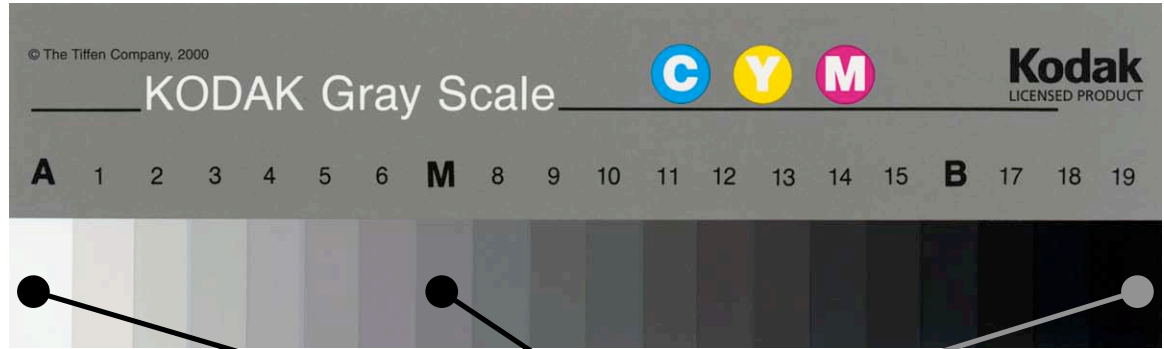
Occasionally, objects being scanned may have a lighter value than the lightest step of the photographic gray scale, usually very bright modern office papers or modern photo papers with a bright-white base. In these cases, you should use a darker aimpoint for the lightest step of the gray scale so the lightest portion of the image area is placed at the normal aimpoint value (for RGB scans, the lightest area of the object being scanned may not be neutral in color and the lightest channel should be placed at the normal aimpoint).

Aimpoints may need to be altered not only for original highlight or shadow values outside the range of the gray scale, but also deficiencies in lighting, especially when scanning photographic intermediates. Excessive flare, reflections, or uneven lighting may need to be accounted for by selecting an alternate value for a patch, or selecting a different patch altogether. At no point should any of the values in any of the color channels of the properly illuminated original fall outside the minimum or maximum values indicated below for scanning without a gray scale.

The aimpoints recommended in the 1998 NARA guidelines have proven to be appropriate for monitor display and for printed output on a variety of printers. The following table provides slightly modified aimpoints to minimize potential problems when printing the image files; the aimpoints described below create a slightly compressed tonal scale compared to the aimpoints in the 1998 guidelines.

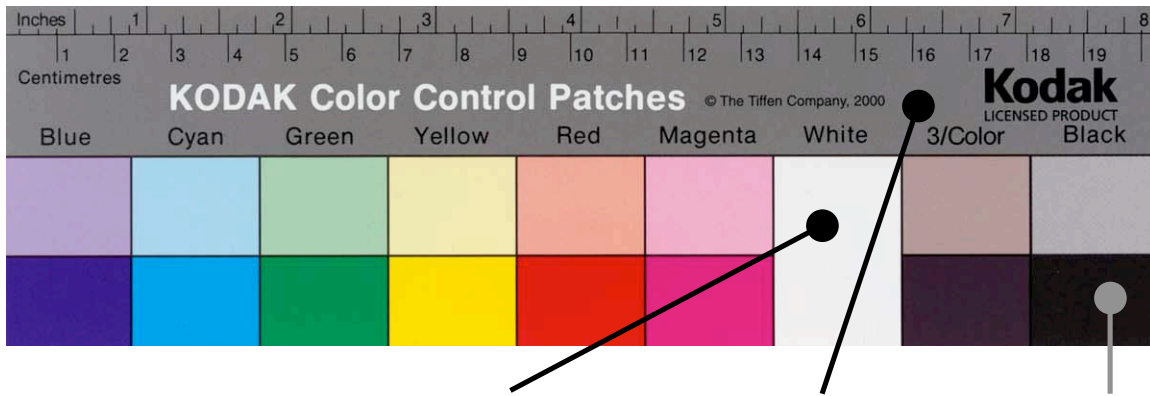
All aimpoint measurements and adjustments should be made using either a 5x5 pixel (25 pixels total) or 3x3 pixel (9 pixels total) sample. Avoid using a point-sample or single pixel measurement.

### Aimpoints for Photographic Gray Scales -



		Neutralized White Point	Neutralized Mid Point*	Neutralized Black Point	Alternative Neutralized Black Point**
Step or Density	Kodak Q-13/Q-14	A	M	19	B
	Visual Density	0.05 to 0.10	0.75 to 0.85	1.95 to 2.05	1.65 to 1.75
Aimpoint	RGB Levels	242-242-242	104-104-104	12-12-12	24-24-24
	% Black	5%	59%	95%	91%
Acceptable Range for Aimpoint	RGB Level	239 to 247	100 to 108	8 to 16	20 to 28
	% Black	3% to 6%	58% to 61%	94% to 97%	89% to 92%
<p>*When using the recommended black point, step 19, the aimpoint for mid point (MP) to be calculated from actual values for white point (WP) and step 19 black point (BP) using the following formula: <math>MP = WP - 0.60(WP - BP)</math></p> <p>**Sometimes there may be problems when trying to use the darkest step of the gray scale - such as excessive flare, reflections, or uneven lighting - and it may be better to use the B-step as the black point. When using the alternative black point, B-step, the aimpoint for mid point (MP) to be calculated from actual values for white point (WP) and step B black point (BP) using the following formula: <math>MP = WP - 0.63(WP - BP)</math></p>					

**Alternative Aimpoints for Kodak Color Control Patches (color bars) -**



		Neutralized White Point	Neutralized Mid Point*	Neutralized Black Point
Color Patch / Area		White	Gray Background	Single Color Black
Aimpoint	RGB Levels	237-237-237	102-102-102	23-23-23
	% Black	7%	60%	91%
Acceptable Range for Aimpoint	RGB Level	233 to 241	98 to 106	19 to 27
	% Black	5% to 9%	58% to 62%	89% to 93%
*Aimpoint for mid point (MP) to be calculated from actual values for white point (WP) and black point (BP) using the following formula: $MP = WP - 0.63(WP - BP)$				

**Aimpoint variability-**

For the three points that have been neutralized and placed at the aimpoint values: no more than +/- 3 RGB level variance from aimpoints and no more than 3 RGB level difference in the individual channels within a patch for RGB scanning and no more than +/- 1% level variance from the aimpoints in % black for grayscale scanning. Again, the image sampler (in Adobe Photoshop or other image processing software) should be set to measure an average of either 5x5 pixels or 3x3 pixels when making these measurements, point sample or single pixel measurements should not be used.

Other steps on the gray scale may, and often will, exhibit a higher degree of variation. Scanner calibration, approaches to scanning / image processing workflow, color management, and variation in the target itself can all influence the variability of the other steps and should be used / configured to minimize the variability for the other steps of the gray scale. Usually the other steps of the gray scale will be relatively consistent for reflection scanning, and significantly less consistent when scanning copy negatives and copy transparencies.

**Minimum and maximum levels-**

The minimum and maximum RGB or % black levels when scanning materials with no reference gray scale or color patches, such as original photographic negatives:

- For RGB scanning the highlight not to go above RGB levels of 247 - 247 - 247 and shadow not to go below RGB levels of 8 - 8 - 8.
- For grayscale scanning the highlight not to go below % black of 3 % and shadow not to go above % black of 97%.

**Color Management Background:**

Digitization is the conversion of analog color and brightness values to discrete numeric values. A number, or set of numbers, designates the color and brightness of each pixel in a raster image. The rendering of these numerical

values, however, is very dependent on the device used for capture, display or printing. Color management provides a context for objective interpretation of these numeric values, and helps to compensate for differences between devices in their ability to render or display these values, within the many limitations inherent in the reproduction of color and tone.

Color management does not guarantee the accuracy of tone and color reproduction. We recommend color management not be used to compensate for poor imaging and/or improper device calibration. As described above, it is most suitable to correct for color rendering differences from device to device.

Every effort should be made to calibrate imaging devices and to adjust scanner/digital camera controls to produce the most accurate images possible in regard to tone and color reproduction (there are techniques for rescuing poorly captured images that make use of profile selection, particularly synthesized profiles, that will not be discussed here. For further information see the writings of Dan Margulis and Michael Kieran). Calibration will not only improve accuracy of capture, but will also ensure the consistency required for color management systems to function by bringing a device to a stable, optimal state. Methods for calibrating hardware vary from device to device, and are beyond the scope of this guidance.

#### **International Color Consortium (ICC) color management system-**

Currently, ICC-based color management is the most widely implemented approach. It consists of four components that are integrated into software (both the operating system and applications):

- PCS (Profile Connection Space)
  - Typically, end users have little direct interaction with the PCS; it is one of two device-independent measuring systems for describing color based on human vision and is usually determined automatically by the source profile. The PCS will not be discussed further.
- Profile
  - A profile defines how the numeric values that describe the pixels in images are to be interpreted, by describing the behavior of a device or the shape and size of a color space.
- Rendering intent
  - Rendering intents determine how out-of-gamut colors will be treated in color space transformations.
- CMM (Color Management Module)
  - The CMM performs the calculations that transform color descriptions between color spaces.

#### **Profiles-**

Profiles are sets of numbers, either a matrix or look up table (LUT), that describe a color space (the continuous spectrum of colors within the gamut, or outer limits, of the colors available to a device) by relating color descriptions specific to that color space to a PCS.

Although files can be saved with any ICC-compliant profile that describes an input device, output device or color space (or with no profile at all), it is best practice to adjust the color and tone of an image to achieve an accurate rendition of the original in a common, well-described, standard color space. This minimizes future effort needed to transform collections of images, as well as streamlines the workflow for repurposing images by promoting consistency. Although there may be working spaces that match more efficiently with the gamut of a particular original, maintaining a single universal working space that covers most input and output devices has additional benefits. Should the profile tag be lost from an image or set of images, the proper profile can be correctly assumed within the digitizing organization, and outside the digitizing organization it can be reasonably found through trial and error testing of the small set of standard workspaces.

Some have argued saving unedited image files in the input device space (profile of the capture device) provides the least compromised data and allows a wide range of processing options in the future, but these files may not be immediately usable and may require individual or small batch transformations. The data available from the scanner has often undergone some amount of adjusting beyond the operator's control, and may not be the best representation of the original. We recommend the creation of production master image files using a standard color space that will be accurate in terms of color and tone reproduction when compared to the original.

The RGB color space for production master files should be gray-balanced, perceptually uniform, and sufficiently large to encompass most input and output devices, while not wasting bits on unnecessary color descriptions. Color spaces that describe neutral gray with equal amounts of red, green and blue are considered to be gray-balanced. A gamma of 2.2 is considered perceptually uniform because it approximates the human visual response to stimuli.

The Adobe RGB 1998 color space profile adequately meets these criteria and is recommended for storing RGB image files. Adobe RGB 1998 has a reasonably large color gamut, sufficient for most purposes when saving files as 24-bit RGB files (low-bit files or 8-bits per channel). Using larger gamut color spaces with low-bit files can cause quantization errors, therefore wide gamut color spaces are more appropriate when saving high-bit or 48-bit RGB files. Gray Gamma 2.2 (available in Adobe products) is recommended for grayscale images.

An ideal workflow would be to scan originals with a calibrated and characterized device, assign the profile of that device to the image file, and convert the file to the chosen workspace (Adobe RGB 1998 for color or Gray Gamma 2.2 for grayscale). Not all hardware and software combinations produce the same color and tonal conversion, and even this workflow will not always produce the best results possible for a particular device or original. Different scanning, image processing and printing applications have their own interpretation of the ICC color management system, and have varying controls that produce different levels of quality. It may be necessary to deviate from the normal, simple color managed workflow to achieve the best results. There are many options possible to achieve the desired results, many of which are not discussed here because they depend on the hardware and software available.

### **Rendering intents-**

When converting images from one color space to another, one of four rendering intents must be designated to indicate how the mismatch of size and shape of source and destination color spaces is to be resolved during color transformations - perceptual, saturation, relative colorimetric, or absolute colorimetric. Of the four, perceptual and relative colorimetric intents are most appropriate for creation of production master files and their derivatives. In general, we have found that perceptual intent works best for photographic images, while relative colorimetric works best for images of text documents and graphic originals. It may be necessary to try both rendering intents to determine which will work best for a specific image or group of images.

When perceptual intent is selected during a color transformation, the visual relationships between colors are maintained in a manner that looks natural, but the appearance of specific colors are not necessarily maintained. As an example, when printing, the software will adjust all colors described by the source color space to fit within a smaller destination space (printing spaces are smaller than most source or working spaces). For images with significant colors that are out of the gamut of the destination space (usually highly saturated colors), perceptual rendering intent often works best.

Relative colorimetric intent attempts to maintain the appearance of all colors that fall within the destination space, and to adjust out-of-gamut colors to close, in-gamut replacements. In contrast to absolute colorimetric, relative colorimetric intent includes a comparison of the white points of the source and destination spaces and shifts all colors accordingly to match the brightness ranges while maintaining the color appearance of all in-gamut colors. This can minimize the loss of detail that may occur with absolute colorimetric in saturated colors if two different colors are mapped to the same location in the destination space. For images that do not contain significant out of gamut colors (such as near-neutral images of historic paper documents), relative colorimetric intent usually works best.

### **Color Management Modules-**

The CMM uses the source and destination profiles and the rendering intent to transform individual color descriptions between color spaces. There are several CMMs from which to select, and each can interact differently with profiles generated from different manufacturers' software packages. Because profiles cannot provide an individual translation between every possible color, the CMM interpolates values using algorithms determined by the CMM manufacturer and each will give varying results.

Profiles can contain a preference for the CMM to be used by default. Some operating systems allow users to designate a CMM to be used for all color transformations that will override the profile tag. Both methods can be superseded by choosing a CMM in the image processing application at the time of conversion. We recommend that you choose a CMM that produces acceptable results for project-specific imaging requirements, and switch only when unexpected transformations occur.

### **Image Processing:**

After capture and transformation into one of the recommended color spaces (referred to as a "working space" at this point in the digitization process), most images require at least some image processing to produce the best digital rendition of the original. The most significant adjustments are color correction, tonal adjustment and sharpening. These processes involve data loss and should be undertaken carefully since they are irreversible once

the file is saved. Images should initially be captured as accurately as possible; image processing should be reserved for optimizing an image, rather than for overcoming poor imaging.

### **Color correction and tonal adjustments-**

Many tools exist within numerous applications for correcting image color and adjusting the tonal scale. The actual techniques of using them are described in many excellent texts entirely devoted to the subject. There are, however, some general principles that should be followed.

- As much as possible, depending on hardware and software available, images should be captured and color corrected in high bit depth.
- Images should be adjusted to render correct highlights and shadows--usually neutral (but not always), of appropriate brightness, and without clipping detail. Also, other neutral colors in the image should not have a color cast (see Aimpoint discussion above).
- Avoid tools with less control that act globally, such as brightness and contrast, and that are more likely to compromise data, such as clipping tones.
- Use tools with more control and numeric feedback, such as levels and curves.
- Despite the desire and all technological efforts to base adjustments on objective measurements, some amount of subjective evaluation may be necessary and will depend upon operator skill and experience.
- Do not rely on "auto correct" features. Most automatic color correction tools are designed to work with color photographic images and the programmers assumed a standard tone and color distribution that is not likely to match your images (this is particularly true for scans of text documents, maps, plans, etc.).

### **Sharpening-**

Digitization utilizes optics in the capture process and the sharpness of different imaging systems varies. Most scans will require some amount of sharpening to reproduce the apparent sharpness of the original. Generally, the higher the spatial resolution, the less sharpening that will be needed. As the spatial resolution reaches a level that renders fine image detail, such as image grain in a photograph, the large features of an image will appear sharp and will not require additional sharpening. Conversely, lower resolution images will almost always need some level of sharpening to match the appearance of the original.

Sharpening tools available from manufacturers use different controls, but all are based on increasing contrast on either side of a defined brightness difference in one or more channels. Sharpening exaggerates the brightness relationship between neighboring pixels with different values, and this process improves the perception of sharpness.

Sharpening of the production master image files should be done conservatively and judiciously; generally it is better to under-sharpen than to over-sharpen. Over-sharpening is irreversible and should be avoided, but it is not objectively measurable. Often over-sharpening will appear as a lighter halo between areas of light and dark.

We recommend using unsharp mask algorithms, rather than other sharpening tools, because they provide the best visual results and usually give greater control over the sharpening parameters. Also-

- Sharpening must be evaluated at an appropriate magnification (1:1 or 100%) and the amount of sharpening is contingent on image pixel dimensions and subject matter.
- Sharpening settings for one image or magnification may be inappropriate for another.
- In order to avoid color artifacts, or fringing, appropriate options or techniques should be used to limit sharpening only to the combined channel brightness.
- The appropriate amount of sharpening will vary depending on the original, the scanner/digital camera used, and the control settings used during digitization.

### **Sample Image Processing Workflow:**

The following provides a general approach to image processing that should help minimize potential image quality defects due to various digital image processing limitations and errors. Depending on the scanner/digital camera, scan/capture software, scanner/digital camera calibration, and image processing software used for post-scan adjustment and/or correction, not all steps may be required and the sequence may need to be modified.

Fewer steps may be used in a high-volume scanning environment to enhance productivity, although this may result in less accurate tone and color reproduction. You can scan a target, adjust controls based on the scan of the target, and then use the same settings for all scans - this approach should work reasonably well for reflection

scanning, but will be much harder to do when scanning copy negatives, copy transparencies, original negatives, and original slides/transparencies.

Consider working in high-bit mode (48-bit RGB or 16-bit grayscale) for as much of the workflow as possible, if the scanner/digital camera and software is high-bit capable and your computer has enough memory and speed to work with the larger files. Conversion to 24-bit RGB or 8-bit grayscale should be done at the end of the sequence.

The post-scan sequence is based on using Adobe Photoshop 7 software.

## WORKFLOW

### Scanning:

Adjust size, scaling, and spatial resolution.

Color correction and tone adjustment-

- o Follow aimpoint guidance - remember there are always exceptions and you may need to deviate from the recommended aimpoints, or to adjust image based on a visual assessment and operator judgement.
- o Recommended – use precision controls in conjunction with color management to achieve the most accurate capture in terms of tone and color reproduction
- o Alternative – if only global controls are available, adjust overall color balance and compress tonal scale to minimize clipping.

Saturation adjustment for color scans.

No sharpening or minimal sharpening (unsharp mask, applied to luminosity preferred).

Color profile conversion (might not be possible at this point, depends on scanner and software)–

- o Convert from scanner space to Adobe RGB 1998 for color images or Gray Gamma 2.2 for grayscale images.
- o Generally, for color image profile conversion – use relative colorimetric rendering intent for near-neutral images (like most text documents) and perceptual rendering intent for photographic and other wide-gamut, high-saturation images.

Check accuracy of scan. You may need to adjust scanner calibration and control settings through trial-and-error testing to achieve best results.

### Post-Scan Adjustment / Correction:

Color profile assignment or conversion (if not done during scanning)–

- o Either assign desired color space or convert from scanner space; use approach that provides best color and tone accuracy.
  - o Adobe RGB 1998 for color images or Gray Gamma 2.2 for grayscale images.
- o Generally, for color image profile conversion – use relative colorimetric rendering intent for near-neutral images (like most text documents) and perceptual rendering intent for photographic and other wide-gamut, high-saturation images.

Color correction-

- o Follow aimpoint guidance - remember there are always exceptions and you may need to deviate from the recommended aimpoints, or to adjust image based on a visual assessment and operator judgment.
- o Recommended - use precision controls (levels recommended, curves alternative) to place and neutralize the black-point, place and neutralize the white-point, and to neutralize mid-point. When color correcting photographic images, levels and curves may both be used.
- o Alternative – try auto-correct function within levels and curves (adjust options, including algorithm, targets, and clipping) and assess results. If auto-correct does a reasonable job, then use manual controls for minor adjustments.
- o Alternative – if only global controls are available, adjust overall color balance.



Tone adjustment, for color files apply correction to luminosity information only-

- o Recommended - use precision controls (levels recommended, curves alternative) to adjust all three aimpoints in iterative process - remember there are always exceptions and you may need to deviate from the recommended aimpoints - or to adjust image based on a visual assessment and operator judgment.
- o Alternative – try auto-correct function within levels and curves (adjust options, including algorithm, targets, and clipping) and assess results. If auto-correct does a reasonable job, then use manual controls for minor adjustments.
- o Alternative – if only global controls are available, adjust contrast and brightness.

Crop and/or deskew.

Check image dimensions and resize.

Convert to 8-bits per channel – either 24-bit RGB or 8-bit grayscale.

Sharpen – Unsharp mask algorithm, applied to approximate appearance of original. For color files, apply unsharp mask to luminosity information only. Version CS (8) of Photoshop has the ability to apply unsharp mask to luminosity in high-bit mode, in this case sharpening should be done prior to the final conversion to 8-bits per channel.

Manual clean up of dust and other artifacts, such as surface marks or dirt on copy negatives or transparencies, introduced during the scanning step. If clean up is done earlier in the image processing workflow prior to sharpening, it is a good idea to check a second time after sharpening since minor flaws will be more obvious after sharpening.

Save file.

Again, the actual image processing workflow will depend on the originals being digitized, the equipment and software being used, the desired image parameters, and the desired productivity. Adjust the image processing workflow for each specific digitization project.

## V. DIGITIZATION SPECIFICATIONS FOR RECORD TYPES

The intent of the following tables is to present recommendations for scanning a variety of original materials in a range of formats and sizes. The tables are broken down into six main categories: textual documents (including graphic illustrations/artworks/originals, maps, plans, and oversized documents); reflective photographic formats (prints); transmissive photographic formats (negatives, slides, transparencies); reflective aerial photographic formats (prints); transmissive aerial photographic formats (negatives, positives); graphic materials (graphic illustrations, drawings, posters); and objects and artifacts.

Because there are far too many formats and document characteristics for comprehensive discussion in these guidelines, the tables below provide scanning recommendations for the most typical or common document types and photographic formats found in most cultural institutions. The table for textual documents is organized around physical characteristics of documents which influence capture decisions. The recommended scanning specifications for text support the production of a scan that can be reproduced as a legible facsimile at the same size as the original (at 1:1, the smallest significant character should be legible). For photographic materials, the tables are organized around a range of formats and sizes that influence capture decisions.

NOTE: We recommend digitizing to the original size of the records following the resolution requirements cited in the tables (i.e. no magnification, unless scanning from various intermediates). Be aware, many Windows applications will read the resolution of image files as 72 ppi by default and the image dimensions will be incorrect.

Workflow requirements, actual usage needs for the image files, and equipment limitations will all be influential factors for decisions regarding how records should be digitized. The recommendations cited in the following section and charts, may not always be appropriate. Again, the intent for these *Technical Guidelines* is to offer a range of options and actual approaches for digitizing records may need to be varied.

### Cleanliness of work area, digitization equipment, and originals-

Keep work area clean. Scanners, platens, and copy boards will have to be cleaned on a routine basis to eliminate the introduction of extraneous dirt and dust to the digital images. Many old documents tend to be dirty and will leave dirt in the work area and on scanning equipment.

See sample handling guidelines, Appendix E, Records Handling for Digitization, for safe and appropriate handling of original records. Photographic originals may need to be carefully dusted with a lint-free, soft-bristle brush to minimize extraneous dust (just as is done in a traditional darkroom or for copy photography).

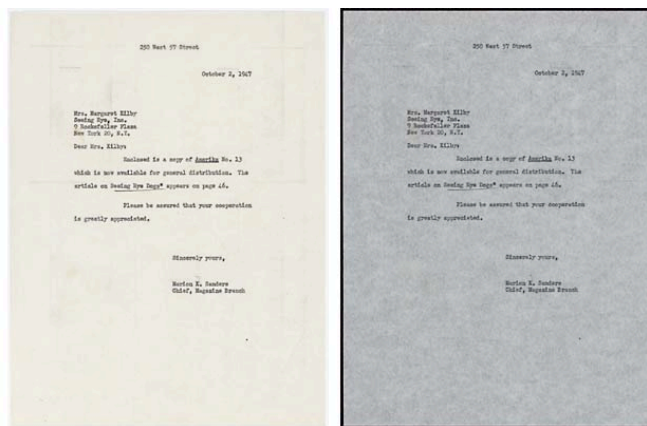
### Cropping-

We recommend the entire document be scanned, no cropping allowed. A small border should be visible around the entire document or photographic image. Careful placement of documents on flatbed scanners may require the originals to be away from platen edge to avoid cropping.

For photographic records - If there is important information on a mount or in the border of a negative, then scan the entire mount and the entire negative including the full border. Otherwise, scan photographs so there is only a small border around just the image area.

### Backing reflection originals-

We recommend backing all originals with a bright white opaque paper (such as a smooth finish cover stock), occasionally, an off-white or cream-colored paper may complement the original document and should be used. For most documents, the bright white backing will provide a lighter shade for scanner auto-ranging and minimize clipping of detail in the paper of the original being scanned. In the graphic arts and photography fields, traditionally items being copied to produce line negatives (somewhat equivalent to 1-bit scanning) have been backed with black to minimize bleed-through from the back. However, this can create very low contrast and/or grayed-out digital images when the paper of the original document is not opaque and when scanning in 8-bit grayscale or 24-bit RGB color. Backing with white paper maximizes the paper brightness of originals and the white border around the originals is much less distracting.



The above document, a carbon copy on thin translucent paper, was scanned and backed with white paper on the left and with black paper on the right. Using a white backing enhances text contrast and allows the scanner to auto-range on the light background, this helps minimize clipping of the tones in the paper of the document. A black backing may minimize bleed-through of text from the back of the page (a technique that may only work well with 1-bit images), but will significantly lower the overall image contrast, gray-out the document image, and show significant paper texture.

### Scanning encapsulated or sleeved originals-

Scanning/digitizing originals that have been encapsulated or sleeved in polyester film can present problems- the visual appearance is changed and the polyester film can cause Newton's rings and other interference patterns.

The polyester film changes the visual appearance of the originals, increasing the visual density. You can compensate for the increase by placing the tone and color reference target (photographic gray scale) into a polyester sleeve (this will increase the visual density of the reference target by the same amount) and scan using the normal aimpoints.

Interference patterns known as Newton's rings are common when two very smooth surfaces are placed in contact, such as placing encapsulated or sleeved documents onto the glass platen of a flatbed scanner; the susceptibility and severity of Newton's rings varies with the glass used, with coatings on the glass, and with humidity in the work area. These patterns will show up in the digital image as multi-colored concentric patterns of various shapes and sizes. Also, we have seen similar interference patterns when digitizing encapsulated documents on a digital copy stand using a scanning camera back, even when there is nothing in contact with the encapsulation. Given the complex nature of these interference patterns, it is not practical to scan and then try to clean-up the image. Some scanners use special glass, known as anti-Newton's ring glass, with a slightly wavy surface to prevent Newton's rings from forming.

To prevent interference patterns, use scanners that have anti-Newton's ring glass and avoid scanning documents in polyester film whenever practical and possible. Some originals may be too fragile to be handled directly and will have to be scanned in the polyester encapsulation or sleeve. One option is to photograph the encapsulated/sleeved document first and then scan the photographic intermediate; generally this approach works well, although we have seen examples of interference patterns on copy transparencies (to a much lesser degree compared to direct digitization).

#### Embossed seals-

Some documents have embossed seals, such as notarized documents, or wax seals that are an intrinsic legal aspect of the documents. Most scanners are designed with lighting to minimize the three dimensional aspects of the original documents being scanned, in order to emphasize the legibility of the text or writing. In most cases, embossed seals or the imprint on a wax seal will not be visible and/or legible in digital images from these scanners, and this raises questions about the authenticity of the digital representation of the documents. Some scanners have a more directed and/or angled lighting configuration that will do a better job reproducing embossed seals. With a few scanners, the operator has the control to turn off one light and scan using lighting from only one direction, this approach will work best for documents with embossed or wax seals. Similarly, when using a digital copy stand, the lighting can be set up for raking light from one direction (make sure the light is still even across the entire document). When working with unidirectional lighting, remember to orient the document so the shadows fall at the bottom of the embossment/seal and of the document.



The close-up on the left shows an embossed seal when scanned on a flatbed scanner with two lights and very even illumination, while the close-up on the right shows the seal from the same document scanned on a flatbed scanner set to use one directional light.

#### Compensating for minor deficiencies-

Scanning at higher than the desired resolution and resampling to the final resolution can minimize certain types of minor imaging deficiencies, such as minor color channel misregistration, minor chromatic aberration, and low to moderate levels of image noise. Conceptually, the idea is to bury the defects in the fine detail of the higher resolution scan, which are then averaged out when the pixels are resampled to a lower resolution. This approach should not be used as a panacea for poorly performing scanners/digital cameras, generally it is better to invest in higher quality digitization equipment. Before using this approach in production, you should run tests to determine there is sufficient improvement in the final image quality to justify the extra time and effort. Generally, we recommend over-scanning at 1.5 times the desired final resolution, as an example- 400 ppi final  $\times$  1.5 = 600 ppi scan resolution.

#### Scanning text-

Guidelines have been established in the digital library community that address the most basic requirements for preservation digitization of text-based materials, this level of reproduction is defined as a "faithful rendering of the underlying source document" as long as the images meet certain criteria. These criteria include completeness, image quality (tonality and color), and the ability to reproduce pages in their correct (original) sequence. As a faithful rendering, a digital master will also support production of a printed page facsimile that is a legible facsimile when produced in the same size as the original (that is 1:1). See the Digital Library Federation's *Benchmark for Faithful Digital Reproductions of Monographs and Serials* at <http://www.diglib.org/standards/bmarkfin.htm> for a detailed discussion.

The Quality Index (QI) measurement was designed for printed text where character height represents the measure of detail. Cornell University has developed a formula for QI based on translating the Quality Index method developed for preservation microfilming standards to the digital world. The QI formula for scanning text relates quality (QI) to character size (h) in mm and resolution (dpi). As in the preservation microfilming standard, the digital QI formula forecasts levels of image quality: barely legible (3.0), marginal (3.6), good (5.0), and excellent (8.0). However, manuscripts and other non-textual material representing distinct edge-based graphics, such as maps, sketches, and engravings, offer no equivalent fixed metric. For many such documents, a better representation of detail would be the width of the finest line, stroke, or marking that must be captured in the digital surrogate. To fully represent such a detail, at least 2 pixels should cover it. (From *Moving Theory into Practice*:

*Digital Imaging for Libraries and Archives*, Anne R. Kenney and Oya Y. Rieger, editors and principal authors. Research Libraries Group, Mountain View, CA: 2000).

Optical character recognition, the process of converting a raster image of text into searchable ASCII data, is not addressed in this document. Digital images should be created to a quality level that will facilitate OCR conversion to a specified accuracy level. This should not, however, compromise the quality of the images to meet the quality index as stated in this document.

#### **Scanning oversized-**

Scanning oversized originals can produce very large file sizes. It is important to evaluate the need for legibility of small significant characters in comparison to the overall file size when determining the appropriate scanning resolution for oversized originals.

#### **Scanning photographs-**

The intent in scanning photographs is to maintain the smallest significant details. Resolution requirements for photographs are often difficult to determine because there is no obvious fixed metric for measuring detail, such as quality index. Additionally, accurate tone and color reproduction in the scan play an equal, if not more, important role in assessing the quality of a scan of a photograph. At this time, we do not feel that there is a valid counterpart for photographic materials to the DLF benchmarks for preservation digitization of text materials.

The recommended scanning specifications for photographs support the capture of an appropriate level of detail from the format, and, in general, support the reproduction, at a minimum, of a high-quality 8"x10" print of the photograph. For photographic formats in particular, it is important to carefully analyze the material prior to scanning, especially if it is not a camera original format. Because every generation of photographic copying involves some quality loss, using intermediates, duplicates, or copies inherently implies some decrease in quality and may also be accompanied by other problems (such as improper orientation, low or high contrast, uneven lighting, etc.).

For original color transparencies, the tonal scale and color balance of the digital image should match the original transparency being scanned to provide accurate representation of the image.

Original photographic negatives are much more difficult to scan compared to positive originals (prints, transparencies, slides, etc.), with positives there is an obvious reference image that can be matched and for negatives there is not. When scanning negatives, for production master files the tonal orientation should be inverted to produce a positive image. The resulting image will need to be adjusted to produce a visually pleasing representation. Digitizing negatives is very analogous to printing negatives in a darkroom and it is very dependent on the photographer's/technician's skill and visual literacy to produce a good image. There are few objective metrics for evaluating the overall representation of digital images produced from negatives.

When working with scans from negatives, care is needed to avoid clipping image detail and to maintain highlight and shadow detail. The actual brightness range and levels for images from negatives are very subject dependent, and images may or may not have a full tonal range.

Often it is better to scan negatives in positive mode (to produce an initial image that appears negative) because frequently scanners are not well calibrated for scanning negatives and detail is clipped in either the highlights and/or the shadows. After scanning, the image can be inverted to produce a positive image. Also, often it is better to scan older black-and-white negatives in color (to produce an initial RGB image) because negatives frequently have staining, discolored film base, retouching, intensification, or other discolorations (both intentional and the result of deterioration) that can be minimized by scanning in color and performing an appropriate conversion to grayscale. Evaluate each color channel individually to determine the channel which minimizes the appearance of any deterioration and optimizes the monochrome image quality, use that channel for the conversion to a grayscale image.

#### **Scanning intermediates-**

Adjust scaling and scan resolution to produce image files that are sized to the original document at the appropriate resolution, or matched to the required QI (legibility of the digital file may be limited due to loss of legibility during the photographic copying process) for text documents.

For copy negatives (B&W and color), if the copy negative has a Kodak gray scale in the image, adjust the scanner settings using the image of the gray scale to meet the above requirements. If there is no gray scale, the scanner software should be used to match the tonal scale of the digital image to the density range of the specific negative

being scanned to provide an image adjusted for monitor representation.

For color copy transparencies and color microfilm, if the color intermediate has a Kodak gray scale in the image, adjust the scanner settings using the image of the gray scale to meet the above requirements. If there is no gray scale, the scanner software should be used to match the tonal scale and color balance of the digital image to the specific transparency being scanned to provide an accurate monitor representation of the image on the transparency.

There are more specific details regarding scanning photographic images from intermediates in the notes following the photo scanning tables.

Generally, for-

- 35mm color copy slides or negatives, a 24-bit RGB digital file of approximately 20 MB would capture the limited information on the film for this small format.
- Approximate maximum scan sizes from color film, 24-bit RGB files (8-bit per channel).<sup>2</sup>

<u>Original Color Film</u>		<u>Duplicate Color Film</u>	
35mm	50 MB	35mm	17 MB
120 square	80 MB	120 square	27 MB
120 6x4.5	60 MB	120 6x4.5	20 MB
120 6x9	90 MB	120 6x9	30 MB
4x5	135 MB	4x5	45 MB
8x10	240 MB	8x10	80 MB

**Scanning microfilm-**

When scanning microfilm, often the desire is to produce images with legible text. Due to photographic limitations of microfilm and the variable quality of older microfilm, it may not be possible to produce what would normally be considered reproduction quality image files. Your scanning approach may vary from the recommendations cited here for textual records and may be more focused on creating digital images with reasonable legibility.

For B&W microfilm, scanner software should be used to match the tonal scale of the digital image to the density range of the specific negative or positive microfilm being scanned. Example: the minimum density of negative microfilm placed at a maximum % black value of 97% and the high density placed at a minimum % black value of 3%.

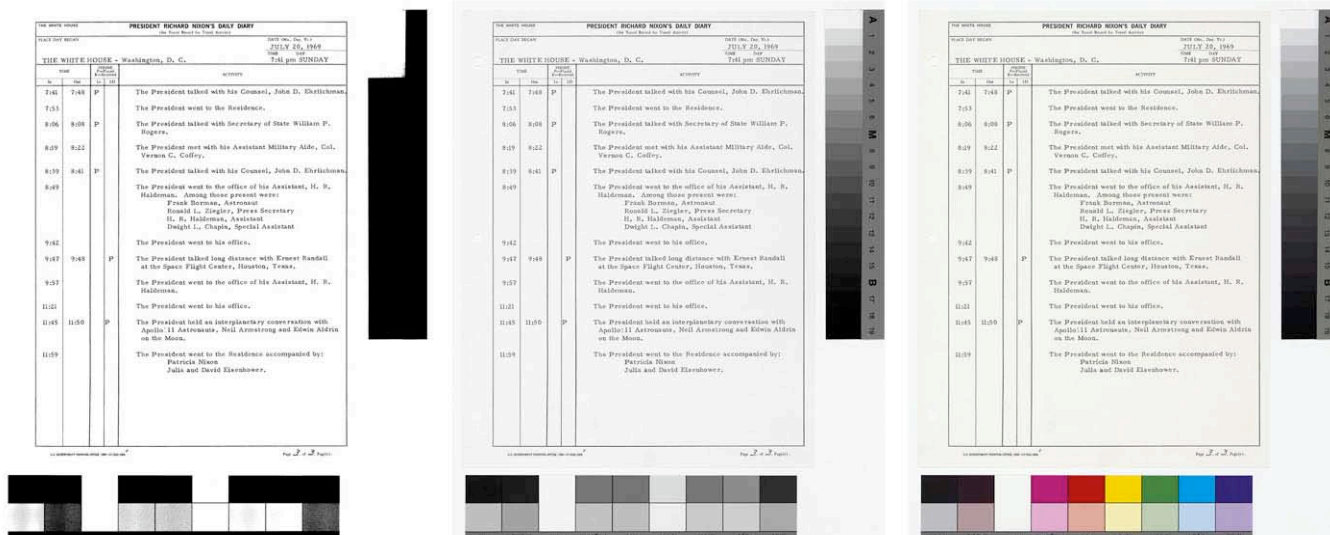
---

<sup>2</sup> From - Digital and Photographic Imaging Services Price Book, Rieger Communications Inc, Gaithersburg, MD, 2001- "In our opinion and experience, you will not achieve better results...than can be obtained from the scan sizes listed....Due to the nature of pixel capture, scanning larger does make a difference if the scan is to be used in very high magnification enlargements. Scan size should not be allowed to fall below 100 DPI at final magnification for quality results in very large prints."

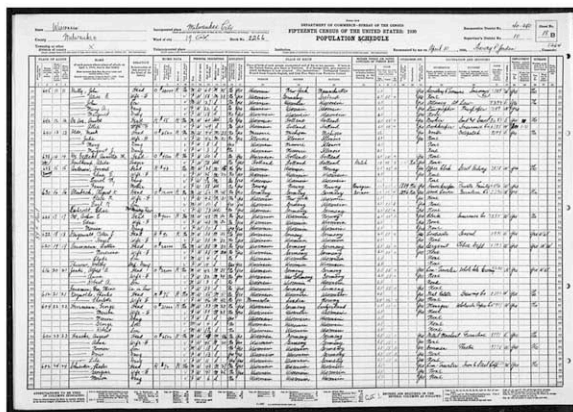
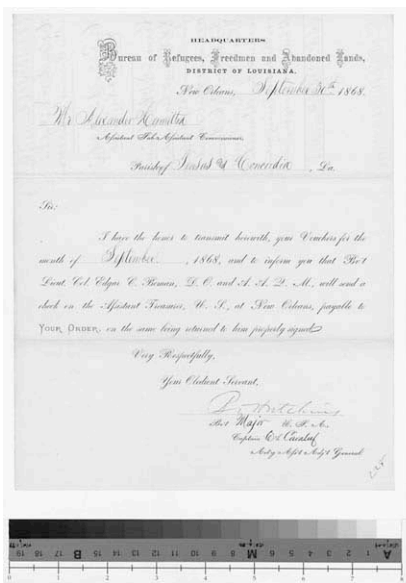


## Illustrations of Record Types:

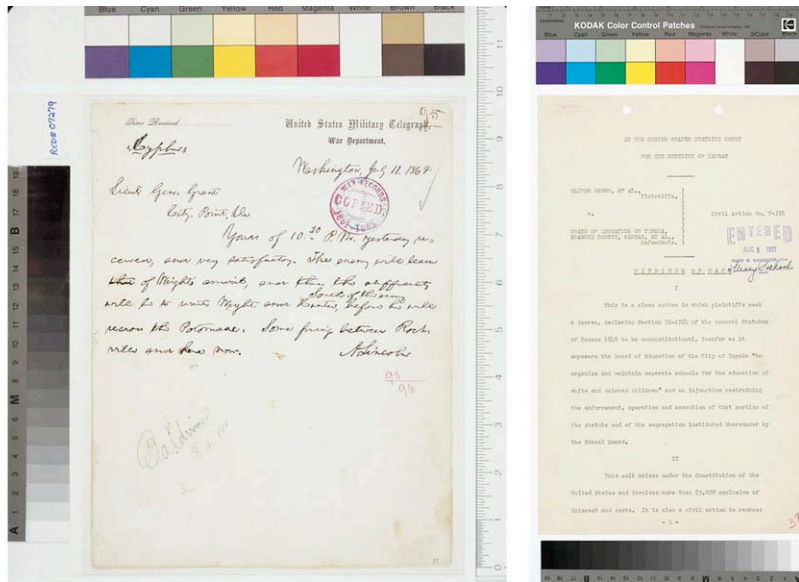
### Textual documents-



Documents with well defined printed type (e.g. typeset, typed, laser printed, etc.), with high inherent contrast between the ink of the text and the paper background, with clean paper (no staining or discoloration), and no low contrast annotations (such as pencil writing) can be digitized either as a 1-bit file (shown on left) with just black and white pixels (no paper texture is rendered), as an 8-bit grayscale file (shown in the center) with gray tones ranging from black to white, or as a 24-bit RGB color image file (shown on right) with a full range of both tones and colors (notice the paper of the original document is an off-white color). [document- President Nixon's Daily Diary, page 3, 7/20/1969, NARA – Presidential Libraries - Nixon Presidential Materials Staff]

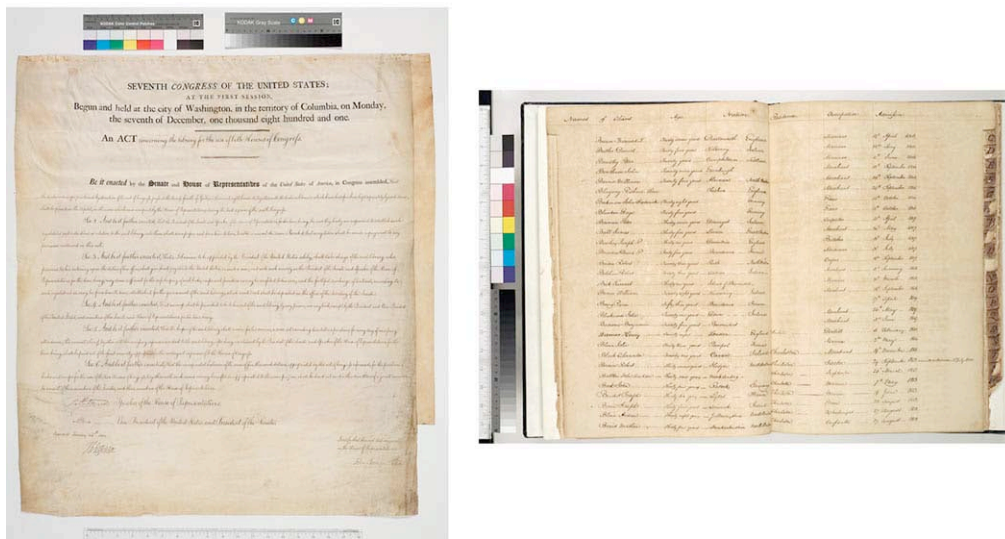


Often grayscale imaging works best for older documents with poor legibility or diffuse characters (e.g. carbon copies, Thermofax/Verifax, etc.), with handwritten annotations or other markings, with low inherent contrast between the text and the paper background, with staining or fading, and with halftone illustrations or photographs included as part of the documents. Many textual documents do not have significant color information and grayscale images will be smaller to store compared to color image files. The document above on the left was scanned directly using a book scanner and the document on the right was scanned from 35mm microfilm using a grayscale microfilm scanner. [document on left- from RG 105, Records of the Bureau of Refugees, Freedmen, and Abandoned Lands, NARA – Old Military and Civil LICON; document on the right- 1930 Census Population Schedule, Milwaukee City, WI, Microfilm Publication T626, Roll 2594, sheet 18B]



For textual documents where color is important to the interpretation of the information or content, or there is a desire to produce the most accurate representation, then scanning in color is the most appropriate approach. The document above on the left was scanned from a 4"x5" color copy transparency using a film scanner and the document on the right was scanned directly on a flatbed scanner. [document on left- Telegram from President Lincoln to General Grant, 07/11/1864, RG 107 Records of the Office of the Secretary of War, NARA – Old Military and Civil LICON; document on the right- Brown v. Board, Findings of Fact, 8/3/1951, RG 21 Records of the District Courts of the United States, NARA – Central Plains Region (Kansas City)]

**Oversized records-**



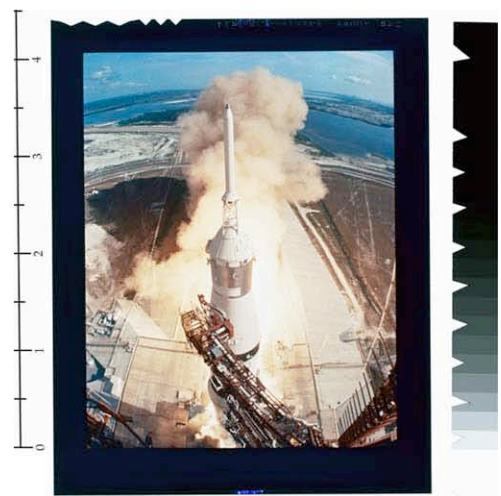
Generally, oversized refers to documents of any type that do not fit easily onto a standard flatbed scanner. The above parchment document on the left and the large book on the right were digitized using a copy stand with a large-format camera and a scanning digital camera back. Books and other bound materials can be difficult to digitize and often require appropriate types of book cradles to prevent damaging the books. [document on the left- Act Concerning the Library for the Use of both Houses of Congress, Seventh Congress of the US, NARA – Center for Legislative Archives; document on the right- Lists of Aliens Admitted to Citizenship 1790-1860, US Circuit and District Courts, District of South Carolina, Charleston, NARA – Southeast Region (Atlanta)]





Maps, architectural plans, engineering plans, etc. are often oversized. Both of the above documents were scanned using a digital copy stand. [document on left- Map of Illinois, 1836, RG 233 Records of the U.S. House of Representatives, NARA – Center for Legislative Archives; document on right- The Mall and Vicinity, Washington, Sheet # 35-23, RG 79 Records of the National Capitol Parks Commission, NARA – Special Media Archives Services Division]

### Photographs-



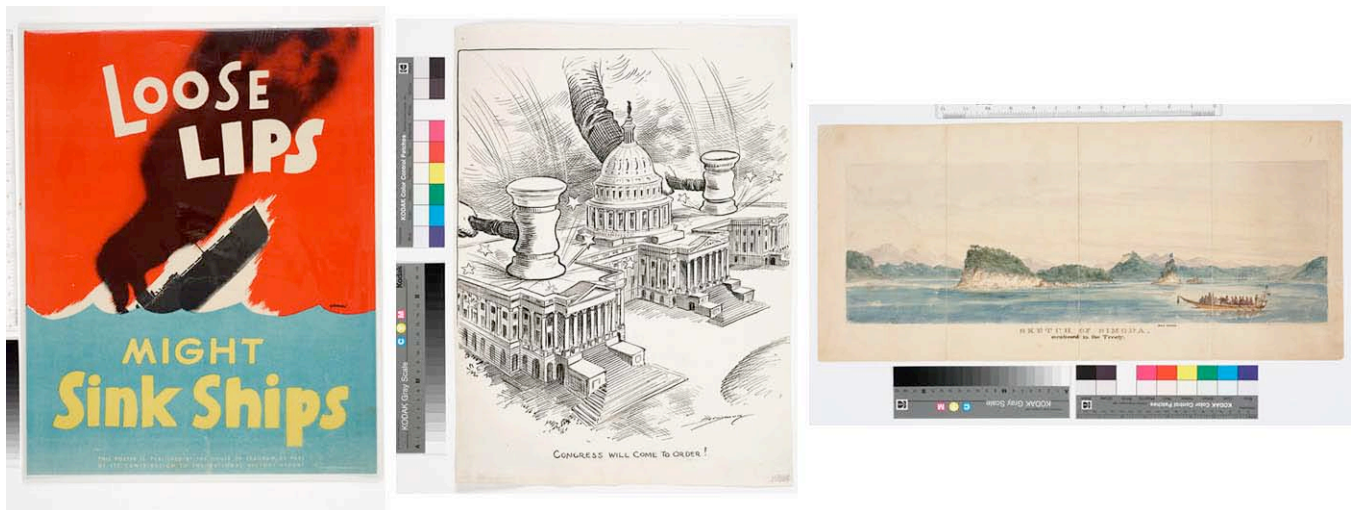
There is a wide variety of photographic originals and different types will require different approaches to digitizing. Above on the left is a scan of a modern preservation-quality film duplicate negative of a Mathew Brady collodion wet-plate negative. Since the modern duplicate negative is in good condition and has a neutral image tone, the negative was scanned as a grayscale image on a flatbed scanner. The photograph in the center is a monochrome print from the 1940s that was scanned in color on a flatbed scanner because the image tone is very warm and there is some staining on the print; many older “black-and-white” prints have image tone and it may be more appropriate to scan these monochrome prints in color. The photo on the right is a 4”x5” duplicate color transparency and was scanned in color using a flatbed scanner. [photograph on left- Gen. Edward O.C. Ord and family, ca. 1860-ca. 1865, 111-B-5091, RG 111 Records of the Office of the Chief Signal Officer, NARA – Special Media Archives Services Division; photograph in center- Alonzo Bankston, electric furnace operator, Wilson Nitrate Plant, Muscle Shoals, Alabama, 1943, RG 142 Records of the Tennessee Valley Authority, NARA – Southeast Region (Atlanta); photograph on right- Launch of the Apollo 11 Mission, 306-AP-A11-5H-69-H-1176, RG 306 Records of the U.S. Information Agency, NARA – Special Media Archives Services Division]

**Aerial photographs-**



Aerial photographs have a lot of fine detail, often require a high degree of enlargement, and may require a higher degree of precision regarding the dimensional accuracy of the scans (compared to textual documents or other types of photographs). The above two grayscale images were produced by scanning film duplicates of the original aerial negatives using a flatbed scanner. The original negative for the image on the left was deteriorated with heavy staining and discoloration, if the original was to be scanned one option would be to scan in color and then to convert to grayscale from an individual color channel that minimizes the appearance of the staining. [photograph on left- Roosevelt Inauguration, 01/1941, ON27740, RG373 Records of the Defense Intelligence Agency, NARA – Special Media Archives Services Division; photograph on the right- New Orleans, LA, French Quarter, 12-15-1952, ON367261/10280628, RG 145 Records of the Farm Service Agency, NARA – Special Media Archives Services Division]

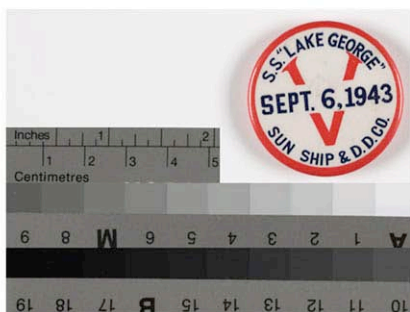
**Graphic illustrations/artwork/originals-**



Some originals have graphic content, and will often have some text information as well. The above examples, a poster on the left, a political cartoon in the center, and an artist’s rendition on the right all fall into this category. The most appropriate equipment to digitize these types of records will vary, and will depend on the size of the originals and their physical condition. [document on left- “Loose Lips Might Sink Ships”, 44-PA-82, RG 44 Records of the Office of Government Reports, NARA – Special Media Archives Services Division; document in center- “Congress Comes to Order” by Clifford K. Berryman, 12/2/1912, Washington Evening Star, D-021, U.S. Senate Collection, NARA – Center for Legislative Archives; document on right- Sketch of Simoda (Treaty of Kanagawa, TS 183 AO, RG 11 General Records of the United States Government, NARA - Old Military and Civil Records LICON]



## Objects and artifacts-



Objects and artifacts can be photographed using either film or a digital camera. If film is used, then the negatives, slides/transparencies, or prints can be digitized. The images on the left were produced using a digital camera and the image on the right was produced by digitizing a 4"x5" color transparency. [objects on top left- Sword and scabbard, Gift from King of Siam, RG 59 General Records of the Department of State, NARA – Civilian Records LICON; object on bottom left- from Buttons Commemorating the Launch of New Ships at Philadelphia Navy Yard, RG 181 Records of the Naval Districts and Shore Establishments, NARA – Mid Atlantic Region (Center City Philadelphia); objects on right- Chap Stick tubes with hidden microphones, RG 460 Records of the Watergate Special Prosecution Force, NARA – Special Access/FOIA LICON]

**Textual documents, graphic illustrations/artwork/originals, maps, plans, and oversized:**

Document Character - Original	Recommended Image Parameters	Alternative Minimum
<p>Clean, high-contrast documents with printed type (e.g. laser printed or typeset)</p>	<p>1-bit bitonal mode or 8-bit grayscale - adjust scan resolution to produce a QI of 8 for smallest significant character</p> <p style="text-align: center;">or</p> <p>1-bit bitonal mode - 600 ppi* for documents with smallest significant character of 1.0 mm or larger</p> <p style="text-align: center;">or</p> <p>8-bit grayscale mode – 400 ppi for documents with smallest significant character of 1.0 mm or larger</p> <p>NOTE: Regardless of approach used, adjust scan resolution to produce a minimum pixel measurement across the long dimension of 6,000 lines for 1-bit files and 4,000 lines for 8-bit files</p> <p style="text-align: center;">*The 600 ppi 1-bit files can be produced via scanning or created / derived from 400 ppi, 8-bit grayscale images.</p>	<p>1-bit bitonal mode - 300 ppi* for documents with smallest significant character of 2.0 mm or larger</p> <p style="text-align: center;">or</p> <p>8-bit grayscale mode - 300 ppi for documents with smallest significant character of 1.5 mm or larger</p> <p style="text-align: right;">*The 300 ppi 1-bit files can be produced via scanning or created / derived from 300 ppi, 8-bit grayscale images.</p>
<p>Documents with poor legibility or diffuse characters (e.g. carbon copies, Thermofax/Verifax, etc.), handwritten annotations or other markings, low inherent contrast, staining, fading, halftone illustrations, or photographs</p>	<p>8-bit grayscale mode - adjust scan resolution to produce a QI of 8 for smallest significant character</p> <p style="text-align: center;">or</p> <p>8-bit grayscale mode - 400 ppi for documents with smallest significant character of 1.0 mm or larger</p> <p>NOTE: Regardless of approach used, adjust scan resolution to produce a minimum pixel measurement across the long dimension of 4,000 lines for 8-bit files</p>	<p>8-bit grayscale mode - 300 ppi for documents with smallest significant character of 1.5 mm or larger</p>
<p>Documents as described for grayscale scanning and/or where color is important to the interpretation of the information or content, or desire to produce the most accurate representation</p>	<p>24-bit color mode - adjust scan resolution to produce a QI of 8 for smallest significant character</p> <p style="text-align: center;">or</p> <p>24-bit RGB mode - 400 ppi for documents with smallest significant character of 1.0 mm or larger</p> <p>NOTE: Regardless of approach used, adjust scan resolution to produce a minimum pixel measurement across the long dimension of 4,000 lines for 24-bit files</p>	<p>24-bit RGB mode - 300 ppi for documents with smallest significant character of 1.5 mm or larger</p>

**Photographs - film / camera originals - black-and-white and color - transmission scanning:**

Format - Original	Recommended Image Parameters	Alternative Minimum
<p><i>Format range:</i></p> <ul style="list-style-type: none"> <li>○ 35 mm and medium-format, up to 4"x5"</li> </ul> <p><i>Size range:</i></p> <ul style="list-style-type: none"> <li>○ Smaller than 20 square inches</li> </ul>	<p><i>Pixel Array:</i></p> <ul style="list-style-type: none"> <li>○ 4000 pixels across long dimension of image area, excluding mounts and borders</li> </ul> <p><i>Resolution:</i></p> <ul style="list-style-type: none"> <li>○ Scan resolution to be calculated from actual image dimensions - approx. 2800 ppi for 35mm originals and ranging down to approx. 800 ppi for originals approaching 4"x5"</li> </ul> <p><i>Dimensions:</i></p> <ul style="list-style-type: none"> <li>○ Sized to match original, no magnification or reduction</li> </ul> <p><i>Bit Depth:</i></p> <ul style="list-style-type: none"> <li>○ 8-bit grayscale mode for black-and-white, can be produced from a 16-bit grayscale file</li> <li>○ 24-bit RGB mode for color and monochrome (e.g. collodion wet-plate negative, pyro developed negatives, stained negatives, etc.), can be produced from a 48-bit RGB file</li> </ul>	<p><i>Pixel Array:</i></p> <ul style="list-style-type: none"> <li>○ 3000 pixels across long dimension for all rectangular formats and sizes</li> <li>○ 2700 pixels by 2700 pixels for square formats regardless of size</li> </ul> <p><i>Resolution:</i></p> <ul style="list-style-type: none"> <li>○ Scan resolution calculated from actual image dimensions – approx. 2100 ppi for 35mm originals and ranging down to the appropriate resolution to produce the desired size file from larger originals, approx. 600 ppi for 4"x5" and 300 ppi for 8"x10" originals</li> </ul>
<p><i>Format range:</i></p> <ul style="list-style-type: none"> <li>○ 4"x5" and up to 8"x10"</li> </ul> <p><i>Size range:</i></p> <ul style="list-style-type: none"> <li>○ Equal to 20 square inches and smaller than 80 square inches</li> </ul>	<p><i>Pixel Array:</i></p> <ul style="list-style-type: none"> <li>○ 6000 pixels across long dimension of image area, excluding mounts and borders</li> </ul> <p><i>Resolution:</i></p> <ul style="list-style-type: none"> <li>○ Scan resolution to be calculated from actual image dimensions – approx. 1200 ppi for 4"x5" originals and ranging down to approx. 600 ppi for 8"x10" originals</li> </ul> <p><i>Dimensions:</i></p> <ul style="list-style-type: none"> <li>○ Sized to match original, no magnification or reduction</li> </ul> <p><i>Bit Depth:</i></p> <ul style="list-style-type: none"> <li>○ 8-bit grayscale mode for black-and-white, can be produced from a 16-bit grayscale file</li> <li>○ 24-bit RGB mode for color and monochrome (e.g. collodion wet-plate negative, pyro developed negatives, stained negatives, etc.), can be produced from a 48-bit RGB file</li> </ul>	<p><i>Dimension:</i></p> <ul style="list-style-type: none"> <li>○ File dimensions set to 10" across long dimension at 300 ppi for rectangular formats and to 9"x9" at 300 ppi for square formats</li> </ul>
<p><i>Format range:</i></p> <ul style="list-style-type: none"> <li>○ 8"x10" and larger</li> </ul> <p><i>Size range:</i></p> <ul style="list-style-type: none"> <li>○ Larger than or equal to 80 square inches</li> </ul>	<p><i>Pixel Array:</i></p> <ul style="list-style-type: none"> <li>○ 8000 pixels across long dimension of image area, excluding mounts and borders</li> </ul> <p><i>Resolution:</i></p> <ul style="list-style-type: none"> <li>○ Scan resolution to be calculated from actual image dimensions – approx. 800 ppi for originals approx. 8"x10" and ranging down to the appropriate resolution to produce the desired size file from larger originals</li> </ul> <p><i>Dimensions:</i></p> <ul style="list-style-type: none"> <li>○ Sized to match original, no magnification or reduction</li> </ul> <p><i>Bit Depth:</i></p> <ul style="list-style-type: none"> <li>○ 8-bit grayscale mode for black-and-white, can be produced from a 16-bit grayscale file</li> <li>○ 24-bit RGB mode for color and monochrome (e.g. collodion wet-plate negative, pyro developed negatives, stained negatives, etc.), can be produced from a 48-bit RGB file</li> </ul>	<p><i>Bit Depth:</i></p> <ul style="list-style-type: none"> <li>○ 8-bit grayscale mode for black-and-white, can be produced from a 16-bit grayscale file</li> <li>○ 24-bit RGB mode for color and monochrome (e.g. collodion wet-plate negative, pyro developed negatives, stained negatives, etc.), can be produced from a 48-bit RGB file</li> </ul>

Duplicate negatives and copy negatives can introduce problems in recommending scanning specifications, particularly if there is no indication of original size. Any reduction or enlargement in size must be taken into account, if possible. In all cases, reproduction to original size is ideal. For copy negatives or

transparencies of prints, use the specifications for that print size. For duplicates (negatives, slides, transparencies), match the original size. However, if original size is not known, the following recommendations are supplied:

- For a copy negative or transparency, scan at a resolution to achieve 4000 pixels across the long dimension.
- For duplicates, follow the scanning recommendations for the size that matches the actual physical dimensions of the duplicate.

For scanning negatives with multiple images on a single negative, see the section on scanning stereographs below. If a ruler has been included in the scan, use it to verify that the image has not been reduced or enlarged before calculating appropriate resolution.

Although many scanning workflows accommodate capturing in 24-bit color, we do not see any benefit at this time to saving the master files of scans produced from modern black-and-white copy negatives and duplicates in RGB. These master scans can be reduced to grayscale in the scanning software or during post-processing editing. However, master scans of camera originals may be kept in RGB, and specifically recommend RGB for any negatives that contain color information as a result of staining, degradation, or intentional color casts.

Scanning Negatives: Often photographic negatives are the most difficult originals to scan. Unlike scanning positives, reflection prints and transparencies/slides, there are no reference images to which to compare scans. Scanning negatives is very much like printing in the darkroom, it is up to the photographer/technician to adjust brightness and contrast to get a good image. Scanning negatives is a very subjective process that is very dependent on the skill of the photographer/technician. Also, most scanners are not as well calibrated for scanning negatives compared to scanning positives.

Often to minimize loss of detail, it is necessary to scan negatives as positives (the image on screen is negative), to invert the images in Photoshop, and then to adjust the images.

If black-and-white negatives are stained or discolored, we recommend making color RGB scans of the negatives and using the channel which minimizes the appearance of the staining/discoloration when viewed as a positive. The image can then be converted to a grayscale image.



On the left is an image of a historic black-and-white film negative that was scanned in color with a positive tonal orientation (the digital image appears the same as the original negative), this represents a reasonably accurate rendition of the original negative. The middle grayscale image shows a direct inversion of the tones, and as shown here, often a direct inversion of a scan of a negative will not produce a well-rendered photographic image. The image on the right illustrates an adjusted version where the brightness and contrast of the image has been optimized (using "Curves" and "Levels" in Adobe Photoshop software) to produce a reasonable representation of the photographic image, these adjustments are very similar to how a photographer prints a negative in the darkroom. [photograph- NRCA-142-INFO01-3169D, RG 142 Records of the TVA, NARA – Southeast Region (Atlanta)]

**Photographs - prints - black-and-white, monochrome, and color - reflection scanning:**

Format - Original	Recommended Image Parameters	Alternative Minimum
<p><i>Format range:</i></p> <ul style="list-style-type: none"> <li>o 8"x10" or smaller</li> </ul> <p><i>Size range:</i></p> <ul style="list-style-type: none"> <li>o Smaller than or equal to 80 square inches</li> </ul>	<p><i>Pixel Array:</i></p> <ul style="list-style-type: none"> <li>o 4000 pixels across long dimension of image area, excluding mounts and borders</li> </ul> <p><i>Resolution:</i></p> <ul style="list-style-type: none"> <li>o Scan resolution to be calculated from actual image dimensions – approx. 400 ppi for 8"x10" originals and ranging up to the appropriate resolution to produce the desired size file from smaller originals, approx. 570 ppi for 5"x7" and 800 ppi for 4"x5" or 3.5"x5" originals</li> </ul> <p><i>Dimensions:</i></p> <ul style="list-style-type: none"> <li>o Sized to match the original, no magnification or reduction</li> </ul> <p><i>Bit Depth:</i></p> <ul style="list-style-type: none"> <li>o 8-bit grayscale mode for black-and-white, can be produced from a 16-bit grayscale file</li> <li>o 24-bit RGB mode for color and monochrome (e.g. albumen prints or other historic print processes), can be produced from a 48-bit RGB file</li> </ul>	<p><i>Pixel Array:</i></p> <ul style="list-style-type: none"> <li>o 3000 pixels across long dimension for all rectangular formats and sizes</li> <li>o 2700 pixels by 2700 pixels for square formats regardless of size</li> </ul> <p><i>Resolution:</i></p> <ul style="list-style-type: none"> <li>o Scan resolution calculated from actual image dimensions – approx. 2100 ppi for 35mm originals and ranging down to the appropriate resolution to produce the desired size file from larger originals, approx. 600 ppi for 4"x5" and 300 ppi for 8"x10" originals</li> </ul>
<p><i>Format range:</i></p> <ul style="list-style-type: none"> <li>o Larger than 8"x10" and up to 11"x14"</li> </ul> <p><i>Size range:</i></p> <ul style="list-style-type: none"> <li>o Larger than 80 square inches and smaller than 154 square inches</li> </ul>	<p><i>Pixel Array:</i></p> <ul style="list-style-type: none"> <li>o 6000 pixels across long dimension of image area, excluding mounts and borders</li> </ul> <p><i>Resolution:</i></p> <ul style="list-style-type: none"> <li>o Scan resolution to be calculated from actual image dimensions – approx. 600 ppi for originals approx. 8"x10" and ranging down to approx. 430 ppi for 11"x14" originals</li> </ul> <p><i>Dimensions:</i></p> <ul style="list-style-type: none"> <li>o Sized to match the original, no magnification or reduction</li> </ul> <p><i>Bit Depth:</i></p> <ul style="list-style-type: none"> <li>o 8-bit grayscale mode for black-and-white, can be produced from a 16-bit grayscale file</li> <li>o 24-bit RGB mode for color and monochrome (e.g. albumen prints or other historic print processes), can be produced from a 48-bit RGB file</li> </ul>	<p><i>Dimension:</i></p> <ul style="list-style-type: none"> <li>o File dimensions set to 10" across long dimension at 300 ppi for rectangular formats and to 9"x9" at 300 ppi for square formats</li> </ul>
<p><i>Format range:</i></p> <ul style="list-style-type: none"> <li>o Larger than 11"x14"</li> </ul> <p><i>Size range:</i></p> <ul style="list-style-type: none"> <li>o Equal to or larger than 154 square inches</li> </ul>	<p><i>Pixel Array:</i></p> <ul style="list-style-type: none"> <li>o 8000 pixels across long dimension of image area, excluding mounts and borders</li> </ul> <p><i>Resolution:</i></p> <ul style="list-style-type: none"> <li>o Scan resolution to be calculated from actual image dimensions – approx. 570 ppi for originals approx. 11"x14" and ranging down to the appropriate resolution to produce the desired size file from larger originals</li> </ul> <p><i>Dimensions:</i></p> <ul style="list-style-type: none"> <li>o Sized to match the original, no magnification or reduction</li> </ul> <p><i>Bit Depth:</i></p> <ul style="list-style-type: none"> <li>o 8-bit grayscale mode for black-and-white, can be produced from a 16-bit grayscale file</li> <li>o 24-bit RGB mode for color and monochrome (e.g. albumen prints or other historic print processes), can be produced from a 48-bit RGB file</li> </ul>	<p><i>Bit Depth:</i></p> <ul style="list-style-type: none"> <li>o 8-bit grayscale mode for black-and-white, can be produced from a 16-bit grayscale file</li> <li>o 24-bit RGB mode for color and monochrome (e.g. collodion wet-plate negative, pyro developed negatives, stained negatives, etc.), can be produced from a 48-bit RGB file</li> </ul>



For stereograph images and other multiple image prints, modified recommended scanning specifications are to scan to original size (length of both photos and mount) and add 2000 pixels to the long dimension, in the event that only one of the photographs is requested for high-quality reproduction. For example, if the stereograph is 8" on the long dimension, a resolution of 500 ppi would be required to achieve 4000 pixels across the long dimension for that size format; in this case, adding 2000 pixels to the long dimension would require that the stereograph be scanned at 750 ppi to achieve the desired 6000 pixels across the long dimension.

For photographic prints, size measurements for determining appropriate resolution are based on the size of the image area only, excluding any borders, frames, or mounts. However, in order to show that the entire record has been captured, it is good practice to capture the border area in the master scan file. In cases where a small image is mounted on a large board (particularly where large file sizes may be an issue), it may be desirable to scan the image area only at the appropriate resolution for its size, and then scan the entire mount at a resolution that achieves 4000 pixels across the long dimension.

**Aerial - transmission scanning:**

Format - Original	Recommended Image Parameters*	Alternative Minimum
<p><i>Format range:</i></p> <ul style="list-style-type: none"> <li>○ 70mm wide and medium format roll film</li> </ul> <p><i>Size range:</i></p> <ul style="list-style-type: none"> <li>○ Smaller than 10 square inches</li> </ul>	<p><i>Pixel Array:</i></p> <ul style="list-style-type: none"> <li>○ 6000 pixels across long dimension of image area, excluding borders</li> </ul> <p><i>Resolution:</i></p> <ul style="list-style-type: none"> <li>○ Scan resolution to be calculated from actual image dimensions – approx. 2700 ppi for 70mm originals and ranging down to the appropriate resolution to produce the desired size file from larger originals</li> </ul> <p><i>Dimensions:</i></p> <ul style="list-style-type: none"> <li>○ Sized to match original, no magnification or reduction</li> </ul> <p><i>Bit Depth:</i></p> <ul style="list-style-type: none"> <li>○ 8-bit grayscale mode for black-and-white, can be produced from a 16-bit grayscale file</li> <li>○ 24-bit RGB mode for color and monochrome (stained negatives), can be produced from a 48-bit RGB file</li> </ul>	<p><i>Pixel Array:</i></p> <ul style="list-style-type: none"> <li>○ 4000 pixels across long dimension of image area</li> </ul> <p><i>Resolution:</i></p> <ul style="list-style-type: none"> <li>○ Scan resolution calculated from actual image dimensions – approx. 1800 ppi for 6cm x 6cm originals and ranging down to the appropriate resolution to produce the desired size file from larger originals, approx. 800 ppi for 4”x5” and 400 ppi for 8”x10” originals</li> </ul>
<p><i>Format range:</i></p> <ul style="list-style-type: none"> <li>○ 127mm wide roll film, 4”x5” and up to 5”x7” sheet film</li> </ul> <p><i>Size range:</i></p> <ul style="list-style-type: none"> <li>○ Equal to 10 square inches and up to 35 square inches</li> </ul>	<p><i>Pixel Array:</i></p> <ul style="list-style-type: none"> <li>○ 8000 pixels across long dimension of image area, excluding borders</li> </ul> <p><i>Resolution:</i></p> <ul style="list-style-type: none"> <li>○ Scan resolution to be calculated from actual image dimensions – approx. 1600 ppi for 4”x5” originals and ranging down to approx. 1100 ppi for 5”x7” originals</li> </ul> <p><i>Dimensions:</i></p> <ul style="list-style-type: none"> <li>○ Sized to match original, no magnification or reduction</li> </ul> <p><i>Bit Depth:</i></p> <ul style="list-style-type: none"> <li>○ 8-bit grayscale mode for black-and-white, can be produced from a 16-bit grayscale file</li> <li>○ 24-bit RGB mode for color and monochrome (stained negatives), can be produced from a 48-bit RGB file</li> </ul>	<p><i>Dimension:</i></p> <ul style="list-style-type: none"> <li>○ File dimensions set to 10” across long dimension at 400 ppi for all formats</li> </ul>
<p><i>Format range:</i></p> <ul style="list-style-type: none"> <li>○ Larger than 127mm wide roll film and larger than 5”x7” sheet film</li> </ul> <p><i>Size range:</i></p> <ul style="list-style-type: none"> <li>○ Larger than 35 square inches</li> </ul>	<p><i>Pixel Array:</i></p> <ul style="list-style-type: none"> <li>○ 10000 pixels across long dimension of image area, excluding borders</li> </ul> <p><i>Resolution:</i></p> <ul style="list-style-type: none"> <li>○ Scan resolution to be calculated from actual image dimensions – approx. 2000 ppi for 5”x5” originals and ranging down to the appropriate resolution to produce the desired size file from larger originals</li> </ul> <p><i>Dimensions:</i></p> <ul style="list-style-type: none"> <li>○ Sized to match original, no magnification or reduction</li> </ul> <p><i>Bit Depth:</i></p> <ul style="list-style-type: none"> <li>○ 8-bit grayscale mode for black-and-white, can be produced from a 16-bit grayscale file</li> <li>○ 24-bit RGB mode for color and monochrome (e.g. stained negatives), can be produced from a 48-bit RGB file</li> </ul>	<p><i>Bit Depth:</i></p> <ul style="list-style-type: none"> <li>○ 8-bit grayscale mode for black-and-white, can be produced from a 16-bit grayscale file</li> <li>○ 24-bit RGB mode for color and monochrome (e.g. stained negatives), can be produced from a 48-bit RGB file</li> </ul>

\*If scans of aerial photography will be used for oversized reproduction, follow the scanning recommendations for the next largest format (e.g., if your original is 70mm wide, follow the specifications for 127mm wide roll film to achieve 8000 pixels across the long dimensions.

**Aerial - reflection scanning:**

Format - Original	Recommended Image Parameters*	Alternative Minimum
<p><i>Format range:</i></p> <ul style="list-style-type: none"> <li>○ Smaller than 8"x10"</li> </ul> <p><i>Size range:</i></p> <ul style="list-style-type: none"> <li>○ Smaller than 80 square inches</li> </ul>	<p><i>Pixel Array:</i></p> <ul style="list-style-type: none"> <li>○ 4000 pixels across long dimension of image area, excluding mounts and borders</li> </ul> <p><i>Resolution:</i></p> <ul style="list-style-type: none"> <li>○ Scan resolution to be calculated from actual image dimensions – approx. 400 ppi for originals approx. 8"x10" and ranging up to the appropriate resolution to produce the desired size file from smaller originals, approx. 570 ppi for 5"x7" and 800 ppi for 4"x5" originals</li> </ul> <p><i>Dimensions:</i></p> <ul style="list-style-type: none"> <li>○ Sized to match the original, no magnification or reduction</li> </ul> <p><i>Bit Depth:</i></p> <ul style="list-style-type: none"> <li>○ 8-bit grayscale mode for black-and-white, can be produced from a 16-bit grayscale file</li> <li>○ 24-bit RGB mode for color and monochrome (e.g. discolored prints), can be produced from a 48-bit RGB file</li> </ul>	<p><i>Pixel Array:</i></p> <ul style="list-style-type: none"> <li>○ 3000 pixels across long dimension of image area</li> </ul> <p><i>Resolution:</i></p> <ul style="list-style-type: none"> <li>○ Scan resolution to be calculated from actual image dimensions – approx. 300 ppi for 8"x10" originals and ranging up to the appropriate resolution to produce the desired size file from smaller originals, approx. 570 ppi for 5"x7" and 800 ppi for 4"x5" or 3.5"x5" originals</li> </ul> <p><i>Dimensions:</i></p> <ul style="list-style-type: none"> <li>○ Sized to match the original, no magnification or reduction</li> </ul> <p><i>Bit Depth:</i></p> <ul style="list-style-type: none"> <li>○ 8-bit grayscale mode for black-and-white, can be produced from a 16-bit grayscale file</li> <li>○ 24-bit RGB mode for color and monochrome (e.g. discolored prints), can be produced from a 48-bit RGB file</li> </ul>
<p><i>Format range:</i></p> <ul style="list-style-type: none"> <li>○ 8"x10" and up to 11"x14"</li> </ul> <p><i>Size range:</i></p> <ul style="list-style-type: none"> <li>○ Equal to 80 square inches and up to 154 square inches</li> </ul>	<p><i>Pixel Array:</i></p> <ul style="list-style-type: none"> <li>○ 6000 pixels across long dimension of image area, excluding mounts and borders</li> </ul> <p><i>Resolution:</i></p> <ul style="list-style-type: none"> <li>○ Scan resolution to be calculated from actual image dimensions – approx. 600 ppi for 8"x10" originals and ranging down to approx. 430 ppi for 11"x14" originals</li> </ul> <p><i>Dimensions:</i></p> <ul style="list-style-type: none"> <li>○ Sized to match the original, no magnification or reduction</li> </ul> <p><i>Bit Depth:</i></p> <ul style="list-style-type: none"> <li>○ 8-bit grayscale mode for black-and-white, can be produced from a 16-bit grayscale file</li> <li>○ 24-bit RGB mode for color and monochrome (e.g. discolored prints), can be produced from a 48-bit RGB file</li> </ul>	<p><i>Resolution:</i></p> <ul style="list-style-type: none"> <li>○ Scan resolution to be calculated from actual image dimensions – approx. 300 ppi for 8"x10" originals and ranging up to the appropriate resolution to produce the desired size file from smaller originals, approx. 570 ppi for 5"x7" and 800 ppi for 4"x5" or 3.5"x5" originals</li> </ul> <p><i>Dimensions:</i></p> <ul style="list-style-type: none"> <li>○ Sized to match the original, no magnification or reduction</li> </ul> <p><i>Bit Depth:</i></p> <ul style="list-style-type: none"> <li>○ 8-bit grayscale mode for black-and-white, can be produced from a 16-bit grayscale file</li> <li>○ 24-bit RGB mode for color and monochrome (e.g. discolored prints), can be produced from a 48-bit RGB file</li> </ul>
<p><i>Format range:</i></p> <ul style="list-style-type: none"> <li>○ Larger than 11"x14"</li> </ul> <p><i>Size range:</i></p> <ul style="list-style-type: none"> <li>○ Larger than 154 square inches</li> </ul>	<p><i>Pixel Array:</i></p> <ul style="list-style-type: none"> <li>○ 8000 pixels across long dimension of image area, excluding mounts and borders</li> </ul> <p><i>Resolution:</i></p> <ul style="list-style-type: none"> <li>○ Scan resolution to be calculated from actual image dimensions – approx. 570 ppi for 11"x14" originals and ranging down to the appropriate resolution to produce the desired size file from larger originals</li> </ul> <p><i>Dimensions:</i></p> <ul style="list-style-type: none"> <li>○ Sized to match the original, no magnification or reduction</li> </ul> <p><i>Bit Depth:</i></p> <ul style="list-style-type: none"> <li>○ 8-bit grayscale mode for black-and-white, can be produced from a 16-bit grayscale file</li> <li>○ 24-bit RGB mode for color and monochrome (e.g. discolored prints), can be produced from a 48-bit RGB file</li> </ul>	<p><i>Resolution:</i></p> <ul style="list-style-type: none"> <li>○ Scan resolution to be calculated from actual image dimensions – approx. 570 ppi for 11"x14" originals and ranging down to the appropriate resolution to produce the desired size file from larger originals</li> </ul> <p><i>Dimensions:</i></p> <ul style="list-style-type: none"> <li>○ Sized to match the original, no magnification or reduction</li> </ul> <p><i>Bit Depth:</i></p> <ul style="list-style-type: none"> <li>○ 8-bit grayscale mode for black-and-white, can be produced from a 16-bit grayscale file</li> <li>○ 24-bit RGB mode for color and monochrome (e.g. discolored prints), can be produced from a 48-bit RGB file</li> </ul>

\*If scans of aerial photography will be used for oversized reproduction, follow the scanning recommendations for the next largest format (e.g., if your original is 8"x10", follow the specifications for formats larger than 8"x10" to achieve 6000 pixels across the long dimensions.

## Objects and artifacts:

Recommended Image Parameters	Alternative Minimum
<p>10 to 16 megapixel 24-bit RGB mode image, can be produced from a 48-bit RGB file.</p> <p>If scanning photographic copies of objects and artifacts, see recommended requirements in the appropriate photo charts above.</p>	<p>6 megapixel 24-bit RGB mode image, can be produced from a 48-bit RGB file.</p> <p>If scanning photographic copies of objects and artifacts, see minimum requirements in the appropriate photo charts above.</p>

### High resolution digital photography requirements:

- Images equivalent to 35mm film photography (6 megapixels to 14 megapixels), to medium format film photography (12 megapixels to 22 megapixels), or to large format film photography (18 megapixels to 200 megapixels).
- Images for photo quality prints and printed reproductions with magazine quality halftones, with maximum image quality at a variety of sizes.
- "Megapixel" is millions of pixels, the megapixel measurement is calculated by multiplying the pixel array values: image width in pixels x image height in pixels.

### Actual pixel dimensions and aspect ratio will vary depending on digital camera - illustrative sizes, dimensions, and proportions are:

- 35mm equivalent - Minimum pixel array of 3,000 pixels by 2,000 pixels (6 megapixels, usual default resolution of 72 ppi at 41.7" by 27.8" or equivalent such as 300 ppi at 10" by 6.7"). Pixel array up to 4,500 pixels by 3,100 pixels (14 megapixels, usual default resolution of 72 ppi at 62.5" by 43" or equivalent such as 300 ppi at 15" by 10.3").
- Medium format equivalent - Minimum pixel array of 4,000 pixels by 3,000 pixels (12 megapixels, usual default resolution of 72 ppi at 55.6" by 41.7" or equivalent such as 300 ppi at 13.3" by 10"). Pixel array up to 5,200 pixels by 4,200 pixels (22 megapixels, usual default resolution of 72 ppi at 72.2" by 58.3" or equivalent such as 300 ppi at 17.3" by 14").
- Large format equivalent - Minimum pixel array of 4,800 pixels by 3,700 pixels (18 megapixels, usual default resolution of 72 ppi at 66.7" by 51.4" or equivalent such as 300 ppi at 16" by 12.5"). Pixel array up to 16,000 pixels by 12,500 pixels (200 megapixels, usual default resolution of 72 ppi at 222.2" by 173.6" or equivalent such as 300 ppi at 53.3" by 41.7").

### File Formats – Image files shall be saved using the following formats:

- Uncompressed TIFF (.tif, sometimes called a raw digital camera file) or LZW compressed TIFF preferred for medium and high resolution requirements.
- JPEG File Interchange Format (JFIF, JPEG, or .jpg) at highest quality (least compressed setting) acceptable for medium and high resolution requirements.
- JPEG Interchange Format (JFIF, JPEG, or .jpg) at any compression setting acceptable for low resolution requirements, depending on the subject matter of the photograph.
- Using the TIFF format and JFIF/JPEG format with high-quality low compression will result in relatively large image file sizes. Consider using larger memory cards, such as 128 MB or larger, or connecting the camera directly to a computer. Select digital cameras that use common or popular memory card formats.

### Image Quality - Digital cameras shall produce high quality image files, including:

- No clipping of image detail in the highlights and shadows for a variety of lighting conditions.
- Accurate color and tone reproduction and color saturation for a variety of lighting conditions.
- Image files may be adjusted after photography using image processing software, such as Adobe Photoshop or JASC Paint Shop Pro. It is desirable to get a good image directly from the camera and to do as little adjustment after photography.
- Digital images shall have minimal image noise and other artifacts that degrade image quality.

- Subject of the photographs shall be in focus, using either auto or manual focus.
- Use of digital zoom feature may have a detrimental effect on the image quality, a smaller portion of the overall image is interpolated to a larger file (effectively lowering resolution).

White Balance – Digital cameras shall be used on automatic white balance or the white balance shall be selected manually to match the light source.

Color Profile – Image files saved with a custom ICC image profile (done in camera or profile produced after photography using profiling software) or a standard color space like sRGB should be converted to a standard wide-gamut color space like Adobe RGB 1998.

Header Data – If camera supports EXIF header data, data in all tags shall be saved.

Image Stitching - Some cameras and many software applications will stitch multiple images into a single image, such as stitching several photographs together to create a composite or a panorama. The stitching process identifies common features within overlapping images and merges the images along the areas of overlap. This process may cause some image degradation. Consider saving and maintaining both the individual source files and the stitched file.

## VI. STORAGE

### File Formats:

We recommend the Tagged Image File Format or TIFF for production master files. Use TIFF version 6, with Intel (Windows) byte order. For additional information on file formats for production masters, see Appendix D, File Format Comparison.

Uncompressed files are recommended, particularly if files are not actively managed, such as storage on CD-ROM or DVD-ROM. If files are actively managed in a digital repository, then you may want to consider using either LZW or ZIP lossless compression for the TIFF files. Do not use JPEG compression within the TIFF format.

### File Naming:

A file naming scheme should be established prior to capture. The development of a file naming system should take into account whether the identifier requires machine- or human-indexing (or both—in which case, the image may have multiple identifiers). File names can either be meaningful (such as the adoption of an existing identification scheme which correlates the digital file with the source material), or non-descriptive (such as a sequential numerical string). Meaningful file names contain metadata that is self-referencing; non-descriptive file names are associated with metadata stored elsewhere that serves to identify the file. In general, smaller-scale projects may design descriptive file names that facilitate browsing and retrieval; large-scale projects may use machine-generated names and rely on the database for sophisticated searching and retrieval of associated metadata.

In general, we recommend that file names-

- Are unique.
- Are consistently structured.
- Take into account the maximum number of items to be scanned and reflect that in the number of digits used (if following a numerical scheme).
- Use leading 0's to facilitate sorting in numerical order (if following a numerical scheme).
- Do not use an overly complex or lengthy naming scheme that is susceptible to human error during manual input.
- Use lowercase characters and file extensions.
- Use numbers and/or letters but not characters such as symbols or spaces that could cause complications across operating platforms.
- Record metadata embedded in file names (such as scan date, page number, etc.) in another location in addition to the file name. This provides a safety net for moving files across systems in the future, in the event that they must be renamed.
- In particular, sequencing information and major structural divisions of multi-part objects should be explicitly recorded in the structural metadata and not only embedded in filenames.
- Although it is not recommended to embed too much information into the file name, a certain amount of information can serve as minimal descriptive metadata for the file, as an economical alternative to the provision of richer data elsewhere.
- Alternatively, if meaning is judged to be temporal, it may be more practical to use a simple numbering system. An intellectually meaningful name will then have to be correlated with the digital resource in the database.

### Directory structure-

Regardless of file name, files will likely be organized in some kind of file directory system that will link to metadata stored elsewhere in a database. Production master files might be stored separately from derivative files, or directories may have their own organization independent of the image files, such as folders arranged by date or record group number, or they may replicate the physical or logical organization of the originals being scanned.

The files themselves can also be organized solely by directory structure and folders rather than embedding meaning in the file name. This approach generally works well for multi-page items. Images are uniquely identified and aggregated at the level of the logical object (i.e., a book, a chapter, an issue, etc.), which requires that the folders or directories be named descriptively. The file names of the individual images themselves are unique only within each directory, but not across directories. For example, book 0001 contains image files 001.tif, 002.tif, 003.tif, etc. Book 0002 contains image files 001.tif, 002.tif, 003.tif. The danger with this approach is that if individual images are separated from their parent directory, they will be indistinguishable from images in a different directory.

In the absence of a formal directory structure, we are currently using meaningful file names. The item being scanned is assigned a 5-digit unique identifier (assigned at the logical level). This identifier has no meaning in the scanning process, but does carry meaning in a system that links the image file(s) to descriptive information. Also embedded in the file name is the year the file was scanned as well as a 3-digit sequential number that indicates multiple pages. This number simply records the number of files belonging to an object; it does not correlate with actual page numbers. The organization is: logical item ID\_scan year\_page or file number\_role of image.tif; e.g., 00001\_2003\_001\_MA.tif.

#### **Versioning-**

For various reasons, a single scanned object may have multiple but differing versions associated with it (for example, the same image prepped for different output intents, versions with additional edits, layers, or alpha channels that are worth saving, versions scanned on different scanners, scanned from different original media, scanned at different times by different scanner operators, etc.). Ideally, the description and intent of different versions should be reflected in the metadata; but if the naming convention is consistent, distinguishing versions in the file name will allow for quick identification of a particular image. Like derivative files, this usually implies the application of a qualifier to part of the file name. The reason to use qualifiers rather than entirely new names is to keep all versions associated with a logical object under the same identifier. An approach to naming versions should be well thought out; adding 001, 002, etc. to the base file name to indicate different versions is an option; however, if 001 and 002 already denote page numbers, a different approach will be required.

#### **Naming derivative files-**

The file naming system should also take into account the creation of derivative image files made from the production master files. In general, derivative file names are inherited from the production masters, usually with a qualifier added on to distinguish the role of the derivative from other files (i.e., "pr" for printing version, "t" for thumbnail, etc.) Derived files usually imply a change in image dimensions, image resolution, and/or file format from the production master. Derivative file names do not have to be descriptive as long as they can be linked back to the production master file.

For derivative files intended primarily for Web display, one consideration for naming is that images may need to be cited by users in order to retrieve other higher-quality versions. If so, the derivative file name should contain enough descriptive or numerical meaning to allow for easy retrieval of the original or other digital versions.

#### **Storage Recommendations:**

We recommend that production master image files be stored on hard drive systems with a level of data redundancy, such as RAID drives, rather than on optical media, such as CD-R. An additional set of images with metadata stored on an open standard tape format (such as LTO) is recommended (CD-R as backup is a less desirable option), and a backup copy should be stored offsite. Regular backups of the images onto tape from the RAID drives is also recommended. A checksum should be generated and should be stored with the image files.

Currently, we use CD-ROMs for distribution of images to external sources, not as a long-term storage medium. However, if images are stored on CD-ROMs, we recommend using high quality or "archival" quality CD-Rs (such as Mitsui Gold Archive CD-Rs). The term "archival" indicates the materials used to manufacture the CD-R (usually the dye layer where the data is recording, a protective gold layer to prevent pollutants from attacking the dye, or a physically durable top-coat to protect the surface of the disk) are reasonably stable and have good durability, but this will not guarantee the longevity of the media itself. All disks need to be stored and handled properly. We have found files stored on brand name CD-Rs that we have not been able to open less than a year after they have been written to the media. We recommend not using inexpensive or non-brand name CD-Rs, because generally they will be less stable, less durable, and more prone to recording problems. Two (or more) copies should be made; one copy should not be handled and should be stored offsite. Most importantly, a procedure for migration of the files off of the CD-ROMs should be in place. In addition, all copies of the CD-ROMs should be periodically checked using a metric such as a CRC (cyclic redundancy checksum) for data integrity. For large-scale projects or for projects that create very large image files, the limited capacity of CD-R storage will be problematic. DVD-Rs may be considered for large projects, however, DVD formats are not as standardized as the lower-capacity CD-ROM formats, and compatibility and obsolescence in the near future is likely to be a problem.

#### **Digital repositories and the long-term management of files and metadata-**

Digitization of archival records and creation of metadata represent a significant investment in terms of time and money. Is it important to realize the protection of these investments will require the active management of both the



image files and the associated metadata. Storing files to CD-R or DVD-R and putting them on a shelf will not ensure the long-term viability of the digital images or the continuing access to them.

We recommend digital image files and associated metadata be stored and managed in a digital repository, see [www.rlg.org/longterm](http://www.rlg.org/longterm), [www.nla.gov.au/padi/](http://www.nla.gov.au/padi/), and [www.dpconline.org/](http://www.dpconline.org/). The Open Archival Information System (OAIS) reference model standard describes the functionality of a digital repository- see [www.rlg.org/longterm/oais.html](http://www.rlg.org/longterm/oais.html) and <http://ssdoo.gsfc.nasa.gov/nost/isoas/overview.html>.

NARA is working to develop a large scale IT infrastructure for the management of, preservation of, and access to electronic records, the Electronic Records Archive (ERA) project. Information is available at [http://www.archives.gov/electronic\\_records\\_archives/index.html](http://www.archives.gov/electronic_records_archives/index.html). ERA will be an appropriate repository for managing and providing access to digital copies of physical records.

## VII. QUALITY CONTROL

Quality control (QC) and quality assurance (QA) are the processes used to ensure digitization and metadata creation are done properly. QC/QA plans and procedures should address issues relating to the image files, the associated metadata, and the storage of both (file transfer, data integrity). Also, QC/QA plans should address accuracy requirements for and acceptable error rates for all aspects evaluated. For large digitization projects it may be appropriate to use a statistically valid sampling procedure to inspect files and metadata. In most situations QC/QA are done in a 2-step process- the scanning technician will do initial quality checks during production and this is followed by a second check by another person.

A quality control program should be initiated, documented, and maintained throughout all phases of digital conversion. The quality control plan should address all specifications and reporting requirements associated with each phase of the conversion project.

### **Completeness-**

We recommend verification that 100% of the required images files and associated metadata have been completed or provided.

### **Inspection of digital image files-**

The overall quality of the digital images and metadata will be evaluated using the following procedures. The visual evaluation of the images shall be conducted while viewing the images at a 1 to 1 pixel ratio or 100% magnification on the monitor.

We recommend, at a minimum, 10 images or 10 % of each batch of digital images, whichever quantity is larger, should be inspected for compliance with the digital imaging specifications and for defects in the following areas:

#### **File Related-**

- o Files open and display
- o Proper format
  - o TIFF
- o Compression
  - o Compressed if desired
  - o Proper encoding (LZW, ZIP)
- o Color mode
  - o RGB
  - o Grayscale
  - o Bitonal
- o Bit depth
  - o 24-bits or 48-bits for RGB
  - o 8-bits or 16-bits for grayscale
  - o 1-bit for bitonal
- o Color profile (missing or incorrect)
- o Paths, channels, and layers (present if desired)

#### **Original/Document Related-**

- o Correct dimensions
- o Spatial resolution
  - o Correct resolution

- Correct units (inches or cm)
- Orientation
  - Document- portrait/vertical, landscape/horizontal
  - Image- horizontally or vertically flipped
- Proportions/Distortion
  - Distortion of the aspect ratio
  - Distortion of or within individual channels
- Image skew
- Cropping
  - Image completeness
  - Targets included
- Scale reference (if present, such as engineering scale or ruler)
- Missing pages or images

Metadata Related - see below for additional inspection requirements relating to metadata-

- Named properly
- Data in header tags (complete and accurate)
- Descriptive metadata (complete and accurate)
- Technical metadata (complete and accurate)
- Administrative metadata (complete and accurate)

Image Quality Related-

- Tone
  - Brightness
  - Contrast
  - Target assessment – aimpoints
  - Clipping – detail lost in high values (highlights) or dark values (shadows) – not applicable to 1-bit images
- Color
  - Accuracy
  - Target assessment – aimpoints
  - Clipping – detail lost in individual color channels
- Aimpoint variability
- Saturation
- Channel registration
  - Misregistration
  - Inconsistencies within individual channels
- Quantization errors
  - Banding
  - Posterization
- Noise
  - Overall
  - In individual channels
  - In areas that correspond to the high density areas of the original
  - In images produced using specific scanner or camera modes
- Artifacts
  - Defects
  - Dust
  - Newton’s rings
  - Missing scan lines, discontinuities, or dropped-out pixels
- Detail
  - Loss of fine detail
  - Loss of texture
- Sharpness
  - Lack of sharpness
  - Over-sharpened
  - Inconsistent sharpness
- Flare
- Evenness of tonal values, of illumination, and vignetting or lens fall-off (with digital cameras)

This list has been provided as a starting point, it should not be considered comprehensive.

### Quality control of metadata-

Quality control of metadata should be integrated into the workflow of any digital imaging project. Because metadata is critical to the identification, discovery, management, access, preservation, and use of digital resources, it should be subject to quality control procedures similar to those used for verifying the quality of digital images. Since metadata is often created and modified at many points during an image's life cycle, metadata review should be an ongoing process that extends across all phases of an imaging project and beyond.

As with image quality control, a formal review process should also be designed for metadata. The same questions should be asked regarding who will review the metadata, the scope of the review, and how great a tolerance is allowed for errors.

Practical approaches to metadata review may depend on how and where the metadata is stored, as well as the extent of metadata recorded. It is less likely that automated techniques will be as effective in assessing the accuracy, completeness, and utility of metadata *content* (depending on its complexity), which will require some level of manual analysis. Metadata quality assessment will likely require skilled human evaluation rather than machine evaluation. However, some aspects of managing metadata stored within a system can be monitored using automated system tools (for example, a digital asset management system might handle verification of relationships between different versions of an image, produce transaction logs of changes to data, produce derivative images and record information about the conversion process, run error detection routines, etc.). Tools such as checksums (for example, the MD5 Message-Digest Algorithm) can be used to assist in the verification of data that is transferred or archived.

Although there are no clearly defined metrics for evaluating metadata quality, the areas listed below can serve as a starting point for metadata review. Good practice is to review metadata at the time of image quality review. In general, we consider:

- *Adherence to standards set by institutional policy or by the requirements of the imaging project.*

Conformance to a recognized standard, such as Dublin Core for descriptive metadata and the *NISO Data Dictionary – Technical Metadata for Digital Still Images* for technical and production metadata, is recommended and will allow for better exchange of files and more straightforward interpretation of the data. Metadata stored in encoded schemes such as XML can be parsed and validated using automated tools; however, these tools do not verify accuracy of the content, only accurate syntax. We recommend the use of controlled vocabulary fields or authority files whenever possible to eliminate ambiguous terms; or the use of a locally created standardized terms list.

- *Procedures for accommodating images with incomplete metadata.*

Often images obtained from various sources are represented among the digital images that NARA manages. Procedures for dealing with images with incomplete metadata should be in place. The minimal amount of metadata that is acceptable for managing images (such as a unique identifier, or a brief descriptive title or caption, etc.) should be determined. If there is no metadata associated with an image, would this preclude the image from being maintained over time?

- *Relevancy and accuracy of metadata.*

How are data input errors handled? Poor quality metadata means that a resource is essentially invisible and cannot be tracked or used. Check for correct grammar, spelling, and punctuation, especially for manually keyed data.

- *Consistency in the creation of metadata and in interpretation of metadata.*

Data should conform to the data constraints of header or database fields, which should be well-defined. Values entered into fields should not be ambiguous. Limit the number of free text fields. Documentation such as a data dictionary can provide further clarification on acceptable field values.

- *Consistency and completeness in the level at which metadata is applied.*

Metadata is collected on many hierarchical levels (file, series, collection, record group, etc.), across many versions (format, size, quality), and applies to different logical parts (item or document level, page level, etc.). Information may be mandatory at some levels and not at others. Data constants can be applied at higher levels and inherited down if they apply to all images in a set.

- *Evaluation of the usefulness of the metadata being collected.*

Is the information being recorded useful for resource discovery or management of image files over time? This is an ongoing process that should allow for new metadata to be collected as necessary.

○ *Synchronization of metadata stored in more than one location.*

Procedures should be in place to make sure metadata is updated across more than one location. Information related to the image might be stored in the TIFF header, the digital asset management system, and other databases, for example.

○ *Representation of different types of metadata.*

Has sufficient descriptive, technical, and administrative metadata been provided? All types must be present to ensure preservation of and access to a resource. All mandatory fields should be complete.

○ *Mechanics of the metadata review process.*

A system to track the review process itself is helpful; this could be tracked using a database or a folder system that indicates status.

Specifically, we consider:

○ *Verifying accuracy of file identifier.*

File names should consistently and uniquely identify both the digital resource and the metadata record (if it exists independently of the file). File identifiers will likely exist for the metadata record itself in addition to identifiers for the digitized resource, which may embed information such as page or piece number, date, project or institution identifier, among others. Information embedded in file identifiers for the resource should parallel metadata stored in a database record or header. Identifiers often serve as the link from the file to information stored in other databases and must be accurate to bring together distributed metadata about a resource. Verification of identifiers across metadata in disparate locations should be made.

○ *Verifying accuracy and completeness of information in image header tags.*

The file browser tool in Adobe Photoshop 7.0 can be used to display some of the default TIFF header fields and IPTC fields for quick review of data in the header; however, the tool does not allow for the creation or editing of header information. Special software is required for editing TIFF header tags.

○ *Verifying the correct sequence and completeness of multi-page items.*

Pages should be in the correct order with no missing pages. If significant components of the resource are recorded in the metadata, such as chapter headings or other intellectual divisions of a resource, they should match up with the actual image files. For complex items such as folded pamphlets or multiple views of an item (a double page spread, each individual page, and a close-up section of a page, for example), a convention for describing these views should be followed and should match with the actual image files.

○ *Adherence to agreed-upon conventions and terminology.*

Descriptions of components of multi-page pieces (i.e., is “front” and “back” or “recto” and “verso” used?) or descriptions of source material, for example, should follow a pre-defined, shared vocabulary.

**Documentation-**

Quality control data (such as logs, reports, decisions) should be captured in a formal system and should become an integral part of the image metadata at the file or the project level. This data may have long-term value that could have an impact on future preservation decisions.

**Testing results and acceptance/rejection-**

If more than 1% of the total number of images and associated metadata in a batch, based on the randomly selected sampling, are found to be defective for any of the reasons listed above, the entire batch should be re-inspected. Any specific errors found in the random sampling and any additional errors found in the re-inspection should be corrected. If less than 1% of the batch is found to be defective, then only the specific defective images and metadata that are found should be redone.

## APPENDIX A: Digitizing for Preservation vs. Production Masters:

In order to consider using digitization as a method of preservation reformatting it will be necessary to specify much more about the characteristics and quality of the digital images than just specifying spatial resolution.

The following chart provides a comparison of image characteristics for preservation master image files and production master image files-

	Preservation Master Files	Production Master Files
Tone reproduction	<p>We need to use well defined, conceptually valid, and agreed upon approaches to tone reproduction that inform current and future users about the nature of the originals that were digitized. At this point in time, no approaches to tone reproduction have been agreed upon as appropriate for preservation digitization.</p> <p>If analog preservation reformatting is used as a model, then one analogous conceptual approach to tone reproduction would be to digitize so the density values of the originals are rendered in a linear relationship to the lightness channel in the LAB color mode. The lightness channel should be correlated to specified density ranges appropriate for different types of originals- as examples, for most reflection scanning a range of 2.0 to 2.2, for transmission scanning of most older photographic negatives a range of 2.0 to 2.2, and for transmission scanning of color transparencies /slides a range of 3.2 to 3.8.</p> <p>Many tone reproduction approaches that tell us about the nature of the originals are likely to produce master image files that are not directly usable on-screen or for printing without adjustment. It will be necessary to make production master derivative files brought to a common rendition to facilitate use. For many types of master files this will be a very manual process (like images from photographic negatives) and will not lend itself to automation.</p> <p>The need for a known rendering in regards to originals argues against saving raw and unadjusted files as preservation masters.</p> <p>For some types of originals, a tone reproduction based upon average or generic monitor display (as described in these <i>Technical Guidelines</i>) may be appropriate for preservation master files.</p>	<p>Images adjusted to achieve a common rendering and to facilitate the use of the files and batch processing.</p> <p>Tone reproduction matched to generic representation – tones distributed in a non-linear fashion.</p>
Tonal Orientation	<p>For preservation digitization the tonal orientation (positive or negative) for master files should be the same as the originals. This approach informs users about the nature of the originals, the images of positive originals appear positive and the images of photographic negatives appear negative. This approach would require production master image files be produced of images of negatives and the tonal orientation inverted to positive images. The master image files of photographic negatives will not be directly usable.</p>	<p>All images have positive tonal orientation.</p>

## Color reproduction

We need to use well defined, conceptually valid, and agreed upon approaches to color reproduction that inform current and future users about the nature of the originals that were digitized. At this point in time, no approaches to color reproduction have been agreed upon as appropriate for preservation digitization.

Device independence and independence from current technical approaches that may change over time (such as ICC color management) are desirable.

Conceptually, LAB color mode may be more appropriate than RGB mode. Although, since scanners/digital cameras all capture in RGB, the images have to be converted to LAB and this process does entail potential loss of image quality. Also, LAB master files would have to be converted back to RGB to be used, another transformation and potential loss of image quality.

Also, the imaging field is looking at multi-spectral imaging to provide the best color reproduction and to eliminate problems like metamerisms. At this time, standard computer software is not capable of dealing with multi-spectral data. Also, depending on the number of bands of wavelengths sampled, the amount of data generated is significantly more than standard 3-channel color digitization. If multi-spectral imaging was feasible from a technical perspective, it would be preferable for preservation digitization. However, at this time there is no simple raster image format that could be used for storing multi-spectral data. The JPEG 2000 file format could be used, but this is a high-encoded wavelet based format that does not save the raster data (it does not save the actual bits that represent the pixels, instead it recreates the data representing the pixels). To use a simple raster image format like TIFF it would probably be necessary to convert the multi-spectral data to 3-channel RGB data; hopefully this would produce a very accurate RGB file, but the multi-spectral data would not be saved.

Images adjusted to achieve a common rendering and to facilitate the use of the files and batch processing.

Color reproduction matched to generic RGB color space. Intent is to be able to use files both within and outside of current ICC color managed process.

## Bit depth

High bit-depth digitization is preferred, either 16-bit grayscale images or 48-bit RGB color images.

Standard 8-bit per channel imaging has only 256 levels of shading per channel, while 16-bit per channel imaging has thousands of shades per channel making them more like the analog originals.

High bit-depth necessary for standard 3-channel color digitization to achieve the widest gamut color reproduction.

Currently, it is difficult to verify the quality of high-bit image files.

Traditional 8-bit grayscale and 24-bit RGB files produced to an appropriate quality level are sufficient.

## Resolution

Requires sufficient resolution to capture all the significant detail in originals.

Currently the digital library community seems to be reaching a consensus on appropriate resolution levels for preservation digitization of text based originals – generally 400 ppi for grayscale and color digitization is considered sufficient as long as a QI of 8 is maintained for all significant text. This approach is

Generally, current approaches are acceptable (see requirements in these *Technical Guidelines*).

based on typical legibility achieved on 35mm microfilm (the current standard for preservation reformatting of text-based originals), and studies of human perception indicate this is a reasonable threshold in regards to the level of detail perceived by the naked eye (without magnification). Certainly all originals have extremely fine detail that is not accurately rendered at 400 ppi. Also, for some reproduction requirements this resolution level may be too low, although the need for very large reproduction is infrequent.

Unlike text-based originals, it is very difficult to determine appropriate resolution levels for preservation digitization of many types of photographic originals. For analog photographic preservation duplication, the common approach is to use photographic films that have finer grain and higher resolution than the majority of originals being duplicated. The analogous approach in the digital environment would be to digitize all photographic camera originals at a resolution of 3,000 ppi to 4,000 ppi regardless of size. Desired resolution levels may be difficult to achieve given limitations of current scanners.

#### File size

The combination of both high bit-depth and high resolution digitization will result in large to extremely large image files. These files will be both difficult and expensive to manage and maintain.

Moderate to large files sizes.

If multi-spectral image is used, file sizes will be even larger. Although, generally it is assumed a compressed format like JPEG 2000 would be used and would compensate for some of the larger amount of data.

#### Other image quality parameters

Preservation master images should be produced on equipment that meets the appropriate levels for the following image quality parameters at a minimum:

- Ability to capture and render large dynamic ranges for all originals.
- Appropriate spatial frequency response to capture accurately fine detail at desired scanning resolutions.
- Low image noise over entire tonal range and for both reflective and transmissive originals.
- Accurate channel registration for originals digitized in color.
- Uniform images without tone and color variation due to deficiencies of the scanner or digitization set-up.
- Dimensionally accurate and consistent images.
- Free from all obvious imaging defects.

Generally, current equipment and approaches are acceptable (see requirements in these *Technical Guidelines*).

We need to use well defined, conceptually valid, and agreed upon approaches to these image quality parameters. At this point in time, no approaches have been agreed upon as appropriate for preservation digitization.



Three dimensional and other physical aspects of documents

We need to acknowledge digitization is a process that converts three-dimensional objects (most of which are very flat, but are three-dimensional nonetheless) into two-dimensional images or representations. Most scanners are designed with lighting to minimize the three dimensional aspects of the original documents being scanned, in order to emphasize the legibility of the text or writing. So not all of the three-dimensional aspects of the documents are recorded well and in many cases are not recorded at all; including properties and features like paper texture and fibers, paper watermarks and laid lines, folds and/or creases, embossed seals, etc. Loss of three-dimensional information may influence a range of archival/curatorial concerns regarding preservation reformatting.

These concerns are not unique to digital reformatting, traditional approaches to preservation reformatting, such as microfilming, photocopying (electrophotographic copying on archival bond), and photographic copying/duplication have the same limitations – they produce two-dimensional representations of three-dimensional originals.

One example of a concern about rendering three-dimensional aspects of documents that has legal implications is documents with embossed seals and questions about the authenticity of the digital representation of the documents when the seals are not visible and/or legible in the digital images (a common problem, see Digitization Specifications for Record Types for a short discussion of lighting techniques to improve legibility of embossed seals).

Other issues that may need to be considered and appropriate approaches defined prior to starting any reformatting include, but limited to, the following:

- Digitize front and/or back of each document or page – even if no information is on one side.
- Reflection and/or transmission scanning for all materials – to record watermarks, laid lines, paper structure and texture, any damage to the paper, etc.
- Use of diffuse and/or raking light – digitize using diffuse light to render text and/or writing accurately, and/or digitize using raking light to render the three-dimensionality of the document (folds, creases, embossed seals, etc.).
- Digitize documents folded and/or unfolded.
- Digitize documents with attachments in place and/or detached as separate documents.
- Digitize documents bound and/or unbound.

The question that needs to be answered, and there will probably not be a single answer, is how many representations are needed for preservation reformatting to accurately document the original records? The digital library community needs to discuss these issues and arrive at appropriate approaches for different types of originals. One additional comment, originals for which it is considered appropriate to have multiple representations in order to be considered preservation reformatting probably warrant preservation in original form.

Generally, digitization limited to one version without consideration of the representation of three-dimensional aspects of the original records.

## APPENDIX B: Derivative Files

The parameters for access files will vary depending on the types of materials being digitized and the needs of the users of the images. There is no set size or resolution for creating derivative access files. The following charts provide some general recommendations regarding image size, resolution, and file formats for the creation of derivative images from production master image files.

From a technical perspective, records that need similar derivatives have been grouped together-

- o textual records and graphic illustrations/artwork/originals
- o photographs and objects/artifacts
- o maps/plans/oversized and aerial photography

The charts have been divided into sections representing two different approaches to web delivery of the derivatives-

- o fixed-sized image files for static access via a web browser
- o dynamic access via a web browser

JPEG compression was designed for photographic images and sacrifices fine detail to save space when stored, while preserving the large features of an image. JPEG compression creates artifacts around text when used with digital images of text documents at moderate to high compression levels. Also, JPEG files will be either 24-bit RGB images or 8-bit grayscale, they can not have lower bit depths.

GIF files use LZW compression (typical compression ratio is 2:1, or the file will be half original size), which is lossless and does not create image artifacts; therefore, GIF files may be more suitable for access derivatives of text documents. The GIF format supports 8-bit (256 colors), or lower, color files and 8-bit, or lower, grayscale files. All color GIF files and grayscale GIF files with bit-depths less than 8-bits are usually dithered (the distribution of pixels of different shades in areas of another shade to simulate additional shading). Well dithered images using an adaptive palette and diffusion dither will visually have a very good appearance, including when used on photographic images. In many cases a well produced GIF file will look better, or no worse, than a highly compressed JPEG file (due to the JPEG artifacts and loss of image sharpness), and for textual records the appearance of a GIF format derivative is often significantly better than a comparable JPEG file.

The following table compares the uncompressed and compressed file sizes for the same image when using GIF format vs. JPEG format:

For an 800x600 pixel access file, assumes 2:1 compression for GIF and 20:1 for JPEG-

	Color Image		Grayscale Image	
	GIF 8-bit	JPEG 24-bit	GIF 4-bit	JPEG 8-bit
Open File Size	480 KB	1.44 MB (3 times larger than open GIF)	240 KB	480 KB (2 times larger than open GIF)
Stored File Size	240 KB (3 times larger stored JPEG)	72 KB	120 KB (5 times larger than stored JPEG)	24 KB

As you can see, when the files are open the GIF file will be smaller due to the lower bit-depth and when stored the JPEG will be smaller due to the higher compression ratio. GIF files will take longer to download, but will decompress quicker and put less demand on the end user's CPU in terms of memory and processor speed. JPEG files will download quicker, but will take longer to decompress putting a greater demand on the end user's CPU. Practical tests have shown a full page of GIF images generally will download, decompress, and display more quickly than the same page full of JPEG versions of the images.

The newer JPEG 2000 compression algorithm is a wavelet compression, and can be used to compress images to higher compression ratios with less loss of image quality compared to the older JPEG algorithm. Generally, JPEG 2000 will not produce the same severity of artifacts around text that the original JPEG algorithm produces.

Access Approach and Derivative File Type		Record Types	
		Textual Records*	Graphic Illustrations/ Artwork/ Originals**
<b>Fixed-Sized Image Files for Static Access via a Web Browser</b>			
Thumbnail*		<ul style="list-style-type: none"> <li>○ <b>File Format:</b> GIF (adaptive/perceptual palette, diffusion/noise dither) or JPG (low to medium quality compression, sRGB profile for color and Gamma 2.2 profile for grayscale)</li> <li>○ <b>Pixel Array:</b> not to exceed an array of 200x200 pixels</li> <li>○ <b>Resolution:</b> 72 ppi</li> </ul>	
Access – requirements for access files will vary depending on the size of the originals, text legibility, and the size of the smallest significant text characters.	Minimum	<ul style="list-style-type: none"> <li>○ <b>File Format:</b> GIF (for smaller originals, adaptive/perceptual palette, diffusion/noise dither) or JPG (for larger originals, low to medium quality compression, sRGB profile for color and Gamma 2.2 for grayscale)</li> <li>○ <b>Image Size:</b> original size</li> <li>○ <b>Resolution:</b> 72 ppi to 90 ppi</li> </ul>	
	Recommended	<ul style="list-style-type: none"> <li>○ <b>File Format:</b> GIF (for smaller originals, adaptive/perceptual palette, diffusion/noise dither) or JPG (for larger originals, medium to high quality compression, sRGB profile for color and Gamma 2.2 for grayscale)</li> <li>○ <b>Image Size:</b> original size</li> <li>○ <b>Resolution:</b> 90 ppi to 120 ppi</li> </ul>	
	Larger Alternative	<ul style="list-style-type: none"> <li>○ <b>File Format:</b> GIF (for smaller originals, adaptive/perceptual palette, diffusion/noise dither) or JPG (for larger originals, medium to high quality compression, sRGB profile for color and Gamma 2.2 for grayscale)</li> <li>○ <b>Image Size:</b> original size</li> <li>○ <b>Resolution:</b> 120 ppi to 200 ppi</li> </ul>	
Printing and Reproduction – for printing full page images from within a web browser and for magazine quality reproduction at approx. 8.5"x11"		<ul style="list-style-type: none"> <li>○ <b>File Format:</b> PDF (JPEG compression at high quality, Adobe 1998 profile for color and Gamma 2.2 for grayscale)</li> <li>○ <b>Image Size:</b> fit within and not to exceed dimensions of 8"x10.5" (portrait or landscape orientation)</li> <li>○ <b>Resolution:</b> 300 ppi</li> </ul>	
<b>Alternative – Dynamic Access via a Web Browser</b>			
Access – High Resolution – requires special server software and allows zooming, panning, and download of high resolution images.		<ul style="list-style-type: none"> <li>○ <b>File Format:</b> JPEG 2000 (wavelet encoding) or traditional raster file formats like TIFF or JPEG (lossy compression at high quality, Adobe 1998 profile for color and Gamma 2.2 for grayscale)</li> <li>○ <b>Image Size:</b> original size</li> <li>○ <b>Resolution:</b> same resolution as production master file</li> </ul>	

\*Many digitization projects do not make thumbnail files for textual records - the text is not legible and most documents look alike when the images are this small, so thumbnails may have limited usefulness. However, thumbnail images may be needed for a variety of web uses or within a database, so many projects do create thumbnails from textual documents.

\*\*Includes posters, artwork, illustrations, etc., generally would include any item that is graphic in nature and may have text as well.

Access Approach and Derivative File Type		Record Types	
		Photographs	Objects and Artifacts
<b>Fixed-Sized Image Files for Static Access via a Web Browser</b>			
Thumbnail		<ul style="list-style-type: none"> <li>○ <b>File Format:</b> GIF (adaptive/perceptual palette, diffusion/noise dither) or JPG (low to medium quality compression, sRGB profile for color and Gamma 2.2 profile for grayscale)</li> <li>○ <b>Pixel Array:</b> not to exceed an array of 200x200 pixels</li> <li>○ <b>Resolution:</b> 72 ppi</li> </ul>	
Access –	Minimum	<ul style="list-style-type: none"> <li>○ <b>File Format:</b> GIF (for smaller originals, adaptive/perceptual palette, diffusion/noise dither) or JPG (for larger originals, low to medium quality compression, sRGB profile for color and Gamma 2.2 for grayscale)</li> <li>○ <b>Pixel Array:</b> array fit within 600x600 pixels at a minimum and up to 800x800 pixels</li> <li>○ <b>Resolution:</b> 72 ppi</li> </ul>	
	Recommended	<ul style="list-style-type: none"> <li>○ <b>File Format:</b> GIF (for smaller originals, adaptive/perceptual palette, diffusion/noise dither) or JPG (for larger originals, medium to high quality compression, sRGB profile for color and Gamma 2.2 for grayscale)</li> <li>○ <b>Pixel Array:</b> array fit within 800x800 pixels at a minimum and up to 1200x1200 pixels</li> <li>○ <b>Resolution:</b> 72 ppi</li> </ul>	
	Larger Alternative	<ul style="list-style-type: none"> <li>○ <b>File Format:</b> GIF (for smaller originals, adaptive/perceptual palette, diffusion/noise dither) or JPG (for larger originals, medium to high quality compression, sRGB profile for color and Gamma 2.2 for grayscale)</li> <li>○ <b>Pixel Array:</b> array fit within 1200x1200 pixels at a minimum and up to 2000x2000 pixels</li> <li>○ <b>Resolution:</b> 72 ppi, or up to 200 ppi</li> </ul>	
Printing and Reproduction – for printing full page images from within a web browser and for magazine quality reproduction at approx. 8.5"x11"		<ul style="list-style-type: none"> <li>○ <b>File Format:</b> PDF (JPEG compression at high quality, Adobe 1998 profile for color and Gamma 2.2 for grayscale)</li> <li>○ <b>Image Size:</b> fit within and not to exceed dimensions of 8"x10.5" (portrait or landscape orientation)</li> <li>○ <b>Resolution:</b> 300 ppi</li> </ul>	
<b>Alternative – Dynamic Access via a Web Browser</b>			
Access – High Resolution – requires special server software and allows zooming, panning, and download of high resolution images.		<ul style="list-style-type: none"> <li>○ <b>File Format:</b> JPEG 2000 (wavelet encoding) or traditional raster file formats like TIFF or JPEG (lossy compression at high quality, Adobe 1998 profile for color and Gamma 2.2 for grayscale)</li> <li>○ <b>Image Size:</b> original size</li> <li>○ <b>Resolution:</b> same resolution as production master file</li> </ul>	

Access Approach and Derivative File Type		Record Types	
		Maps, Plans, and Oversized	Aerial Photography
<b>Fixed-Sized Image Files for Static Access via a Web Browser</b>			
Thumbnail		<ul style="list-style-type: none"> <li>○ <b>File Format:</b> GIF (adaptive/perceptual palette, diffusion/noise dither) or JPG (low to medium quality compression, sRGB profile for color and Gamma 2.2 profile for grayscale)</li> <li>○ <b>Pixel Array:</b> not to exceed an array of 200 pixels by 200 pixels</li> <li>○ <b>Resolution:</b> 72 ppi</li> </ul>	
Access –	Minimum	<ul style="list-style-type: none"> <li>○ <b>File Format:</b> GIF (for smaller originals, adaptive/perceptual palette, diffusion/noise dither) or JPG (for larger originals, low to medium quality compression sRGB profile for color and Gamma 2.2 for grayscale)</li> <li>○ <b>Pixel Array:</b> array fit within 800x800 pixels at a minimum and up to 1200x1200 pixels</li> <li>○ <b>Resolution:</b> 72 ppi</li> </ul>	
	Recommended	<ul style="list-style-type: none"> <li>○ <b>File Format:</b> GIF (for smaller originals, adaptive/perceptual palette, diffusion/noise dither) or JPG (for larger originals, medium to high quality compression, sRGB profile for color and Gamma 2.2 for grayscale)</li> <li>○ <b>Pixel Array:</b> array fit within 1200x1200 pixels at a minimum and up to 2000x2000 pixels</li> <li>○ <b>Resolution:</b> 72 ppi, or up to 200 ppi</li> </ul>	
	Larger Alternative	<ul style="list-style-type: none"> <li>○ <b>File Format:</b> GIF (for smaller originals, adaptive/perceptual palette, diffusion/noise dither) or JPG (for larger originals, medium to high quality compression, sRGB profile for color and Gamma 2.2 for grayscale)</li> <li>○ <b>Pixel Array:</b> array fit within 2000x2000 pixels at a minimum and up to 3000x3000 pixels</li> <li>○ <b>Resolution:</b> 72 ppi, or up to 300 ppi</li> </ul>	
Printing and Reproduction – for printing full page images from within a web browser and for magazine quality reproduction at approx. 8.5"x11"		<ul style="list-style-type: none"> <li>○ <b>File Format:</b> PDF (JPEG compression at high quality, Adobe 1998 profile for color and Gamma 2.2 for grayscale)</li> <li>○ <b>Image Size:</b> fit within and not to exceed dimensions of 8"x10.5" (portrait or landscape orientation)</li> <li>○ <b>Resolution:</b> 300 ppi</li> </ul>	
<b>Recommended Alternative – Dynamic Access via a Web Browser</b>			
Access – High Resolution – requires special server software and allows zooming, panning, and download of high resolution images.		<ul style="list-style-type: none"> <li>○ <b>File Format:</b> JPEG 2000 (wavelet encoding) or traditional raster file formats like TIFF or JPEG (lossy compression at high quality, Adobe 1998 profile for color and Gamma 2.2 for grayscale)</li> <li>○ <b>Image Size:</b> original size</li> <li>○ <b>Resolution:</b> same resolution as production master file</li> </ul>	

**APPENDIX C: Mapping of LCDRG Elements to Unqualified Dublin Core  
(Mandatory Metadata Elements excerpted from the Lifecycle Data Requirements Guide [LCDRG]<sup>3</sup>)**

*Mandatory Elements for Record Groups and Collections*

Record Group	Collection	Dublin Core	LCDRG Notes
Title	Title	Title	May be an assigned name that differs from the original name
	Collection Identifier	Identifier	
Record Group Number		Identifier	
Inclusive Start Date	Inclusive Start Date	Date	
Inclusive End Date	Inclusive End Date	Date	
Description Type	Description Type		Level of aggregation

*Mandatory Elements for Series, File Units, and Items*

Series	File Unit	Item	Dublin Core	LCDRG Notes
Title	Title	Title	Title	
Function and Use	<input type="checkbox"/>	<input type="checkbox"/>	Description	Only mandatory for newly created descriptions of organizational records
Inclusive Start Date	<input type="checkbox"/>	<input type="checkbox"/>	Date	Inclusive Start Date for File Unit and Item inherited from Series description
Inclusive End Date	<input type="checkbox"/>	<input type="checkbox"/>	Date	Inclusive End Date for File Unit and Item inherited from Series description
General Records Type	General Records Type	General Records Type	Type	Uses NARA-controlled values
Access Restriction Status	Access Restriction Status	Access Restriction Status	Rights	
Specific Access Restrictions	Specific Access Restrictions	Specific Access Restrictions	Rights	Mandatory if value is present in Access Restriction Status

<sup>3</sup> Lifecycle Data Requirements Guide. Second Revision, January 18, 2002

[http://www.nara-at-work.gov/archives\\_and\\_records\\_mgmt/archives\\_and\\_activities/accessioning\\_processing\\_description/lifecycle/mandatoryelements.html](http://www.nara-at-work.gov/archives_and_records_mgmt/archives_and_activities/accessioning_processing_description/lifecycle/mandatoryelements.html)

Series (cont.)	File Unit (cont.)	Item (cont.)	Dublin Core (cont.)	LCDRG Notes (cont.)
Security Classification	Security Classification	Security Classification	Rights	
Use Restriction Status	Use Restriction Status	Use Restriction Status	Rights	
Specific Use Restrictions	Specific Use Restrictions	Specific Use Restrictions	Rights	Mandatory if value is present in Use Restriction Access
Creating Individual	<input type="checkbox"/>	<input type="checkbox"/>	Creator	Creators at the File Unit and the Item level are inherited from Series description
Creating Individual Type	<input type="checkbox"/>	<input type="checkbox"/>		Most Recent/Predecessor
Creating Organization	<input type="checkbox"/>	<input type="checkbox"/>	Creator	Creators at the File Unit and the Item level are inherited from Series description
Creating Organization Type	<input type="checkbox"/>	<input type="checkbox"/>		Most Recent/Predecessor
Description Type	Description Type	Description Type		Level of Aggregation
Copy Status	Copy Status	Copy Status		Role or purpose of physical occurrence
Extent	<input type="checkbox"/>	<input type="checkbox"/>	Coverage	
GPRA Indicator	<input type="checkbox"/>	<input type="checkbox"/>		
Holdings Measurement Type	<input type="checkbox"/>	<input type="checkbox"/>		Unit by which archival materials are physically counted
Holdings Measurement Count	<input type="checkbox"/>	<input type="checkbox"/>		Numeric value. Quantity of archival materials
Location Facility	Location Facility	Location Facility	Publisher	
Reference Unit	Reference Unit	Reference Unit	Publisher	
Media Type	Media Type	Media Type	Format	Describes both physical occurrence and individual media occurrences

#### *Mandatory Elements for Archival Creators*

Organization Elements	Person Elements	Dublin Core	LCDRG Notes
Organization Name	Name	Title	
Abolish Date	<input type="checkbox"/>	Date	
Establish Date	<input type="checkbox"/>	Date	

Note: Many of the LCDRG elements above use authority lists for data values that may not necessarily map into recommended Dublin Core Metadata Initiative typology for vocabulary terms, data values, and syntax or vocabulary encoding schemes. Please consult the LCDRG for acceptable data values.

This table suggests a simple mapping only. It is evident that Dublin Core elements are extracted from a much richer descriptive set outlined in the LCDRG framework. Dublin Core elements are repeatable to accommodate multiple LCDRG fields; however, repeatability of fields is not equivalent to the complex structure of archival collections that the LCDRG attempts to capture. As a result, mapping to Dublin Core may result in a loss of information specificity and/or meaning in an archival context. A more detailed analysis of how LCDRG values are being implemented in Dublin Core will be necessary.



## APPENDIX D: File Format Comparison

As stated earlier, the choice of file format has a direct affect on the performance of the digital image as well as implications for long term management of the image. Future preservation policy decisions, such as what level of preservation service to apply, are often made on a format-specific basis\*. A selection of file formats commonly used for digital still raster images are listed below. The first table lists general technical characteristics to consider when choosing an appropriate file format as well as a statement on their recommended use in imaging projects. Generally, these are all well-established formats that do not pose a big risk to the preservation of content information; however, it is advised that an assessment of the potential longevity and future functionality of these formats be undertaken for any digital imaging project. The second table attempts to summarize some of these concerns.

File Format	Technical Considerations	Recommended Use
TIFF	<ul style="list-style-type: none"> <li>-“De facto” raster image format used for master files</li> <li>-Simply encoded raster-based format</li> <li>-Accommodates internal technical metadata in header/ extensible and customizable header tags</li> <li>-Supports Adobe’s XMP (Extensible Metadata Platform)</li> <li>-Accommodates large number of color spaces and profiles</li> <li>-Supports device independent color space (CIE L*a*b)</li> <li>-Uncompressed; lossless compression (Supports multiple compression types for 1-bit files). JPEG compression not recommended in TIFF file</li> <li>-High-bit compatible</li> <li>-Can support layers, alpha channels</li> <li>-Accommodates large file sizes</li> <li>-Anticipate greater preservation support in repository settings; preferred raster image format for preservation</li> <li>-Widely supported and used</li> <li>-Long track record (format is over 10 years old)</li> <li>-Potential loss of Adobe support of TIFF in favor of PDF?</li> <li>-Not suitable as access file—no native support in current web browsers</li> </ul>	Preferred format for production master file
PNG	<ul style="list-style-type: none"> <li>-Simple raster format</li> <li>-High-bit compatible</li> <li>-Lossless compression</li> <li>-Supports alpha channels</li> <li>-Not widely adopted by imaging community</li> <li>-Native support available in later web browsers as access file</li> </ul>	Possible format for production master file—not currently widely implemented
JPEG 2000	<ul style="list-style-type: none"> <li>-Not yet widely adopted</li> <li>-More complex model for encoding data (content is not saved as raster data)</li> <li>-Supports multiple resolutions</li> <li>-Extended version supports color profiles</li> <li>-Extended version supports layers</li> <li>-Includes additional compression algorithms to JPEG (wavelet, lossless)</li> <li>-Support for extensive metadata encoded in XML “boxes;” particularly technical, descriptive, and rights metadata. Supports IPTC information; mapping to Dublin Core.</li> </ul>	Possible format for production master file—not currently widely implemented
GIF	<ul style="list-style-type: none"> <li>-Lossy (high color) and lossless compression</li> <li>-Limited color palette</li> <li>-8-bit maximum, color images are dithered</li> <li>-Short decompression time</li> </ul>	Access derivative file use only—recommend for text records

File Format	Technical Considerations	Recommended Use
JFIF/JPEG	<ul style="list-style-type: none"> <li>-Lossy compression, but most software allows for adjustable level of compression</li> <li>-Presence of compression artifacts</li> <li>-Smaller files</li> <li>-High-bit compatible</li> <li>-Longer decompression time</li> <li>-Supports only a limited set of internal technical metadata</li> <li>-Supports a limited number of color spaces</li> <li>-Not suitable format for editing image files—saving, processing, and resaving results in degradation of image quality after about 3 saves</li> </ul>	Access derivative file use only—not recommended for text or line drawings
PDF	<ul style="list-style-type: none"> <li>-Intended to be a highly structured page description language that can contain embedded objects, such as raster images, in their respective formats.</li> <li>-Works better as a container for multiple logical objects that make up a coherent whole or composite document</li> <li>-More complex format due to embedded/externally linked objects</li> <li>-Implements Adobe’s XMP specification for embedding metadata in XML</li> <li>-Can use different compression on different parts of the file; supports multiple compression schemes</li> <li>-Supports a limited number of color spaces</li> </ul>	Not recommended for production master files
[ASCII]	<ul style="list-style-type: none"> <li>-For image files converted to text</li> <li>-Potential loss to look and feel of document/formatting</li> </ul>	N/A
[XML]	<ul style="list-style-type: none"> <li>-For image files converted to text</li> <li>-Hierarchical structure</li> <li>-Good for encoding digital library-like objects or records</li> <li>-Allows for fast and efficient end-user searching for text retrieval</li> <li>-Easily exchanged across platforms/systems</li> </ul>	N/A

\* For example, DSpace directly associates various levels of preservation services with file formats—categorized as supported formats, known formats, and unknown formats. See <http://dspace.org/faqs/index.html#preserve>. The Florida Center for Library Automation (FCLA) specifies preferred, acceptable, and bit-level preservation only categories for certain file formats for their digital archive. See <http://www.fcla.edu/digitalArchive/pdfs/recFormats.pdf>.

For additional information on research into file format longevity, see *Digital Formats for Library of Congress Collections: Factors to Consider When Choosing Digital Formats* by Caroline Arms and Carl Fleischhauer at: <http://memory.loc.gov/ammem/techdocs/digform/>, from which many of the considerations below were taken; see also the Global Digital Format Registry (GDFR) at <http://hul.harvard.edu/gdfr/> for discussion of a centralized, trusted registry for information about file formats.

#### **Longevity Considerations**

- *Documentation:* For both proprietary and open standard formats, is deep technical documentation publicly and fully available? Is it maintained for older versions of the format?
- *Stability:* Is the format supported by current applications? Is the current version backward-compatible? Are there frequent updates to the format or the specification?
- *Metadata:* Does the format allow for self-documentation? Does the format support extensive embedded metadata beyond what is necessary for normal rendering of a file? Can the file support a basic level of descriptive, technical, administrative, and rights metadata? Can metadata be encoded and stored in XML or other standardized formats? Is metadata easily extracted from the file?
- *Presentation:* Does the format contain embedded objects (e.g. fonts, raster images) and/or link out to external objects? Does the format provide functionality for preserving the layout and structure of document, if this is important?
- *Complexity:* Simple raster formats are preferred. Can the file be easily unpacked? Can content be easily separated from the container? Is “uncompressed” an option for storing data? Does the format incorporate external programs (e.g., Javascript, etc.)? Complexity of format is often associated with risk management—more complex formats are assumed to be harder to decode. However, some formats are by necessity complex based on their purpose and intended functionality. Complex formats should not be avoided solely on the basis that they are forecast to be difficult to preserve, at the expense of using the best format for the use of the data it contains.
- *Adoption:* Is the format widely used by the imaging community in cultural institutions? How is it generally used by these stakeholders—as a master format, a delivery format?
- *Continuity:* How long has the format been in existence? Is the file format mature (most of the image formats in the table above have been in existence for over 10 years).
- *Protection:* Does the format accommodate error detection and correction mechanisms and encryption options? These are related to complexity of the file. In general, encryption and digital signatures may deter full preservation service levels.
- *Compression algorithms:* Does the format use standard algorithms? In general, compression use in files may deter full preservation service levels; however, this may have less to do with file complexity and more to do with patent issues surrounding specific compression algorithms.
- *Interoperability:* Is the format supported by many software applications/ OS platforms or is it linked closely with a specific application? Are there numerous applications that utilize this format? Have useful tools been built up around the format? Are there open source tools available to use and develop the format? Is access functionality improved by native support in web browsers?
- *Dependencies:* Does the format require a plug-in for viewing if appropriate software is not available, or rely on external programs to function?
- *Significant properties:* Does the format accommodate high-bit, high-resolution (detail), color accuracy, multiple compression options? (These are all technical qualities important to master image files).
- *Ease of transformation/preservation:* Is it likely that the format will be supported for full functional preservation in a repository setting, or can guarantees currently only be made at the bitstream (content data) level (where only limited characteristics of the format are maintained)?
- *Packaging formats:* In general, packaging formats such as zip and tar files should be acceptable as transfer mechanisms for image file formats. These are not normally used for storage/archiving.

## APPENDIX E: Records Handling for Digitization

All digitization projects should have pre-established handling guidelines. The following provides general guidance on the proper handling of archival materials for digitization projects. This appendix is provided for informational purposes and does not constitute a policy. Handling guidelines may need to be modified for specific projects based on the records being digitized and their condition.

### 1. Physical Security

As records are received for digitization, they should be logged into the lab area for temporary storage. The log should include-

- o date and time records received
- o job or project title (batch identification if applicable)
- o item count
- o NARA citation/identification (including custodial unit or LICON)
- o media or physical description of the records
- o person dropping off records (archivist/technician/etc. and unit)
- o lab personnel log-in or acceptance of records
- o requested due date
- o special instructions
- o date completed
- o date and time records picked-up
- o person picking up records (archivist/technician/etc. and unit)
- o lab personnel log-out of records

The above list is not intended to be comprehensive, other fields may be required or desirable.

Records should be stored in a secure area that provides appropriate physical protection. Storage areas should meet all NARA requirements and environmental standards for records storage or processing areas - see 36 CFR, Part 1228, Subpart K, Facility Standards for Records Storage Facilities at [http://www.archives.gov/about\\_us/regulations/part\\_1228\\_k.html](http://www.archives.gov/about_us/regulations/part_1228_k.html)

### 2. Equipment

- a. Preservation Programs, NWT, shall review and approve all equipment prior to beginning projects.
- b. The unit/partner/contractor shall not use automatic feed devices, drum scanners or other machines that require archival materials to be fed into rollers or wrapped around rollers, that place excessive pressure on archival materials, or require the document to be taped to a cylinder. Motorized transport is acceptable when scanning microfilm.
- c. The unit's/partner's/contractor's equipment shall have platens or copy boards upon which physical items are supported over their entire surface.
- d. The unit/partner/contractor shall not use equipment having devices that exert pressure on or that affix archival materials to any surface. The unit/partner/contractor shall ensure that no equipment comes into contact with archival materials in a manner that causes friction. The unit/partner/contractor shall not affix pressure sensitive adhesive tape, nor any other adhesive substance, to any archival materials.
- e. The unit/partner/contractor shall not use equipment with light sources that raise the surface temperature of the physical item being digitized. The unit/partner/contractor shall filter light sources that generate ultraviolet light. Preservation Programs, NWT shall have the right to review the lighting parameters for digitizing, including the number of times a single item can be scanned, the light intensity, the ultraviolet and infrared content, and the duration of the scan.
- f. The scanning/digitization area shall have sufficient space and flat horizontal work-surfaces (tables, carts, shelves, etc.) to work with and handle the records safely.

### 3. Procedures

- a. Custodial units shall maintain written records of pulling archival materials for digitization and of the receipt of materials when returned to the custodial units. The unit/partner/contractor shall keep any tracking paperwork with the archival materials and/or their containers.
- b. The unit/partner/contractor shall keep all archival materials in their original order and return them to their original jackets or containers. The unit/partner/contractor shall not leave archival materials unattended or uncovered on digitizing equipment or elsewhere. The unit/partner/contractor shall return archival materials left un-digitized, but needed for the next day's work, to their jackets and containers and place them in the appropriate secure storage areas in the unit's/partner's/contractor's work area. The unit/partner/contractor shall return completed batches of archival materials to NARA staff in the unit's/partner's/contractor's work area.
- c. Review of the condition of the records should take place prior to the beginning of the digitization project and shall be done in consultation with Preservation Programs, NWT. During digitization, the unit/partner/contractor shall report archival materials that are rolled (excluding roll film), folded, or in poor condition and cannot be safely digitized, and seek further guidance from NARA custodial staff and Preservation Programs, NWT, before proceeding.
- d. The unit/partner/contractor shall not remove encapsulated archival materials from their encapsulation or sleeved documents from L-sleeves. The unit/partner/contractor may remove L-sleeves with permission of custodial staff.
- e. The unit/partner/contractor shall place archival materials flat on the platen- rolling, pulling, bending, or folding of archival materials is not permitted, and items shall be supported over their entire surface on the platen- no part of an item shall overhang the platen so that it is unsupported at any time. The unit/partner/contractor shall not place archival materials that may be damaged, such as rolled, folded, warped, curling, or on warped and/or fragile mounts, on the platen. The unit/partner/contractor shall place only one physical item at a time on a surface appropriate for the item's size and format, except when scanning 35mm slides in a batch mode on a flatbed scanner. The unit/partner/contractor shall handle archival materials in bound volumes carefully and not force them open or place them face down. The unit/partner/contractor shall use book cradles to support volumes, and volumes shall be digitized in a face up orientation on cradles.
- f. The unit/partner/contractor shall not place objects such as books, papers, pens, and pencils on archival materials or their containers. The unit/partner/contractor shall not lean on, sit on, or otherwise apply pressure to archival materials or their containers. The unit/partner/contractor shall use only lead pencils as writing implements near archival materials or their containers. The unit/partner/contractor shall not write on or otherwise mark archival materials, jackets, or containers. The unit/partner/contractor shall not use Tacky finger, rubber fingers, or other materials to increase tackiness that may transfer residue to the records.
- g. The unit/partner/contractor shall not smoke, drink, or eat in the room where archival materials or their containers are located. The unit/partner/contractor shall not permit anyone to bring tobacco, liquids, and food into the room where archival materials or their containers are located.
- h. Unit/partner/contractor staff shall clean their hands prior to handling records and avoid the use of hand lotions before working with archival materials. Unit/partner/contractor staff shall wear clean white cotton gloves at all times when handling photographic film materials, such as negatives, color transparencies, aerial film, microfilm, etc. The unit/partner/contractor shall provide gloves. For some types of originals using cotton gloves can inhibit safe handling, such as when working with glass plate negatives.
- i. The unit/partner/contractor shall reinsert all photographic negatives, and other sheet film, removed from jackets in proper orientation with the emulsion side away from the seams. The unit/partner/contractor shall unwind roll film carefully and rewind roll film as soon as the digitizing is finished. The unit/partner/contractor shall rewind any rolls of film with the emulsion side in and with the head/start of the roll out.
- j. NARA custodial staff and Preservation Programs, NWT, shall have the right to inspect, without notice, the unit/partner/contractor work areas and digitizing procedures or to be present at all times when archival materials are being handled. Units/partners/contractors are encouraged to consult with Preservation

Programs, NWT, staff for clarification of these procedures or when any difficulties or problems arise.

#### **4. Training**

Training shall be provided by Preservation Programs, NWT, for archival material handling and certification of unit/partner/contractor staff prior to beginning any digitization. Any new unit/partner/contractor staff assigned to this project after the start date shall be trained and certified before handling archival materials.

## APPENDIX F: Resources

### Scope -

NARA 816 – *Digitization Activities for Enhanced Access*, at [http://www.nara-at-work.gov/nara\\_policies\\_and\\_guidance/directives/0800\\_series/nara816.html](http://www.nara-at-work.gov/nara_policies_and_guidance/directives/0800_series/nara816.html) (NARA internal link only)

Records Management information on the NARA website - [http://www.archives.gov/records\\_management/index.html](http://www.archives.gov/records_management/index.html) and [http://www.archives.gov/records\\_management/initiatives/erm\\_overview.html](http://www.archives.gov/records_management/initiatives/erm_overview.html)

Records transfer guidance for scanned textual documents - [http://www.archives.gov/records\\_management/initiatives/scanned\\_textual.html](http://www.archives.gov/records_management/initiatives/scanned_textual.html)

Records transfer guidance for scanned photographs and digital photography image files - [http://www.archives.gov/records\\_management/initiatives/digital\\_photo\\_records.html](http://www.archives.gov/records_management/initiatives/digital_photo_records.html)

Code of Federal Regulations at – <http://www.gpoaccess.gov/cfr/index.html>, see 36 CFR 1200

NARA’s Electronic Records Archive project - [http://www.archives.gov/electronic\\_records\\_archives/index.html](http://www.archives.gov/electronic_records_archives/index.html)

### Introduction -

#### General Resources –

*Moving Theory into Practice*, Cornell University Library, available at – <http://www.library.cornell.edu/preservation/tutorial/>

*HANDBOOK FOR DIGITAL PROJECTS: A Management Tool for Preservation and Access*, Northeast Document Conservation Center, available at – <http://www.nedcc.org/digital/dighome.htm>

*Guides to Quality in Visual Resource Imaging*, Digital Library Federation and Research Libraries Group, July 2000, available at - <http://www.rlg.org/visguides>

*The NINCH Guide to Good Practice in the Digital Representation and Management of Cultural Heritage Materials*, Humanities Advanced Technology and Information Institute, University of Glasgow, and National Initiative for a Networked Cultural Heritage, available at – <http://www.nyu.edu/its/humanities/ninchguide/index.html>

#### Project Management Outlines –

“NDLP Project Planning Checklist,” Library of Congress, available at <http://lcweb2.loc.gov/ammem/prjplan.html>

“Considerations for Project Management,” by Stephen Chapman, *HANDBOOK FOR DIGITAL PROJECTS: A Management Tool for Preservation and Access*, Northeast Document Conservation Center, available at – <http://www.nedcc.org/digital/dighome.htm>

“Planning an Imaging Project,” by Linda Serenson Colet, *Guides to Quality in Visual Resource Imaging*, Digital Library Federation and Research Libraries Group, available at - <http://www.rlg.org/visguides/visguide1.html>

Digitization resources, Colorado Digitization Program, available at <http://cdphheritage.org/resource/index.html>



## Metadata -

## Common Metadata Types -

Dublin Core Metadata Initiative -

<http://dublincore.org/usage/terms/dc/current-elements/>

NARA's Lifecycle Data Requirements Guide (LCDRG), January 2002, (NARA internal link only)

[http://www.nara-at-work.gov/archives\\_and\\_records\\_mgmt/archives\\_and\\_activities/accessioning\\_processing\\_description/lifecycle/](http://www.nara-at-work.gov/archives_and_records_mgmt/archives_and_activities/accessioning_processing_description/lifecycle/)

Official EAD site at the Library of Congress -

<http://lcweb.loc.gov/ead/>

Research Library Group's Best Practices Guidelines for EAD -

<http://www.rlg.org/rlgead/eadguides.html>

Harvard University Library's Digital Repository Services (DRS) User Manual for Data Loading, Version 2.04 -

[http://hul.harvard.edu/ois/systems/drs/drs\\_load\\_manual.pdf](http://hul.harvard.edu/ois/systems/drs/drs_load_manual.pdf)

Making of America 2 (MOA2) Digital Object Standard: Metadata, Content, and Encoding -

<http://www.cdlib.org/about/publications/CDLObjectStd-2001.pdf>

Dublin Core initiative for administrative metadata -

<http://metadata.net/admin/draft-iannella-admin-01.txt>

Data Dictionary for Administrative Metadata for Audio, Image, Text, and Video Content to Support the Revision of Extension Schemas for METS -

<http://lcweb.loc.gov/rr/mopic/avprot/extension2.html>

Metadata Encoding and Transmission Standard (METS) rights extension schema available at -

<http://www.loc.gov/standards/rights/METSRights.xsd>

Peter B. Hirtle, "Archives or Assets?" -

<http://techreports.library.cornell.edu:8081/Dienst/UI/1.0/Display/cul.lib/2003-2>

June M. Besek, Copyright Issues Relevant to the Creation of a Digital Archive: A Preliminary Assessment, January 2003 - <http://www.clir.org/pubs/reports/pub112/contents.html>

Adrienne Muir, "Copyright and Licensing for Digital Preservation," -

<http://www.cilip.org.uk/update/issues/jun03/article2june.html>

Karen Coyle, Rights Expression Languages, A Report to the Library of Congress, February 2004, available at - [http://www.loc.gov/standards/Coylereport\\_final1single.pdf](http://www.loc.gov/standards/Coylereport_final1single.pdf)

MPEG-21 Overview v.5 contains a discussion on intellectual property and rights at -

<http://www.chiariglione.org/mpeg/standards/mpeg-21/mpeg-21.htm>

Peter Hirtle, "When Works Pass Into the Public Domain in the United States: Copyright Term for Archivists and Librarians," -

[http://www.copyright.cornell.edu/training/Hirtle\\_Public\\_Domain.htm](http://www.copyright.cornell.edu/training/Hirtle_Public_Domain.htm)

Mary Minow, "Library Digitization Projects: Copyrighted Works that have Expired into the Public Domain" - <http://www.librarylaw.com/DigitizationTable.htm>

Mary Minow, *Library Digitization Projects and Copyright* -

<http://www.llrx.com/features/digitization.htm>

NISO Data Dictionary - Technical Metadata for Digital Still Images –  
[http://www.niso.org/standards/resources/Z39\\_87\\_trial\\_use.pdf](http://www.niso.org/standards/resources/Z39_87_trial_use.pdf).

Metadata for Images in XML (MIX) –  
<http://www.loc.gov/standards/mix/>

TIFF 6.0 Specification -  
<http://partners.adobe.com/asn/developer/pdfs/tn/TIFF6.pdf>

Digital Imaging Group's DIG 35 metadata element set –  
[http://www.i3a.org/i\\_dig35.html](http://www.i3a.org/i_dig35.html)

Harvard University Library's Administrative Metadata for Digital Still Images data dictionary -  
[http://hul.harvard.edu/ldi/resources/ImageMetadata\\_v2.pdf](http://hul.harvard.edu/ldi/resources/ImageMetadata_v2.pdf)

Research Library Group's "Automatic Exposure" Initiative -  
<http://www.rlg.org/longterm/autotechmetadata.html>

Global Digital Format Registry –  
<http://hul.harvard.edu/gdfr/>

Metadata Encoding Transmission Standard (METS) –  
<http://www.loc.gov/standards/mets/>

Flexible and Extensible Digital Object Repository Architecture (FEDORA) –  
<http://www.fedora.info/documents/master-spec-12.20.02.pdf>

Open Archival Information System (OAIS) -  
<http://ssdoo.gsfc.nasa.gov/nost/isoas/overview.html>

A Metadata Framework to Support the Preservation of Digital Objects -  
[http://www.oclc.org/research/projects/pmwg/pm\\_framework.pdf](http://www.oclc.org/research/projects/pmwg/pm_framework.pdf)

Preservation Metadata for Digital Objects: A Review of the State of the Art -  
[http://www.oclc.org/research/projects/pmwg/presmeta\\_wp.pdf](http://www.oclc.org/research/projects/pmwg/presmeta_wp.pdf)

PREMIS (Preservation Metadata Implementation Strategies) -  
<http://www.oclc.org/research/projects/pmwg/>

OCLC Digital Archive Metadata –  
[http://www.oclc.org/support/documentation/pdf/da\\_metadata\\_elements.pdf](http://www.oclc.org/support/documentation/pdf/da_metadata_elements.pdf)

Florida Center for Library Automation Preservation Metadata –  
[http://www.fcla.edu/digitalArchive/pdfs/Archive\\_data\\_dictionary20030703.pdf](http://www.fcla.edu/digitalArchive/pdfs/Archive_data_dictionary20030703.pdf)

Technical Metadata for the Long-Term Management of Digital Materials -  
[http://dvl.dtic.mil/metadata\\_guidelines/TechMetadata\\_26Mar02\\_1400.pdf](http://dvl.dtic.mil/metadata_guidelines/TechMetadata_26Mar02_1400.pdf)

National Library of New Zealand, Metadata Standard Framework, Preservation Metadata –  
[http://www.natlib.govt.nz/files/4initiatives\\_metaschema\\_revised.pdf](http://www.natlib.govt.nz/files/4initiatives_metaschema_revised.pdf)

National Library of Medicine Permanence Ratings -  
<http://www.nlm.nih.gov/pubs/reports/permanence.pdf>  
and <http://www.rlg.org/events/pres-2000/byrnes.html>

Design Criteria Standard for Electronic Records Management Software Applications (DOD 5015.2)  
<http://www.dtic.mil/whs/directives/corres/html/50152std.htm>

Assessment of Metadata Needs for Imaging Projects -

Guidelines for implementing Dublin Core in XML -  
<http://dublincore.org/documents/2002/09/09/dc-xml-guidelines/>

Adobe's Extensible Metadata Platform (XMP) –  
<http://www.adobe.com/products/xmp/main.html>

#### Local Implementation -

Adobe's Extensible Metadata Platform (XMP) –  
<http://www.adobe.com/products/xmp/main.html>

#### Technical Overview -

##### Glossaries of Technical Terms –

Technical Advisory Service for Images, available at –  
[http://www.tasi.ac.uk/glossary/glossary\\_technical.html](http://www.tasi.ac.uk/glossary/glossary_technical.html)

Kodak Digital Learning Center, available at –  
<http://www.kodak.com/US/en/digital/dlc/book4/chapter2/index.shtml>

##### Raster Image Characteristics –

“Introduction to Imaging (Getty Standards Program),” Getty Information Institute, available at –  
[http://www.getty.edu/research/conducting\\_research/standards/introimages/homepage.html](http://www.getty.edu/research/conducting_research/standards/introimages/homepage.html)

“Handbook for Digital Projects – Section VI Technical Primer,” by Steven Puglia, available at –  
<http://www.nedcc.org/digital/vi.htm>

##### Digitization Environment -

###### Standards-

ISO 3664 *Viewing Conditions- For Graphic Technology and Photography*

ISO 12646 *Graphic Technology – Displays for Colour Proofing – Characteristics and Viewing Conditions* (currently a draft international standard or DIS)

These standards can be purchased from ISO at <http://www.iso.ch> or from IHS Global at <http://global.ihs.com>.

“Digital Imaging Production Services at the Harvard College Library,” by Stephen Chapman and William Comstock, *DigiNews*, Vol. 4, No. 6, Dec. 15, 2000, available at <http://www.rlg.org/legacy/preserv/diginews/diginews4-6.html>

##### Quantifying Scanner/Digital Camera Performance –

###### Standards-

ISO 12231 Terminology

ISO 14524 Opto-electronic Conversion Function

ISO 12233 Resolution: Still Picture Cameras

ISO 16067-1 Resolution: Print Scanners

ISO 16067-2 Resolution: Film Scanners

ISO 15739 Noise: Still Picture Cameras

ISO 21550 Dynamic Range: Film Scanners

These standards can be purchased from ISO at <http://www.iso.ch> or from IHS Global at <http://global.ihs.com>.

“Debunking Specsmanship” by Don Williams, Eastman Kodak, in *RLG DigiNews*, Vol. 7, No. 1, Feb. 15, 2003, available at [http://www.rlg.org/preserv/diginews/v7\\_v1\\_feature1.html](http://www.rlg.org/preserv/diginews/v7_v1_feature1.html)

"Image Quality Metrics" by Don Williams, Eastman Kodak, in *RLG DigiNews*, Vol. 4, No. 4, Aug. 15, 2000, available at <http://www.rlg.org/legacy/preserv/diginews/diginews4-4.html#technical1>

"What is an MTF...and Why Should You Care" by Don Williams, Eastman Kodak, in *RLG DigiNews*, Vol. 2, No. 1, Feb. 15, 1998, available at <http://www.rlg.org/preserv/diginews/diginews21.html#technical>

*Guides to Quality in Visual Resource Imaging*, Digital Library Federation and Research Libraries Group, July 2000, available at <http://www.rlg.org/visguides>

- Guide 2 – "Selecting a Scanner" by Don Williams
- Guide 3 – "Imaging Systems: the Range of Factors Affecting Image Quality" by Donald D'Amato
- Guide 4 – "Measuring Quality of Digital Masters" by Franziska Frey

"Digital Imaging for Photographic Collections" by Franziska Frey and James Reilly, Image Permanence Institute, 1999, available at - [http://www.rit.edu/~661www1/sub\\_pages/digibook.pdf](http://www.rit.edu/~661www1/sub_pages/digibook.pdf)

"Image Capture Beyond 24-bit RGB" by Donald Brown, Eastman Kodak, in *RLG DigiNews*, Vol. 3, No. 5, Oct. 15, 1999, available at <http://www.rlg.org/preserv/diginews/diginews3-5.html#technical>

#### Color Management –

*Real World Color Management*, by Bruce Fraser, Chris Murphy, and Fred Bunting, Peachpit Press, Berkeley, CA, 2003 – <http://www.peachpit.com>

#### Image Processing Workflow –

*Adobe Photoshop CS Studio Techniques*, by Ben Wilmore, Adobe Press, Berkeley, CA, 2004 – <http://www.digitalmastery.com/book> or <http://www.adobepress.com>

"Digital Imaging for Photographic Collections" by Franziska Frey and James Reilly, Image Permanence Institute, 1999, available at - [http://www.rit.edu/~661www1/sub\\_pages/digibook.pdf](http://www.rit.edu/~661www1/sub_pages/digibook.pdf)

#### Digitization in Production Environments –

"Imaging Production Systems at Corbis Corporation," by Sabine Süssstrunk, *DigiNews*, Vol. 2, No. 4, August 15, 1998, available at <http://www.rlg.org/legacy/preserv/diginews/diginews2-4.html>

"Imaging Pictorial Collections at the Library of Congress," by John Stokes, *DigiNews*, Vol. 3, No. 2, April 15, 1999, available at <http://www.rlg.org/legacy/preserv/diginews/diginews3-2.html>

### Digitization Specifications for Record Types -

#### Imaging guidelines -

*Benchmark for Faithful Digital Reproductions of Monographs and Serials*, Digital Library Federation, available at <http://www.diglib.org/standards/bmarkfin.htm>

"Managing Text Digitisation," by Stephen Chapman. *Online Information Review*, Volume 27, Number 1, 2003, pp. 17-27. Available for purchase at: <http://www.emeraldinsight.com/1468-4527.htm>

Library of Congress - <http://memory.loc.gov/ammem/techdocs/index.html> and <http://lcweb2.loc.gov/ammem/formats.html>

Colorado Digitization Program - [http://www.cdphheritage.org/resource/scanning/documents/WSDIBP\\_v1.pdf](http://www.cdphheritage.org/resource/scanning/documents/WSDIBP_v1.pdf)

California Digital Library - <http://www.cdlib.org/news/pdf/CDLImageStd-2001.pdf>

Imaging Techniques –

Copying and Duplicating: Photographic and Digital Imaging Techniques, Kodak Publication M-1, CAT No. E152 7969, Sterling Publishing, 1996.

Storage and Digital Preservation -

*Digital Preservation & Ensured Long-term Access*, Research Libraries Group, available at – [http://www.rlg.org/en/page.php?Page\\_ID=552](http://www.rlg.org/en/page.php?Page_ID=552)

Digital Preservation Coalition, available at – <http://www.dpconline.org/graphics/index.html>

Preserving Access to Digital Information, available at – <http://www.nla.gov.au/padi/>

National Digital Information Infrastructure and Preservation Program, Library of Congress, available at – <http://www.digitalpreservation.gov/>

NARA's Electronic Records Archive project - [http://www.archives.gov/electronic\\_records\\_archives/index.html](http://www.archives.gov/electronic_records_archives/index.html)

*Digital Preservation*, Digital Library Federation, available at – <http://www.diglib.org/preserve.htm>

Open Archival Information System, available at – <http://ssdoo.gsfc.nasa.gov/nost/isoas/overview.html>

"Trusted Digital Repositories: Attributes and Responsibilities," Research Libraries Group and OCLC, May 2002, available at [www.rlg.org/longterm/repositories.pdf](http://www.rlg.org/longterm/repositories.pdf)

Quality Control, Testing Results, and Acceptance/Rejection

*Conversion Specifications*, American Memory, Library of Congress, available at – <http://memory.loc.gov/ammem/techdocs/conversion.html>

"NDLP Project Planning Checklist," Library of Congress, available at <http://lcweb2.loc.gov/ammem/prjplan.html>



The National Archives

**Digital  
Preservation  
Guidance  
Note:**

**1**

## **Selecting File Formats for Long-Term Preservation**

## Document Control

Author: Adrian Brown, Head of Digital Preservation Research

Document  
Reference: DPGN-01

Issue: 2

Issue Date: August 2008



## Contents

<b>1</b>	<b>INTRODUCTION</b> .....	<b>4</b>
<b>2</b>	<b>SELECTION ISSUES</b> .....	<b>4</b>
2.1	Ubiquity .....	6
2.2	Support.....	6
2.3	Disclosure .....	6
2.4	Documentation quality .....	6
2.5	Stability.....	6
2.6	Ease of identification and validation .....	6
2.7	Intellectual Property Rights .....	7
2.8	Metadata Support.....	7
2.9	Complexity .....	7
2.10	Interoperability.....	7
2.11	Viability.....	8
2.12	Re-usability .....	8
<b>3</b>	<b>EVALUATING FORMATS: SOURCES OF INFORMATION</b> .....	<b>8</b>
3.1	Ubiquity .....	8
3.2	Support.....	8
3.3	Disclosure .....	9
3.4	Documentation quality .....	9
3.5	Stability.....	9
3.6	Ease of identification and validation .....	9
3.7	Intellectual Property Rights .....	9
3.8	Metadata Support.....	9
3.9	Complexity .....	10
3.10	Interoperability.....	10
3.11	Viability.....	10
3.12	Re-usability .....	10
<b>4</b>	<b>CONCLUSION</b> .....	<b>10</b>

## 1 Introduction

This document is one of a series of guidance notes produced by The National Archives, giving general advice on issues relating to the preservation and management of electronic records. It is intended for use by anyone involved in the creation of electronic records that may need to be preserved over the long term, as well as by those responsible for preservation.

This guidance note provides information for the creators and managers of electronic records about file format selection. Please note that The National Archives does not specify or require the use of any particular file formats for records which are to be transferred. Choice of file format should always be determined by the functional requirements of the record creation process. Record creators should be aware however, that long-term sustainability will become a requirement, both for ongoing business purposes and archival preservation. Sustainability costs are inevitably minimised when this factor is taken into account prior to data creation. Failure to do so often makes later attempts to bring electronic records into a managed and sustainable regime an expensive, complex and, generally, less successful process.

This guidance note sets out a range of criteria the aim of which is to help data creators and archivists make informed choices about file format issues.

## 2 Selection issues

File formats encode information into forms that can only be processed and rendered comprehensible by very specific combinations of hardware and software. The accessibility of that information is therefore highly vulnerable in today's rapidly evolving technological environment. This issue is not solely the concern of digital archivists, but of all those responsible for managing and sustaining access to electronic records over even relatively short timescales.

The selection of file formats for creating electronic records should therefore be determined not only by the immediate and obvious requirements of the situation, but also with long-term sustainability in mind. An electronic record is not fully fit-for-purpose unless it is sustainable throughout its required life cycle.













The practicality of managing large collections of electronic records, whether in a business or archival context, is greatly simplified by minimising the number of separate file formats involved. It is useful to identify a minimal set of formats which meet both the active business needs and the sustainability criteria below, and restrict data creation to these formats.

This guidance note is primarily concerned with the selection of file formats for data *creation*, rather than the conversion of existing data into 'archival' formats. However, the criteria described are equally applicable to the latter.

Selecting file formats for migration introduces some additional issues. Formats for migration must meet the requirements for both preservation of authenticity and ease of access. For example, the data elements of a word-processed document could be preserved as plain ASCII text, together with any illustrations as separate image files. However, this would result in a loss of structure (e.g. the formatting of the text), and of some context (e.g. the internal pointers to the illustrations).

There is also a subtly different conflict between the need for data formats that can be accessed and those that can be re-used. From a preservation and re-use perspective, data must be maintained in a form that can be processed. For the purposes of access, however, control of the formatting may well be the most important criteria, and in some cases it may be desirable for the data not to be able to be processed by end users. In some cases it may only be possible to reconcile these differences by using different formats for preservation and presentation purposes.

The following criteria should be considered by data creators when selecting file formats:

-  Ubiquity
-  Support
-  Disclosure
-  Documentation quality
-  Stability
-  Ease of identification and validation
-  Intellectual Property Rights
-  Metadata Support
-  Complexity
-  Interoperability
-  Viability
-  Re-usability

These criteria are elaborated in the following sections:

## 2.1 Ubiquity

The laws of supply and demand dictate that formats which are well established and in widespread use will tend to have broader and longer-lasting support from software suppliers than those that have a niche market. There is also likely to be more comprehensive community support amongst users. Popular formats are therefore preferable in many cases.

## 2.2 Support

The extent of current software support is a major factor for consideration. The availability of a wide range of supporting software tools removes dependence on any single supplier for access, and is therefore preferable. In some cases however, this may be counterbalanced by the ubiquity of a single software tool.

## 2.3 Disclosure

Those responsible for the management and long-term preservation of electronic records require access to detailed technical information about the file formats used. Formats that have technical specifications available in the public domain are recommended. This is invariably the case with open standards, such as JPEG. The developers of proprietary formats may also publish their specifications, either freely (for example, PDF), or commercially (as is the case with the Adobe Photoshop format specification, which is included as part of the Photoshop Software Development Kit). The advantages of some open formats may come at the cost of some loss in structure, context, and functionality (e.g. ASCII), or the preservation of formatting at the cost some reusability (e.g. PDF). Proprietary formats frequently support features of their creating software, which open formats do not. The tension between these needs is sometimes unavoidable, although the range and sophistication of open formats is increasing all the time. The use of open standard formats is however highly recommended wherever possible.

## 2.4 Documentation quality

The availability of format documentation is not, in itself, sufficient; documentation must also be comprehensive, accurate and comprehensible. Specifically, it should be of sufficient quality to allow interpretation of objects in the format, either by a human user or through the development of new access software.

## 2.5 Stability

The format specification should be stable and not subject to constant or major changes over time. New versions of the format should also be backwards compatible.

## 2.6 Ease of identification and validation

The ability to accurately identify the format of a data file and confirm that it is a valid example of that format, is vital to continued use. Well-designed formats facilitate identification through the use of 'magic numbers' and version information within the file structure. The availability of tools to validate the format is also a consideration.

## 2.7 Intellectual Property Rights

Formats may utilise technologies encumbered by patents or other intellectual property constraints, such as image compression algorithms. This may limit present or future use of objects in that format. In particular, 'submarine patents' (when previously undisclosed patent claims emerge), can be a concern. Formats that are unencumbered by patents are recommended.

## 2.8 Metadata Support

Some file formats make provision for the inclusion of metadata. This metadata may be generated automatically by the creating application, entered by the user, or a combination of both. This metadata can have enormous value both during the active use of the data and for long-term preservation, where it can provide information on both the provenance and technical characteristics of the data. For example, a TIFF file may include metadata fields to record details such as the make and model of scanner, the software and operating system used, the name of the creator, and a description of the image. Similarly, Microsoft Word documents can include a range of metadata to support document workflow and version control, within the document properties. The value of such metadata will depend upon:

- The degree of support provided by the software environment used to create the files,
- The extent to which externally stored metadata is used in its place. (For example if records are stored within an Electronic Records Management System).

In general, formats that offer metadata support are preferable to those that do not.

## 2.9 Complexity

Formats should be selected for use on the basis that they support the full range of features and functionality required for their designated purpose. It is equally important, however to avoid choosing over-specified formats. Generally speaking the more complex the format, the more costly it will be to both manage and preserve.

## 2.10 Interoperability

The ability to exchange electronic records with other users and IT systems is also an important consideration. Formats that are supported by a wide range of software or are platform-independent are most desirable. This also tends to

support long-term sustainability of data by facilitating migration from one technical environment to another.

## 2.11 Viability

Some formats provide error-detection facilities, to allow detection of file corruption that may have occurred during transmission. Many formats include a CRC (Cyclic Redundancy Check) value for this purpose, but more sophisticated techniques are also used. For example, the PNG format incorporates byte sequences to check for three specific types of error that could be introduced. Formats that provide facilities such as these are more robust, and thus preferable.

## 2.12 Re-usability

Certain types of data must retain the ability to be processed if they are to have any re-use value. For example, conversion of a spreadsheet into PDF format effectively removes much of its ability to be processed. The requirement to maintain a version of the record that can be processed must also be considered.

# 3 Evaluating formats: sources of information

A variety of practical information sources are available to support the evaluation of formats in accordance with these criteria. PRONOM, The National Archives' technical registry, is particularly designed as an impartial and authoritative source of advice on this subject, and is freely available online at [www.nationalarchives.gov.uk/pronom/](http://www.nationalarchives.gov.uk/pronom/)

The following sections indicate sources for evaluating formats:

## 3.1 Ubiquity

The relative popularity of a format tends to be a comparatively subjective measure, but is likely to be widely known within a particular user community.

## 3.2 Support

This requires consideration of the number of software tools which currently support the format, and of the ubiquity of those tools. In PRONOM, the level of software support for a format may be assessed using the 'Compatible software' search facility on the 'File format' tab. This will return a list of software known to support a given format. This can be supplemented by additional research, as PRONOM may not provide comprehensive coverage for all formats. In addition, this factor must be considered in conjunction with the ubiquity of the format (see 3.1).

### **3.3 Disclosure**

In PRONOM, the degree of disclosure may be ascertained from the 'Availability' field on the 'Documentation' tab of a format record.

### **3.4 Documentation quality**

PRONOM provides links to known documentation that is available for a format. In PRONOM, an initial assessment of the comprehensiveness of available documentation may be gained from the 'Disclosure' field on the 'Summary' tab of a format record. The authoritativeness may be ascertained from the 'Type' field on the 'Documentation' tab of a format record. A detailed judgement of documentation quality will require evaluation of the documentation itself.

### **3.5 Stability**

The stability of a format may be judged by its age, and the frequency with which new versions are released. The number of versions of a format may be determined in PRONOM by searching on the format name: all known versions of the format will be listed.

PRONOM also records the dates on which versions of formats were released and withdrawn from current support – these may be used to judge the longevity of each format version.

### **3.6 Ease of identification and validation**

In PRONOM, the availability of existing identification and validation tools for a format may be determined by using the 'Compatible software' search facility on the 'File format' tab. The search can then be filtered by software which can 'identify' or 'validate' a given format respectively.

### **3.7 Intellectual Property Rights**

In PRONOM, known IPR restrictions for a format will be listed under the 'Rights' tab of any format record.

### **3.8 Metadata Support**

Determining the degree of metadata support offered by a format may require a review of its technical documentation. PRONOM may be of assistance for locating such documentation (see 3.4).



### **3.9 Complexity**

Complexity is a subjective measure, and can generally only be determined with reference to the relevant technical documentation. PRONOM may be of assistance for locating such documentation (see 3.4).

### **3.10 Interoperability**

In PRONOM, the general level of interoperability for a given format may be judged by reviewing the number of software products which are available to create or render files in that format (see 3.2).

### **3.11 Viability**

In PRONOM, the provision of error detection and correction mechanisms may be noted in the 'Description' field of the format record. Otherwise, this will need to be determined with reference to the relevant technical documentation. PRONOM may be of assistance for locating such documentation (see 3.4).

### **3.12 Re-usability**

Re-usability is a complex measure and will vary depending on the requirements of a particular community of users. It can generally only be determined with reference to the relevant technical documentation. PRONOM may be of assistance for locating such documentation (see 3.4).

## **4 Conclusion**

There are many issues to be considered when selecting file formats extending beyond the immediate and obvious requirements of the situation. It may not be possible to select formats that meet all criteria in every case; however, new formats and revisions of existing formats are constantly being developed. This guidance note should assist data creators to make informed decisions about file format selection from the ever-changing choices available.

The adoption of sustainable file formats for electronic records brings benefits to data creators, data managers and digital archivists. Selection decisions informed by the criteria described above will greatly enhance the sustainability of the records created.

*Digitization of books in the  
National Library*

*– methodology and lessons learned*

**National Library of Norway**

**September 2007**

# 1. The digital national library

The vision of the National Library of Norway is to be a living memory bank, by being a "Multimedia Centre of Knowledge" with focus not only on preservation but also on mediation.

To succeed with this ambition, one of our main goals is to be a digital national library, as the core of a Norwegian digital library. A digital National Library is just another way of being a National Library. It is then important to have as much digital material as possible, not only historical material but also the modern part of the cultural heritage, to give access to as much material as possible to as many as possible whenever required.

The Norwegian National Library has therefore started a systematic digitisation of the entire collection. Based on a modern Legal Deposit Act we receive everything that is produced and of interest to the public, be it books, newspapers, periodicals, photos, films, music or broadcasting. All broadcasters with a license in Norway may be asked to deliver copies of their programmes for preservation, and we have an extensive collaboration with the national broadcaster NRK about preservation and dissemination. We are also to preserve those digital signals that are never converted into anything else before they reach the user. Starting in 2005 we have harvested large parts of the Norwegian web domain .no.

Some of this material is delivered to us in digital formats, and we are about to have a relatively large digital collection when it comes to the audio-visual, but little when it comes to printed material.

The issue now is how we shall take advantage of having established a digital repository for preservation in order also to give access to this material both to scholars, students and the public. This is a challenge in many aspects, but mainly with regard to copyright, technical quality and dissemination pedagogic.

Our aim to establish a unified Digital Library for all our different media, with our Digital Long Term Repository as a basis, will require the following strategies:

- As soon as possible (depending on funding, of course) to digitize all our collections. We do this systematically
- To digitize on demand
- To negotiate with the publishers to get as much material as possible deposited in digital formats
- To be a digital archive for other institutions, e.g. publishers and newspapers
- To find strategic partners to cooperate with, financially and on know-how
- To be a trusted repository for digital material in the Norwegian society
- To give access to as much as possible of our cultural heritage, e.g. through search engines
- To negotiate with the right-holders to give access to material that is not yet in the free, so that also modern books, films, and music can be searched

To be able to do this, it is also important to take part in discussions and developing the pedagogic of the net. Libraries' knowledge about the users' needs is a good experience basis on which to develop search methodology in the crossing between metadata and the methodology of search engines. Our aim is to give access to information, knowledge and experience on a given topic across media types.

## 2. Digitizing the Norwegian cultural heritage

### Scope

For more than 10 years the National Library has been digitizing a wide range of media. The main focus for this digitization has been photographs, sound recordings and microfilm (newspapers). As a result, the digital collection today contains more than 150,000 hours of radio, more than 300,000 photographs and more than 1,000,000 newspaper pages. In addition, we have digitized more than 25,000 books over the last year.

Still, these are modest numbers compared to what we have now set out to accomplish. If we include the estimated growth of the collections through analogue legal deposit until digital legal deposit is up and running, we expect the amount of material to be digitized in the National Library digitization programme to be:

- 450,000 books
- 2,000,000 periodicals
- 4,700,000 newspapers (more than 60,000,000 pages)
- 1,300,000 pictures (photographs and postcards)
  - 60,000 posters
  - 200,000 maps
- 4,000,000 manuscripts
  - 200,000 units of sheet music
- 1,900,000 leaflets
- 1,000,000 hours of radio
  - 80,000 hours of music
- 250,000 hours of film and television

Gifts, purchases or deposited material during the program period will add to these numbers.

In order to accomplish such an ambitious program of digitization during a foreseeable period of time, the digitization activity has been greatly increased. As much as possible, we are now streamlining the process from when an object is selected for digitization until it is placed digitally in our digital long term repository, simultaneously offering web access for authorized users. With the multimedia collection of the National Library this poses special challenges, as we need to establish separate production lines for the digitization of different types of material. In addition, there are usually several variants within each type of material, which again demand different adjustments.

Putting these things in place is a challenge, both in terms of technology, logistics, organization, manpower and financing.

### ***Legal Rights***

The National Library has a legal right to digitize the collection for preservation. However, to make the collection available in the digital library, it is necessary to make agreements with the copyright holders whenever the material is still under copyright protection.

The old material with no copyright restrictions will be made available for everyone in the digital library.

The Legal Deposit Act states that every object subject to legal deposit may be made available for research and documentation. This is also valid for the digital material subject to legal deposit. However, in the digital domain there are still several unresolved challenges, e.g. privacy questions related to the interconnection of several extensive data sources (e.g. Norwegian internet pages), and the risk and consequences related to misuse through illegal copying of digital content, which are much more extensive than for paper based documents. These are the reasons why we still do not have good solutions for access to digital legal deposited material via the Internet.

When the documents are published in a traditional way but deposited in digital formats, we have to make agreements with the copyright holders to be able to give access to the digital documents. In this case we try to get agreements giving us at least the same rights that we have for material subject to legal deposit.

The National Library also works to get agreements with the copyright holders making it possible to give a broader access to the modern part of the cultural heritage. An example is the [agreement](#) between the National Library and several copyright holders organisations on giving access to books and journal articles related to "The High North". This agreement gives the National Library rights to make available approximately 1 400 complete works in the digital library. These works will be used to evaluate user behaviour and the frequency of use of the digital material. The result of the evaluation will form basis to negotiate on more permanent agreements with the copyright holders.

## ***Financing***

Preliminary estimates suggest a total cost of around 1 billion NOK for the National Library's digitization programme. Around 60% of the cost is for the digital storage, the purchase of digitization equipment and software, the development and integration of systems that will be part of the process of digitization and post-processing, in addition to wages for carrying out the digitization and money for some external commissions. The remaining 40% of the cost will be needed for the indexing of the material in order to establish the metadata required for retrieval, and for fetching the material from the collections, for the necessary conservation, and for the return of the material to the collections.

Through a re-channelling of activities towards the strategic initiatives, the National Library has prioritized 12 M NOK a year for the digitization programme. In addition, we have received a grant from the Ministry of 3 M NOK for digitization in 2007. Finally, on top of this, we have a budget of 13 M NOK for the purchase of digital storage in 2007, so the total budget for the initiative is 28 M NOK this year.

Preliminary estimates show that the whole digitization programme can be carried out in 15 - 20 years. Naturally, there is considerable uncertainty associated with an extensive and long-term effort such as this. We expect developments in technology both in terms of digital storage, digitization equipment and of tools that will enable retrieval in a way that reduces the need for manual indexing. However, it is difficult to accurately predict the consequences of such a development, in terms of both cost and increased efficiency.

## ***Status***

Large parts of 2006 were spent on carrying out the process of inviting tenders for the purchase of digitization equipment, software, digital storage and other ICT solutions required. In addition, a project manager and some system developers were hired late 2006.

In the new digitization programme, books were given priority. A closer scrutiny of this work can be found in a separate chapter.

Following an invitation for tenders we have outsourced the digitization of microfilm to a German company (CCS). So far 400,000 pages from the newspapers *Aftenposten* and *Adresseavisen* have been digitized. However, this number will be increased in the fall of 2007.

In addition, the digitization of photos has been made more efficient through the purchase of software for automatic post processing of digitized material, and through an upgrade of the digital cameras used for this work. We have also bought efficient equipment for the digitization of 35mm roll film negatives.

We are also working on an upgrade of equipment for the digitization of music, radio and moving images. An important part of this is a more automated connection between the digitized material and metadata that describes the material.

Other types of material will be digitized on demand, but so far we have not established adequate production lines that will enable speedy transfer of digitized materials into our long term digital repository, and immediate availability in our digital library.

The systematic digitization is the foundation of the digitizing process. In addition, we will carry out on-demand digitization which can be both user initiated and based on the strategic priorities of the National Library. On-demand digitization will be given priority over the systematic work. This applies to all parts of the digitization process. However, at this time it is only true for those types of material for which production lines have already been established (books, photography and sound).

### ***Future plans***

Of most importance in the short term is to make operative all the production lines for the digitization of books. What remains to be done is mainly to establish a regime for quality assurance of the digitized books, and an adequate handling of exceptions through all of the production line. This includes all stages of the process starting with ordering material out from the stacks and ending with the digital versions of the material residing in the digital long term repository and becoming accessible via the NB digital library.

We have also started work on establishing production lines to preserve both digitized newspaper pages and digitally deposited newspapers in the digital long term repository, and to make them available in our digital library. In addition to the newspapers that have been digitized from microfilm, we have established a pilot of daily legal deposit from two leading national newspapers in preservation quality PDF format. When this pilot gets into regular operations, we will work towards extending it to more newspaper titles.

We also plan to start the work of establishing a production line for the digitization of manuscripts. This is a type of material in heavy demand from our users and which will be of relevance in connection with upcoming authors' anniversaries in the years ahead. For this material, in addition to digitization, we must establish the necessary metadata in order to achieve satisfactory search and retrieval in the digital library.

Later again, we will establish production lines for the digitization of periodicals and posters. In addition, we plan to automate parts of the existing production line for the digitization of photography, in order to increase efficiency.

Another very important activity will be to initiate legal deposit of new publications in digital formats of preservation quality. Besides the pilot setup with daily legal deposit from two newspapers, today we have an operative solution for the digital deposit of radio in preservation quality. Also, audio books produced in digital format are delivered from the Norwegian Library of Talking Books and Braille.

We also have an ongoing dialogue with publishers and TV broadcasters, planning to start digital deposit of books, periodicals and television during the coming year.

### **3. Establishing a production line for the digitization of books**

#### ***3.1 Starting out***

In establishing a brand new process for the digitization of books, it was important to achieve a well integrated production line capable of covering all the steps through which the book would pass before it was preserved in the National Library's digital long term repository, and also becoming available to authorized users in the digital library. We wanted to automate as much as possible of the data flow and processing of the digital book, at the same time leaving enough flexibility in order to accommodate new process stages in the production line if the need should arise.

The processes we wanted to include were: selection for digitization and ordering of the books in question, fetching of material from the stacks, transport, extraction of metadata from the catalogue, digitization, OCR treatment and structural analysis, format conversion, the generation of preservation objects, ingest to the digital long term repository, notifying the catalogue of the digital object, and indexing OCR text and metadata in our search engine.

We also saw that different types of scanner technology gave very different efficiency in the digitization. If books could be dismounted, the scanning would become at least 10 times faster.

#### ***3.2 Fundamental decisions***

##### **Preservation – quality and format decisions**

The digitization programme is a part of the National Library's strategy for preservation and dissemination. Digitization will make preservation of the collection more efficient and less vulnerable to physical deterioration. This means that the digitization must be performed in a quality so high that, after digital preservation, it must be possible to satisfactorily recreate the properties of the original at a later time. At the same time, the digitization and the chosen formats must fulfil the requirements for dissemination of the material.

We have chosen to digitize books at a resolution of 400 dpi and a colour depth of 24 bits. Our preservation format is JPEG2000 with lossless compression. The preserved image is not processed or reduced in any way after the scanning process.

By choosing losslessly compressed JPEG2000 instead of uncompressed TIFF as a preservation format, we will reduce the need for digital storage by about 50%. For the whole digitization programme this means savings on the order of NOK 70 M. Through practical tests we have demonstrated that we are able to convert from the JPEG2000 format back to uncompressed TIFF with absolutely no loss of information.



An argument against using JPEG2000 is that one bit error in a JPEG2000 image file will be able to destroy the whole image, whereas a bit error in an uncompressed TIFF image file will not affect more than one pixel. With the storage policy in our digital long term repository, we find the risk of bit errors to be negligible.

The quality requirements for preservation are far stricter than those normally used for dissemination. Also, the formats used for dissemination have a shorter lifespan, partially because new formats are developed with more advanced compression algorithms that offer better quality with less data, and partially because there are new versions developed of the existing formats, with better algorithms and better quality. Therefore, we have chosen to generate the dissemination format from the preservation file at the moment a user asks for the image. Using this strategy we will easily be able to switch dissemination formats by replacing the algorithm that generates the dissemination format.

In today's solution we generate a JPEG file of the desired quality for viewing (typically around 200 Kbytes) from the JPEG2000 file located in the digital long term repository (typically around 20 Mbytes).

### **Perform our own digitization or hire others?**

Today, there are several organizations offering cultural institutions to digitize their book collections for free or at a very low price, in return asking for the right to store the digital copies and to offer search and display of the books. Examples of this are Google and the Internet Archive. In the National Library's book collection there are relatively few books which are no longer protected by copyright, and to which free access accordingly can be offered. Out of the 450,000 titles that will be digitized under the digitization programme, at present only around 5,000 titles are no longer copyrighted. It is important to the National Library that books still under copyright will only be stored in the National Library's digital long term repository. Access to such books will only be given in accordance with agreements made with the right holders. Also, it has been a fundamental principle for the National Library that other service providers must be given equal opportunity to offer services based on our collections. The sum of these considerations has meant that we have not deemed it natural to cooperate with this type of organizations for digitization. Still, we will invite them to disseminate what we have digitized, on par with other actors.

It has also been part of the picture that the National Library has been able to reassign existing staff to tasks associated with the digitization production line. This has meant that the total cost of in-house digitization of books is lower than it would have been for outsourcing the digitization.

For other types of material we have chosen to outsource digitization. For instance, this applies to digitization, OCR and structural analysis of newspapers on microfilm.

### **Dismounting books**

In order to achieve efficient digitization, we have chosen to dismount books for digitization if we have at least three copies in our repository library. The dismounted copy is then thrown out after digitization.

When we have fewer copies, the books are scanned manually, with operators opening the books and scanning two pages at a time. The most vulnerable books are scanned under the supervision of a conservator, and any necessary conservation measures are performed before or in connection with digitization.

The process of preparing books for dismounted scanning, is more labour intensive than preparing books for manual scanning. Special operators are required for deconstructing the books (separating the binding from the rest of the book, removing glue with a hydraulic cutter), and the scanning of the binding is a separate process. Because of this, 4 operators are needed in order to feed one scanner of dismounted books. Still, the total picture means lower cost and higher production than if the same resources had been used for manual scanning.

As of today, around a quarter of the National Library's book collection can be dismounted for digitization. In order to improve this ratio, the National Library plans to invite Norwegian libraries to contribute copies of books we have too few of. In this way, book digitization can be carried out faster and more efficiently in the National Library, and the libraries will be able to free up space in their stacks.

We have not yet tried out scanners that automatically turn the pages of whole books that have not been dismounted. Such scanners are developing rapidly, and this technology is becoming very interesting. For books that can not be dismounted this technology may offer considerably higher production per operator, both because the scanners are fast and treat the material gently, and because one operator can serve several such scanners at the same time. However, investment costs for this type of scanner are still high.

## **OCR and structural analysis**

In order to allow for full text search, all digitized books undergo a process of optical character recognition (OCR). In the regular production this process is fully automated, and there is no manual quality control or correction phase. The text that results from the OCR treatment is indexed in our search engine together with the metadata. If a search gives results from the text, the page of the book where the text was found will be displayed, and the user can browse the book from that page.

Also, an automated structural analysis is performed, during which any table of contents is annotated, and page numbers in the book are verified so that the user interface relates to the actual pagination of the book. This is also an automatic process. The software allows for very advanced structural analysis, but at this stage it is not feasible to increase complexity without also applying extensive manual quality control and post control. For selected parts of the collection we will perform more advanced structural analysis including annotation of several parts of the documents, again allowing for more advanced navigation of the books in the user interface.

At present, about 2,000 – 3,000 books are digitized every month in the National Library. With this volume it is not possible in practice to perform manual post control of the OCR and structural treatment.

Both OCR and structural analysis are performed using software called docWorks.

## **The National Library's digital long term repository**

The digital long term repository is an infrastructure for the long term preservation of digital objects. Everything that is digitized as part of the National Library's digitization programme is to be preserved as digital objects in the National Library's digital long term repository.

The digital long term repository separates the use of digital content from the technology which is employed for the storage. This allows for easy migration to new generations of storage technology without affecting the systems for retrieval of the digital content. This is very important in a 1000 year perspective.

All digital content is stored in three copies on two separate storage media in the digital long term repository. At present one copy is stored on disk while two are on tape.

## **Search engine**

In order to realize search across large aggregations of data, the National Library has chosen to employ search engine technology rather than traditional data base solutions. Both metadata and full text are indexed by the search engine, and searches are performed without regard to types of material. We have also implemented a so-called drill-down search in the metadata. Metadata for objects satisfying the search criteria is analyzed in real time during searches, and alternate paths of navigation and different ways of narrowing the search results are built and displayed to the user.

The search engine used is delivered by FAST.

## **Authentication and authorization – access control**

The National Library has chosen to employ role based access control. We have chosen to cooperate with FEIDE, which is the national infrastructure for authentication and authorization of users at Norwegian universities and other institutions of higher education. This means that we can open for access to defined groups of scientists at the universities and institutions, without knowing every individual. The universities are responsible for the authentication of those who satisfy the requirements for the different roles which have been defined within the system.

If we had picked an access control based on user names and passwords with associated rights for individuals, the National Library would have needed to spend considerable resources on the administration of users and the maintenance of access rights.

At present we have not implemented this access control solution for digital books. Therefore, we have only indexed and made available books that are either out of copyright, or covered by agreements with the right-holders on giving open access.

The software used for user authorization and authentication was supplied by SUN Microsystems.

## ***3.3 Production lines***

### **Prioritization**

The basis for the digitization is the systematic selection. We have chosen to start with the oldest material in order to quickly get the material which is out of copyright into our digital library. In addition to the systematic selection we give priority to material on the basis of internal needs and external requests. Especially prioritized material is placed at the head of the queue, in front of the systematic selection.

A special case is the agreement between the National Library and several rights organizations regarding books and articles related to the "High North". These works have been given special priority in the digitization process.

### **Ordering and extraction from the stacks**

In order to achieve an efficient extraction of material for digitization, a special function has been implemented in BIBSYS, our book cataloguing system. Here we can order a given number of titles for dismounting, chosen automatically by the system among titles that there are a sufficient number

of copies of, starting with the oldest. In addition, we can order single titles to be given special priority (both on extraction from the stacks and through the whole production line). Adaptations have also been made to the software that runs our automatic storage system for books, so that the operators can give top priority to remote loans, and then extract books for digitization. This system is integrated with the catalogue, so that books ordered for digitization automatically appear in the interface for the operators of the automatic storage system.

Already we have spent more than one man-year on system adaptations of the catalogue system and the software for the automatic storage system.

## **Digitization**

For the books to be dismantled we have two hydraulic cutters, three binding scanners (i2s Copibook) and two auto-feed scanners (Agfa S655). For the page-turner scanning we use i2s Digibook Suprascan. Five of these are A2 scanners for normal page-turner scanning and one is an A0 scanner for special material. The A0 scanner is operated by conservators.

Before the bindings are scanned, all metadata for the book is retrieved from the catalogue (BIBSYS) by way of a bar code assigned to every book in BIBSYS. A digital ID for the book is generated and inserted into an XML file together with the metadata obtained from the catalogue.

In the case of autoscan, a sheet of paper with the bar code of the digital ID is printed out after the scanning of the binding. This sheet is put on top of the stack of the dismantled book. When the bar code later is sent through the autoscanner, it is identified, thus making an automatic connection between the metadata file and the scanned binding.

In the case of page-turner scanning, the binding and the content of the book are scanned on the same equipment. This process too fetches metadata from the catalogue and generates an XML file with metadata that accompanies the book through the rest of the process.

## **OCR/DSA**

After digitization the digital book and its accompanying metadata will be placed on temporary storage ready for further processing. The books must be manually imported into the docWorks software, but from there on the processing of most books is fully automated. Manual operators are only employed for handling of exceptions when the software calls attention to errors in processing (i.e. when processing failed to stay within the defined tolerance limits).

In addition, operators are used for quality control of special parts of the collection that we want to process further.

Books of high priority are placed in special folders that are imported before the standard systematic digitization.

After OCR treatment and document structural analysis, losslessly compressed JPEG2000 files are generated from all the image files of the book. This is the format used for preservation.

## **Digital preservation**

After processing in docWorks, a METS object containing metadata, the digital book, the OCR processed text and structural information is generated. This object is placed in the National Library's digital long term repository for preservation.

Simultaneously the catalogue is updated with the digital ID of the book.

## **Indexing**

At regular intervals, an OAI import of catalogue data is performed. If this import finds that a book has been updated with a digital ID, a process is started in order to fetch the metadata and text of the book from the digital long term repository and index both, making the book available for search in the digital national library.

## **4. Important lessons learned this far**

### **Scope, complexity and implementation**

Implementing an integrated production line for the digitization of books with a high degree of automation turned out to be far more extensive and complicated than anticipated.

In order not to waste time, as soon as the decision to establish a production line for the digitization of books had been made, we launched activities aimed at realizing the first part of the chain of production (ordering, extraction, transportation and the digitization itself). In order to realize the efficient extraction of material, adaptations had to be made both to BIBSYS, which is the National Library's book cataloguing system, and to the software that runs the automated storage system from which the books are retrieved. This made us dependent on two external suppliers, which had consequences for the rate of development.

Carrying out an invitation for tenders for the purchase of scanners is also a long and time consuming process. We could not develop the method of digitization until it was clear what kind of equipment we would be using, and then we had to get the first scanners installed before we were able to terminate the implementation and start testing.

After the first part of the production line was in place, we began test production. Having aimed at a high production rate, we soon found ourselves with large amounts of data in temporary storage. While waiting for the rest of the production line to be put in place, we had to establish temporary routines for the safeguarding of the digital content.

In order to set up the rest of the production line and the functionality of our digital library that would facilitate the search and display of books, a multitude of development activities had to be launched. Some examples: Installing and putting into operation the software for OCR and structural analysis of documents, and the integration of this system into the production line, the generation of preservation objects based on the METS standard, the process of ingesting the METS objects into the digital long term repository, a setup for updating the catalogue system with the digital ID, OAI harvesting of metadata from the catalogue system, the process which retrieves text and metadata from the digital long term repository for books which have a digital ID, the indexing of these, and the development of necessary functionality for the search and retrieval of books in the digital library. At the same time there arose great pressure to quickly get to see the results of the digitization already under way, which led to the pushing of deadlines and, during one period, a very high level of stress among the section responsible for development.

After the functionality for the display of books in the digital library was in place, it soon became clear that this would become a very interesting service which would provide a "lift" to our digital library. Also, we had received considerable media coverage of the digitization, and expectations of seeing results were great, both from external users and from inside the National Library. We

therefore decided to launch the service, even though the production line still was under development, meaning that many manual operations were needed to speed a book through the full production line. The service has functioned well, but the expectations of quickly reaching a large volume of digital books in the service were not fulfilled. This was mainly due to the fact that the production line was not fully implemented, and accordingly not put into regular service.

With hindsight it is easy to see that we ought to have had a greater focus from the beginning on the unity of both the whole production line and the necessary functionality in the digital library, and that from the start we should have better assigned ownership to the timeline in this development work in the whole organization.

### **Actual efficiency**

With basis in the specifications of the digitizing equipment, we assigned production targets already from the start. These took into consideration that this was a ramp-up phase. Still, it turned out that some factors we had not been aware of reduced the overall efficiency. This became most apparent when looking at the automated scanners.

The book pages were on the whole thicker than the reference paper used to measure the scanners' specifications. This led to a decrease in paper feed speed, which made a great impact on the daily production numbers.

Since we started with the oldest books, we had an issue with them being very dusty. This meant that the scanners needed considerably more cleaning and maintenance than expected by the supplier. This in turn meant reduced production time per day, and therefore reduced daily production compared to expectations.

Our plan called for the scanners to work as continuously as possible through the whole working day. This was to be realized through operators relieving each other at the scanning stations, taking their breaks at different times. This was a new and unfamiliar way of working, creating some resistance among the operators. In practice we have not been able to realize this well enough, and also this has contributed to reduced production time per day on the scanners compared to our prognoses.

Actual production has been between 60% and 80% of our stated production goals.

### **Quality**

During initial testing, all pages of the books digitized were subjected to quality control. Since then we have not had routines for quality assurance of the digitization work. After the functionality of the digital library became capable of displaying the books, there have been random quality checks of the quality of what is available through the service.

Books that were displayed in the digital library revealed that the compression of the digital display copies had been performed using inadequate settings. The visual quality of the images was lower than expected. After adjustments had been made, the results became much better.

Obviously, the quality of the digital books is closely related to the quality of the originals, Our algorithm for the automatic selection of books for digitization does not take this into account, and we accordingly run the risk of extracting inferior specimens for digitization even if there actually are very good copies in the collection.

The scanners that digitize the dismantled books scan both sides of every sheet in one operation. This means that two different digitizing units are processing the two sides of a sheet. It has proved quite difficult to calibrate these two units identically, resulting in colour variations between the pages. This has improved immensely since the test phase, but the problem has not yet been fully solved. Experiments have been made with scanning a reference sheet at the start of every book, to facilitate subsequent automatic colour adjustment. So far these experiments have not yielded the desired results, but we will strive towards a solution.

When the production line enters a phase of regular production, we plan to establish random quality control of digitized books.

## **OCR/DSA**

From the outset, we had planned for fully automated use of OCR and document analysis tools. These kinds of tasks had never before been performed at such a large scale, and we had no prior experience using tools of such advanced nature.

The first challenge was to establish a large scale production setup with eight instances of the software on eight blade servers. This was necessary in order to achieve sufficient processing capacity, but it turned out to be harder than expected to make this stable and operative.

The next challenge arose when it turned out to be impossible to run the system fully automated. This created an unforeseen need for resources, and manning this task caused us some headaches. We spent some time figuring this out, and had to postpone the planned training on the system. This in turn gave us a short-term competence problem since it is a complex system with very advanced functionality. The answer was in part found through close contact with the supplier. This challenge has now been overcome, and training has taken place.

Our initial expectations of precision in the fully automated structural analysis have so far not been met. Advanced structural analysis can be done, but with such a degree of uncertainty that manual quality control is absolutely necessary. The more advanced the analysis you want to employ, the more manual quality control you will need. Accordingly, we will use this only as an exception, in special dissemination projects. A simple calculation shows that a post control that takes on average 15 seconds per page, will in total require an effort equal to 18 man days per day at the present production level. For this we don't have resources. So we have been forced to stay at an absolute minimum level by only requiring correct pagination in the digital library service and that the table of contents must be directly linkable whenever a table of content is present.

So far we have focused on OCR and DSA of publications in Latin letters. Here we have acceptable precision in the letter recognition. For Gothic letters the results are worse, but even there we have a degree of recognition that opens up interesting possibilities for free text search. We are running separate configurations of the system for Latin and Gothic letters. Books are categorized when the bindings are scanned, and then they are routed to the correct configuration. We expect there to be a potential for improvement by further training of the software and more advanced configuration of the system.

## **The digital long term repository – scaling and performance**

So far we have operated a separate instance of the DSM (digital long term repository) for the digitization of books. The use so far does not show performance problems for the DSM, but the usage is expected to rise considerably compared to the present traffic, as the volume of digital books offered in the service increases.



Since we do not have in place an access solution (user authorization and authentication) for books, we have so far chosen to only place in the DSM books that we are allowed to give full access to in the digital library. Some of the logic today is that when a book is placed in the DSM its digital ID is entered into the catalogue. This in turn lets the book be fetched automatically from the DSM for indexing to the search index of our digital library.

This strategy means that most of the digitized books are still in temporary storage, albeit subject to the same storage policy as those in the DSM.

When a user asks for a certain book page at a certain quality (at present the interface gives a choice of three quality levels), a JPEG file of the desired quality is generated from the JPEG2000 file in the DSM. So far we have not implemented intelligent pre-caching or placing pages in a buffer outside the DSM to increase performance as perceived by the user, but this may be an interesting future development.

So far the rate of progress in technology has given us enough flexibility that we can operate a one machine solution for the digital long term repository (DSM). But this makes it vulnerable when errors arise. We are therefore continually evaluating the possibility for more robust solutions.

### **Statistical tools – production supervision tools**

So far we have been using simple UNIX tools to generate the statistics necessary in order to supervise production. We are now considering the development of more advanced general production supervision tools that will at any time give us updated information of where a given object is in the process. We should also be able to use them to generate production statistics from all stages of production.

### **Exception handling**

With a few exceptions everything in the production line that can be automated, has been automated. Still, some times there can arise deviations in all stages of the production line, and these deviations must be handled and followed up manually. This has been one of the greatest challenges in the production so far. We are now working on establishing routines and assigning precise responsibility for such follow-up in the line organization.

## **5. Summary**

The implementation of a production line for books has had its fair share of problems. However, we have learned from our mistakes, and today we have in place an advanced production line for books. By the end of the year the production line will come into regular operation, and the last pieces of assigning responsibility and exception handling will fall into place.

In spite of the challenges we have met on the way, we have had a considerable production during our test year. Close to 26,000 books with an average of more than 200 pages per book have now been digitized, and most of these have also undergone OCR and structural analysis. A little more than 1,500 books are freely accessible in their entirety in our digital library, where they are also searchable in full text.

The challenge we now face is to establish production lines for all the types of material that are to be digitized, so that we truly will be able to establish the multimedia digital national library.

# **Alternative File Formats for Storing Master Images of Digitisation Projects**

**Date:** March 7, 2008  
**Authors:** Robèrt Gillesse  
Judith Rog  
Astrid Verheusen  
National Library of the Netherlands  
Research & Development Department  
**Version:** 2.0  
**Status:** Final  
**File Name:** Alternative File Formats for Storing Masters 2.0.doc

## Management Summary

This document is the end result of a study regarding alternative formats for storing master files of digitisation projects of the Koninklijke Bibliotheek (KB) in The Hague, the Netherlands. The study took place in the context of reviewing the KB's storage strategy. The magnitude of digitisation projects is increasing to such a degree – the estimate is a production of 40 million images (counting master versions only) in the next four years – that a revision of this strategy is deemed necessary. Currently, master files are stored in uncompressed TIFF file format, a format used world wide. 650 TB of storage space will be necessary to store 40 million files in this format. The objective of the study was to describe alternative file formats in order to reduce the necessary storage space. The desired image quality, long-term sustainability and functionality had to be taken into account during the study.

The following four file formats were reviewed:

- JPEG 2000 part 1 (lossless and lossy)
- PNG 1.2
- Basic JFIF 1.02 (JPEG)
- TIFF LZW

For each file format, we examined what the consequences would be for the following if that format were to be selected:

1. The required storage capacity
2. The image quality
3. The long-term sustainability
4. The functionality

The KB distinguished three main reasons for wanting to store the master files for a long or even indefinite period:

1. Substitution (the original is susceptible to deterioration and another alternative, high-quality carrier – preservation microfilm – is not available)
2. Digitisation has been so costly and time consuming that redigitisation is no option
3. The master file is the basis for access, or in other words: the master file is identical to the access file

The recommendations for the choice of file format were made on the basis of these three reasons.

The study made use of existing knowledge and expertise at the KB's Research & Development Department. The *quantifiable file format risk assessment method* recently developed by the KB was employed for examining the long-term sustainability of the formats. The results of the study were presented to a selection of national and international specialists

in the area of digital preservation, file formats and file management. Their comments are incorporated into the final version of this document.

The main conclusions of this study are as follows:

*Reason 1: Substitution*

JPEG 2000 lossless and PNG are the best alternatives for the uncompressed TIFF file format from the perspective of long-term sustainability. When the storage savings (PNG 40%, JPEG 2000 lossless 53%) and the functionality are factored in, the scale tips in favour of JPEG 2000 lossless.

*Reason 2: Redigitisation Is Not Desirable*

JPEG 2000 and JPEG are the best alternatives for the uncompressed TIFF file format. If no image information may be lost, then JPEG 2000 lossless and PNG are the two recommended options.

*Reason 3: Master File is the Access File*

JPEG 2000 lossy and JPEG with greater compression are the most suitable formats.

# Table of Contents

<b>MANAGEMENT SUMMARY</b> .....	<b>2</b>
<b>1 INTRODUCTION</b> .....	<b>6</b>
1.1 CONSEQUENCES .....	7
1.2 THREE REASONS FOR LONG TERM STORAGE OF MASTER FILES.....	9
1.3 CONCLUSION.....	9
1.4 REVIEW BY SPECIALISTS.....	9
1.5 FOLLOW-UP STUDY .....	10
<b>2 JPEG 2000</b> .....	<b>12</b>
2.1 WHAT IS JPEG 2000? .....	12
2.1.1 General Information .....	12
2.1.2 JPEG 2000 Parts.....	12
2.2 HOW DOES IT WORK? .....	15
2.2.1 Structure .....	15
2.2.2 Encoding and Decoding .....	15
2.3 CONSEQUENCES FOR THE REQUIRED STORAGE CAPACITY.....	18
2.4 CONSEQUENCES FOR THE IMAGE QUALITY .....	18
2.5 CONSEQUENCES FOR THE LONG-TERM SUSTAINABILITY.....	19
2.6 CONSEQUENCES FOR THE FUNCTIONALITY .....	20
2.7 CONCLUSION.....	21
<b>3 PNG</b> .....	<b>24</b>
3.1 WHAT IS PNG?.....	24
3.2 HOW DOES IT WORK?.....	25
3.2.1 Structure .....	25
3.2.2 Encoding and Decoding/Filtering and Compression .....	25
3.3 CONSEQUENCES FOR THE REQUIRED STORAGE CAPACITY.....	25
3.4 CONSEQUENCES FOR THE IMAGE QUALITY .....	26
3.5 CONSEQUENCES FOR THE LONG-TERM SUSTAINABILITY.....	26
3.6 CONSEQUENCES FOR THE FUNCTIONALITY .....	26
3.7 CONCLUSION.....	27
<b>4 JPEG</b> .....	<b>29</b>
4.1 WHAT IS JPEG? .....	29
4.2 HOW DOES IT WORK? .....	30
4.2.1 Structure .....	30
4.2.2 Encoding and Decoding/Filtering and Compression .....	30
4.3 CONSEQUENCES FOR THE REQUIRED STORAGE CAPACITY.....	31
4.4 CONSEQUENCES FOR THE IMAGE QUALITY .....	31
4.5 CONSEQUENCES FOR THE LONG-TERM SUSTAINABILITY.....	32
4.6 CONSEQUENCES FOR THE FUNCTIONALITY .....	32
4.7 CONCLUSION.....	33
<b>5 TIFF LZW</b> .....	<b>35</b>
5.1 WHAT IS TIFF LZW?.....	35
5.2 HOW DOES IT WORK? .....	36
5.2.1 Structure .....	36
5.2.2 Encoding and Decoding/Filtering and Compression .....	36
5.3 CONSEQUENCES FOR THE REQUIRED STORAGE CAPACITY.....	36
5.4 CONSEQUENCES FOR THE IMAGE QUALITY .....	36
5.5 CONSEQUENCES FOR THE LONG-TERM SUSTAINABILITY.....	36
5.6 CONSEQUENCES FOR THE FUNCTIONALITY .....	37

5.7 CONCLUSION.....	38
<b>6 CONCLUSION .....</b>	<b>40</b>
<b>APPENDIX 1: USE OF ALTERNATIVE FILE FORMATS.....</b>	<b>45</b>
<b>APPENDIX 2: FILE FORMAT ASSESSMENT METHOD – OUTPUT .....</b>	<b>48</b>
<b>APPENDIX 3 FILE FORMAT ASSESSMENT METHOD – EXPLAINED.....</b>	<b>50</b>
<b>APPENDIX 4: STORAGE TESTS .....</b>	<b>62</b>
<b>BIBLIOGRAPHY .....</b>	<b>63</b>

# 1 Introduction

The study took place in the context of reviewing the Koninklijke Bibliotheek's (KB) storage strategy for digitisation projects. The magnitude of digitisation projects is increasing to such a degree – the estimate is 40 million images and 650 TB in uncompressed data storage by 2011, counting master versions only – that a revision of this strategy is deemed necessary. The central questions are whether all master files of digitisation projects actually have to be stored in the long-term storage system, what the costs are of long-term storage and what the alternatives are besides expensive, uncompressed, high-resolution storage in TIFF file format.

This study focuses on the last question. Its objective is to describe alternative file formats besides uncompressed TIFF for storing image master files.

In this study the context is that of digitized low contrast material – which means originals like for instance older printed text, engravings, photographs and paintings. Higher contrast materials – read: (relatively) modern, non illustrated printed material – are out of the scope of this study. The classification of different types of originals on the basis of their information value, the selection of a suited digitization quality connected to this value and, subsequently, the choice of using either lossy compression, lossless compression or no compression at all, are issues that have not been elaborated upon in this study. These two subjects will have to be part of a possible second version of this study.

Master images are defined as followed: Raster images that are a high quality (in either colour, tonality or resolution) copy from the original source from which in most cases derivatives are made for access use.

The following images are excluded from the scope of this study:

- Vector images
- 3D images
- Moving images
- Images in various editing layers (not identical to multiresolution images<sup>1</sup>)
- Multipage files (PDF, multipage TIFF are dropped from consideration)
- Multispectral, hyperspectral images<sup>2</sup>

The following four file/compression formats will be reviewed:

1. JPEG 2000 part 1 (lossless and lossy)<sup>3</sup>
2. PNG 1.2

---

<sup>1</sup> Photoshop .psd or TIFF multilayer files, for example.

<sup>2</sup> This is because multispectral imaging has been no serious consideration for KB digitisation projects. This is not to say that this will not be case in the future. For sake of the argument now, multispectral images are not relevant.

<sup>3</sup> A review of JPEG 2000 as an alternative file format was already conducted in large measure in 2007 by Judith Rog: *Notitie over JPEG 2000 voor de KB (Note regarding JPEG 2000 for the RL)*, version 0.2 (August 2007).



3. Basic JFIF 1.02 (JPEG)
4. TIFF LZW

The arguments for selecting precisely these four file formats reside in the following requirements for an *alternative* master file:

- Software support (very new or rarely used formats such as Windows Media Photo/JPEG XR and JPEG-LS are dropped from consideration).
- Sufficient bit depth: A minimum of 8 bits greyscale or 24 bits colour (bitonal, 1 bit, TIFF G4/JBIG files are dropped from consideration<sup>4</sup>, as well as GIF due to 8 bits, limited colour palette).
- Possibility for lossless or high-end lossy compression (BMP excluded).

TIFF with lossless ZIP compression is excluded from this study out of sheer shortage of time. In the second version of this study TIFF zip will have to be included.

### 1.1 Consequences

This report has an individual section for each of the four file formats listed above. A summary description of the format and how it works is followed by subsections describing the consequences of using the format in the following areas:

1. Consequences for the required storage capacity
2. Consequences for the image quality
3. Consequences for the long-term sustainability
4. Consequences for the functionality

Sub 1: This section provides an outline of the storage consequences of the format choice. The storage gain of the compressed compared to the uncompressed TIFF file is calculated in term of percentage: if necessary, a differentiation shall be made between lossy and lossless compression. Two sets comprising about one hundred scans are employed for the calculation... On these tests two limitations have been set:

- Only 24 bit, RGB (8 bit per colour channel) files have been tested
- Only two sets of originals have been tested: a set low contrast text material and a set of photographs

These limitations were born out of fact that the great majority of KB files that were made and shall be made in the near future are of this nature: 24 bit RGB files of a low contrast nature. Of course higher (and maybe lower) bit depths and high contrast materials (modern print), which will yield other compression ratio's, will have to be included a in later versions of this study.

See Appendix 4 for the results of the text set.

Sub 2. This section attempts to outline the difference with regard to the uncompressed master file using as many quantifiable terms as possible (among other things by means of the Peak Signal-to-Noise Ratio – PSNR<sup>5</sup> – and Modulation Transfer Function – MTF<sup>6</sup>).

---

<sup>4</sup> Whether bitonal files actually fall outside the scope of image master files is not yet certain. It may be that the loss of brightness values is considered acceptable for some access projects (see below) of (relatively) modern, unillustrated material.

The following technical targets and tools are used to determine the possible decrease of image quality.

- Possible loss of detail is measured by means of the QA-62 SFR and OECF test chart.
- Possible loss of greyscale is measured using the Kodak Greyscale.
- Possible loss of colour is measured using the MacBeth ColorChecker.
- Artefacts are determined through visual inspection.

Sub 3: This section employs the *quantifiable file format risk assessment method* recently developed by Judith Rog, Caroline van Wijk and Jeffrey van der Hoeven for the KB. Using this method, file formats can be measured based on seven widely accepted sustainability criteria. The criteria are as follows: Openness, Adoption, Complexity, Technical Protection Mechanism, Self-documentation, Robustness and Dependencies. Each file format receives a sustainability score in this method. These seven main criteria are subdivided into measurable sub-characteristics. For example, the main criterion “Openness” is subdivided into the characteristics “Standardization,” “Restrictions on the interpretation of the file format” and “Reader with freely available source”. Each format receives a score between 0 and 2 for each characteristic. The method precisely defines how the score is determined. For example, a format will receive the maximum score of 2 for the “Standardization” characteristic if it is a “de jure standard”, a score of 1.5 if it is a “de facto standard” and so on down to a score of 0. The scores are subsequently multiplied with a weighing factor that is attributed to each main criterion or characteristic. The weights that are assigned to the criteria and their characteristics are not fixed. They depend on the local policy of an institution. A weight of 0 can be assigned if an institution chooses to ignore the characteristic. The weights that are used in the examples in this paper are the weights as assigned by the KB based on its local policy, general digital preservation literature and common sense. For example, the sub-characteristics “Standardization,” “Restrictions on the interpretation of the file format” and “Reader with freely available source” of the “Openness” criterion receive a weighing factor at the KB of 9 (Standardization), 9 (Restrictions on the interpretation of the file format) and 7 (Reader with freely available source), and all sub-characteristics of the criterion “Self-documentation,” which includes the option of adding metadata to files, receive a weighing factor of 1. The KB will initially not employ metadata that is stored in the files themselves. This is the reason for the relatively low weighing factor for this criterion. However, this may be different for other institutions. In this method, each file format ultimately receives a sustainability score between 0 and 100. The higher the score, the more suitable the format is for long-term storage and permanent access.

Appendix 2 of this document contains the interpretation of the method for the formats that are discussed in this report. Appendix 3 explains the method. For this study all discussed formats received a score of 0 for the “Support for file corruption detection” characteristic because the time and expertise to research this was lacking at this point in time. We are aware that PNG does provide a level of corruption detection in the file header, but lacked the time to research whether and to what level this is case for the other formats. Because all formats received the same score, this ultimately plays no role in the relative final scores of the formats with regard to each other.

Because the method was developed so recently and feedback is still awaited from colleague institutions, the ultimate choice for an alternative format is not solely based on the File

---

<sup>5</sup> Cit.“[...] the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation.” Wikipedia: <http://en.wikipedia.org/wiki/PSNR>

<sup>6</sup> MTF is a measurement of detail reproduction of an optical system. Output is in reproduced line pairs (or cycles) per millimetre.

Format Assessment Method. The results of the method are tested against other knowledge and experience in this area.

Sub 4: This section outlines the consequences for functionality. The section deals with questions of usage such as the following:

- Is the file format suitable as a high-resolution access master?
- Are there options for including bibliographic and technical (EXIF) metadata?
- Are the Library of Congress criteria regarding quality and functionality factors complied with? Normal display, clarity (support of colour spaces, possible bit depths), colour maintenance (option to include gamma correction and ICC colour profiles), support of graphical effects and typography (Alpha channel where transparency information can be stored) and functionality beyond the normal reproduction (animation, multipage and multiresolution support)<sup>7</sup>?

## 1.2 Three Reasons for Long Term Storage of Master Files

As said above the master file is a high quality copy from the original source from which in most cases derivatives are made for access use. It is possible to delete the master files after the access derivatives have been made. In which case, when other, more demanding use of the files is needed, digitisation will have to be performed again.

The KB distinguishes three main reasons for wanting to store the master files for a long or even indefinite period:

1. Substitution (the original is susceptible to deterioration and another alternative, high-quality carrier – preservation microfilm – is not available)
2. Digitisation has been so costly and time consuming that redigitisation is no option
3. The master file is the basis for access, or in other words: the master file is identical to the access file

These three reasons form the basis on which the recommendations for the different file formats are made.

## 1.3 Conclusion

Each analysis is followed by a conclusion per format and after all analyses have been discussed, an overall conclusion is presented where the various formats are compared and recommendations are made for the selection of an alternative file format. The above mentioned reasons for long term storage will be included in this.

## 1.4 Review by Specialists

A panel of national and international specialists in the area of digital preservation, file formats and file management was asked to critically review this study and provide comments, where required. Their input is incorporated into this version of the report. The panel consisted of the following persons:

---

<sup>7</sup> Sustainability of Digital Formats - Planning for Library of Congress Collections - Still Images Quality and Functionality Factors [http://www.digitalpreservation.gov/formats/content/still\\_quality.shtml](http://www.digitalpreservation.gov/formats/content/still_quality.shtml)

- Stephen Abrams (Harvard University Library/University of California-California Digital Library )
- Caroline Arms (Library of Congress, US)
- Martijn van den Broek (Nederlands Fotomuseum [Netherlands Photo Museum], the Netherlands)
- Adrian Brown (National Archives, UK)
- Robert R. Buckley (Xerox Corporation)
- Aly Conteh (British Library)
- Carl Fleischhauer (Library of Congress)
- Rose Holley (National Library of Australia)
- Marc Holtman (City Archive of Amsterdam)
- Rene van Horik (DANS, the Netherlands)
- Dr. Klaus Jung (Luratech Imaging GmbH)
- Ulla Bøgvad Kejser (Kongelige Bibliotek Denmark)
- Rory McLeod (British Library)
- Andrew Stawowczyk Long (National Library of Australia)
- Boudewijn de Ridder (Nederlands Fotomuseum [Netherlands Photo Museum], the Netherlands)
- Brian Thurgood (Libraries and Archives of Canada)
- Thomas Zellmann (LuraTech Europe GmbH)

We would like to thank all experts for their very useful feedback that improved this report considerably. The amount of feedback that we received was overwhelming and showed us that the problem that was the immediate cause for this report is relevant at many other institutions as well.

## **1.5 Follow-up Study**

All the feedback that we have received confirmed us in the idea that the report is far from complete as it is. We could easily have spent several more months on further, in-depth study on all the topics that are being addressed in the report. Unfortunately we lack the time to do so.

Among others, these items remain open for further study:

- A classification of different types of originals on the basis of their information value, the selection of a suited digitization quality connected to this value and, subsequently, the choice of using either lossy compression, lossless compression or no compression at all
- Further compression tests including:
  - High contrast, textual material
  - 16 bit files
  - Greyscale files
  - Using alternative compression software for JPEG2000 and PNG

- PSNR – Peak Signal To Noise Ratio
- Structure of the JPEG file
- Functioning of LZW compression
- Further work on the “File Format Assessment Method” that is being used in this report to assess file formats on their long-term sustainability. On the basis of the feedback from experts mentioned above, we have already adjusted and refined the method, but it will need our further and constant attention.

We are very much open to all input from others on one of these or all other topics from this study.

## 2 JPEG 2000

### 2.1 What is JPEG 2000?

#### 2.1.1 General Information

JPEG 2000 is a standard (ISO/IEC 15444-1/ITU-T Rec. T.800) developed by the JPEG (Joint Photographic Experts Group) as a joint effort of the ISO, IEC and ITU-T standardization organizations. These groups are comprised of representatives of various commercial parties and academic institutes from the four corners of the globe.

The objective of the JPEG group was to develop a new image standard with the following basic principles:

- Complete openness of the format.
- An improved lossy compression algorithm compared to the current JPEG compression.
- An option for lossless compression.
- Comprehensive options for bundling metadata in the image file.
- Storage of several resolutions within one file.

These basic principles were implemented in the JPEG 2000 standard.

#### 2.1.2 JPEG 2000 Parts

JPEG 2000 in the year 2007 is divided into twelve standards that are all more or less derivations of or supplements to the first standard: Part 1. This concerns still images (part 1 .jp2 and part 2 .jpx), documents (part 6 .jpm) and moving images (part 3 .mj2). The employed wavelet compression technology is the connecting element.

Only parts 1, 2, 4, 6 and 8 seem to be relevant for storing masters of still images.

The following contains an overview of the twelve parts in a summarized form<sup>8</sup>

#### **Part 1**

As its name suggests, Part 1 defines the core of JPEG 2000. This includes the syntax of the JPEG 2000 codestream and the necessary steps involved in encoding and decoding JPEG 2000 images. The later parts of the standard are all concerned with extensions of various kinds, and none of them is essential to a basic JPEG 2000 implementation. A number of existing implementations use only Part 1.

Part 1 also defines a basic file format called JP2. This allows metadata such as colour space information (which is essential for accurate rendering) to be included with a JPEG 2000 codestream in an interoperable way. JP2 uses an extensible architecture shared with the other file formats in the JPEG 2000 family defined in later parts of the standard.

Part 1 also includes guidelines and examples, a bibliography of technical references, and a list of companies from whom patent statements have been received by ISO. JPEG 2000 was developed with the intention that Part 1 could be implemented without the payment of licence

---

<sup>8</sup> See for full description of the parts the JPEG 2000 homepage: <http://www.jpeg.org/jpeg2000/>.

fees or royalties, and a number of patent holders have waived their rights toward this end. However, the JPEG committee cannot make a formal guarantee, and it remains the responsibility of the implementer to ensure that no patents are infringed.

Part 1 became an International Standard (ISO/IEC 15444-1) in December 2000.

A second edition of Part 1 was published in 2004. Among other things, a standard colour spaces (YCC) was added.

## **Part 2**

Part 2 defines various extensions to Part 1, including:

- More flexible forms of wavelet decomposition and coefficient quantization
- An alternative way of encoding regions of particular interest (ROIs)
- A new file format, JPX, based on JP2 but supporting multiple compositing layers, animation, extended colour spaces and more
- A rich metadata set for photographic imagery (based on the DIG35 specification)

Most of the extensions in Part 2 operate independently of each other. To assist interoperability, mechanisms are provided at both the codestream and the JPX file format level for signalling the use of extensions.

Part 2 became an International Standard (ISO/IEC 15444-2) in November 2001.

## **Part 3**

Part 3 defines a file format called MJ2 (or MJP2) for motion sequences of JPEG 2000 images. Support for associated audio is also included.

Part 3 became an International Standard (ISO/IEC 15444-3) in November 2001

## **Part 4**

JPEG 2000 Part 4 is concerned with testing conformance to JPEG 2000 Part 1. It specifies test procedures for both encoding and decoding processes, including the definition of a set of decoder compliance classes. The Part 4 test files include both bare codestreams and JP2 files.

Note that JPEG 2000 Part 4 explicitly excludes from its scope acceptance, performance or robustness testing.

Part 4 became an International Standard (ISO/IEC 15444-4) in May 2002.

## **Part 5**

JPEG 2000 Part 5 (ISO/IEC 15444-5:2003) consists of a short text document, and two source code packages that implement JPEG 2000 Part 1. The two codecs were developed alongside Part 1 and were used to check it and to test interoperability. One is written in C and the other in Java. They are both available under open-source type licensing.

Part 5 became an International Standard (ISO/IEC 15444-5) in November 2001.



## **Part 6**

Part 6 of JPEG 2000 defines the JPM file format for document imaging, which uses the Mixed Raster Content (MRC) model of ISO/IEC 16485. JPM is an extension of the JP2 file format defined in Part 1: it uses the same architecture and many of the same boxes defined in Part 1 (for JP2) and Part 2 (for JPX).

JPM can be used to store multi-page documents with many objects per page. Although it is a member of the JPEG 2000 family, it supports the use of many other coding or compression technologies as well. For example, JBIG2 could be used for regions of text, and JPEG could be used as an alternative to JPEG 2000 for photographic images.

Part 6 became an International Standard (ISO/IEC 15444-6) in April 2003.

## **Part 7**

This part has been abandoned.

## **Part 8**

JPEG 2000 Secured (JPSEC) or Part 8 of the standard is standardizing tools and solutions in terms of specifications in order to ensure the security of transaction, protection of contents (IPR), and protection of technologies (IP), and to allow applications to generate, consume, and exchange JPEG 2000 Secured bitstreams. The following applications are addressed: encryption, source authentication, data integrity, conditional access, ownership protection.

Part 8 became an International Standard (ISO/IEC 15444-8) in July 2006

## **Part 9**

The main component of Part 9 is a client-server protocol called JPIP. JPIP may be implemented on top of HTTP, but is designed with a view to other possible transports.

Part 9 became an International Standard (ISO/IEC 15444-9) in October 2004.

## **Part 10**

Part 10 is at the end of the Approval Stage (50.60).. It is concerned with the coding of three-dimensional data, the extension of JPEG 2000 from planar to volumetric images.

## **Part 11**

To address this issue, JPEG 2000 Wireless (JPWL) or Part 11 of the standard is standardizing tools and methods to achieve the efficient transmission of JPEG 2000 imagery over an error-prone wireless network. More specifically, JPWL extends the elements in the core coding system described in Part 1 with mechanisms for error protection and correction. These extensions are backward compatible in the sense that decoders which implement Part 1 are able to skip the extensions defined in JPWL.

Part 11 became an International Standard (ISO/IEC 15444-11) in June 2007.

## **Part 12**

Part 12 of JPEG 2000, ISO/IEC 15444-12, has a common text with Part 12 of the MPEG-4 standard, ISO/IEC 14496-12. It is a joint JPEG and MPEG initiative to create a base file format for future applications. The format is a general format for timed sequences of media data. It uses the same underlying architecture as Apple's QuickTime file format and the JPEG 2000 file format.

Part 12 became an International Standard (ISO/IEC 15444-12) in July 2003.

## **Part 13 - An entry level JPEG 2000 encoder**

Part 13 defines a royalty- and license-fee free entry-level JPEG 2000 encoder with widespread applications. There is no Final Committee Draft available yet.

### **2.2 How does it work?**

#### **2.2.1 Structure**

A JPEG 2000 file is comprised of a succession of boxes. A box can contain other boxes and is then called a superbox.<sup>9</sup> The boxes are of variable length. The length is determined by the first four bytes. Each box has a type that is determined by the second sequence of the four bytes.

Each file of the JPEG 2000 family begins with a JPEG 2000 signature box, followed by a file type box which determines, among other things, the type (e.g. JP2) and the version. This is followed by the header box, which contains various boxes in which the resolution, bit depth and colour specifications are set down, among other things. Optional are boxes in which XML and non-XML structured metadata can be determined about the file. This is followed by a “contiguous codestream” box which contains the image data.<sup>10</sup>

#### **2.2.2 Encoding and Decoding**

JPEG 2000 encoding takes place in six steps<sup>11</sup>:

##### *Step 1: Colour Component Transformation (optional)*

First, the RGB colour space is changed to another colour space. This is an optional step, but mostly used and recommended for RGB-like colour spaces. Two options are possible for this:

1. Irreversible Colour Transform (ICT) to the YCbCr colour space
2. Reversible Colour Transform (RCT) to the YUV colour space

The first method is used for lossy compression and includes a simplification of the colour information and can bring about quantification errors.

##### *Step: 2 Tiling*

After the colour transformation the image is divided into so-called tiles. The advantage of this is that the decoder requires less memory in order to create the image. The size of the tiles can even be selected (if the encoding software offers this advanced option). If the tiles are made too small, or if the compression factor is very high, the same blocking effect can occur as with JPEG (this only applies to lossy compression). The size of the tiles has a minimal effect on

---

<sup>9</sup> This box structure is related to the Quicktime and MPEG-4 format. Boxes are “atoms” in these formats.

<sup>10</sup> For a comprehensive overview of the box structure of JP2, see the Florida Digital Archive description of JP 2 - section 1.14: [http://www.fcla.edu/digitalArchive/pdfs/action\\_plan\\_bgounds/jp2\\_bg.pdf](http://www.fcla.edu/digitalArchive/pdfs/action_plan_bgounds/jp2_bg.pdf).

<sup>11</sup> Wikipedia, [JPEG2000: http://en.wikipedia.org/wiki/JPEG\\_2000#Technical\\_discussion](http://en.wikipedia.org/wiki/JPEG_2000#Technical_discussion).

the file size: when a smaller tile is chosen, the file becomes a bit larger.<sup>12</sup> This step is optional as well, in the sense that you can use one single tile that has the same dimensions as the whole image. This would prevent the blocking effect/tiling artefacts, mentioned earlier.

#### *Step 3: Wavelet Transformation*

The tiles are then transformed with Discrete Wavelet Transformation (DWT).<sup>13</sup>

There are two possibilities for this:

1. Lossy (or visual lossless) compression by means of the 9/7 floating point wavelet filter.
2. Lossless compression by means of the 5/3 integer wavelet filter.<sup>14</sup>

#### *Step 4: Quantification (for lossy compression only)*

Scalar quantification of the coefficients in order to decrease the quantity of bits that represent them. The result is a set of whole numbers that must be encoded. The so-called quantification step is a flexible parameter: the larger this step, the greater the compression and the loss of quality.

#### *Step 5: Encoding*

Encoding includes a hierarchical succession of continually smaller “units”:

1. Sub-bands – frequency range and spatial area. These elements are split into:
2. Precincts – rectangular regions in the wavelet domain. These elements are split into the smallest JPEG 2000 element:
3. Code blocks: square blocks in a sub-band. The bits of the code blocks are encoded by means of the EBCOT (Embedded Block Coding with Optimal Truncation) scheme. The significant bits are encoded first and then the less significant bits. The encoding itself takes place in three steps (coding passes), whereby the less relevant bits are filtered out in the lossy version.

#### *Step 6: Packetizing*

This is the process whereby the codestream is divided into “packets” and “layers” that can be sorted by resolution, quality, colour or position within a tile.

Packets contain the compressed data of the code blocks of a specific position of a given resolution of a component of a tile.

The packets, in turn, are a component of a layer: a layer is a collection of packets, one of each position, for each resolution.<sup>15</sup>

By arranging these layers in a certain way, it is possible during decoding/access to stipulate that certain information be made available first and other information later. This particularly plays a role for access via the Web.

---

<sup>12</sup> Robert Buckley, *JPEG 2000 for Image Archiving, with Discussion of Other Popular Image Formats*. Tutorial IS&T Archiving 2007 Conference, p. 41, slide 81.

<sup>13</sup> Instead of Discrete Cosine Transformation (DCT), which is used for JPEG. The DCT technique works in blocks of 8x8 pixels, which renders the image pixelated with higher compression.

<sup>14</sup> Robert Buckley, *JPEG 2000 for Image Archiving, with Discussion of Other Popular Image Formats*. Tutorial IS&T Archiving 2007 Conference, p. 42, slide 83.

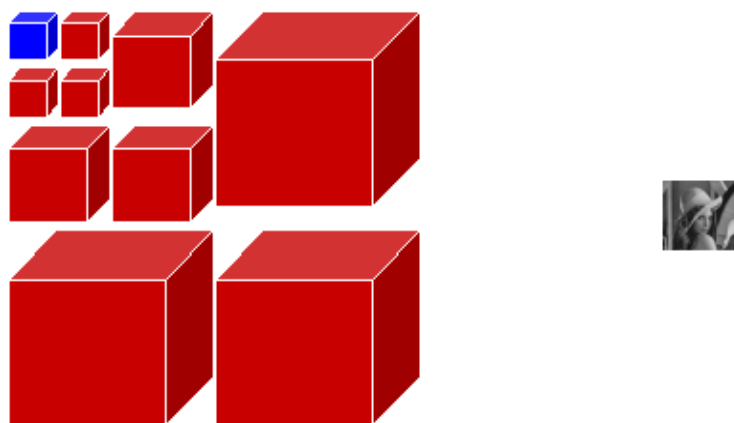
<sup>15</sup> *Ibidem*, p. 32, slide 64.

For example, if you choose to arrange the decoding per resolution, then you can first offer a low-resolution image during access, with larger-resolution images becoming available as the decoding goes on. If you arrange the codestream by quality, then you can repeatedly offer more quality/bit depth. Arranged by colour channels, you can always offer various colours and arranged by location, you can show certain parts of the image first. For example, the codestream can be arranged so that access takes place first by Quality (L), then Resolution (R), then Colour Channel (C) and then Position (P). The order is then LRCP. Other possible orders are: RLCP, RPCL, PCRL and CPRL. A special option of LRCP (LRCP with Region of Interest Encoding) is constructing a certain part of the image first.<sup>16</sup>

The two illustrations below<sup>17</sup> show how continually decoding more blocks results in a continuously higher resolution (RPCL).

## Scalability: Progressive By Resolution

---



**Figure 1: Resolution with one block decoded**

---

<sup>16</sup>Buckley, *JPEG 2000 Image Archiving*, page 34, slide 68.

<sup>17</sup> *Ibidem*, p. 28, slide 55, 56.

## Scalability: Progressive By Resolution

---

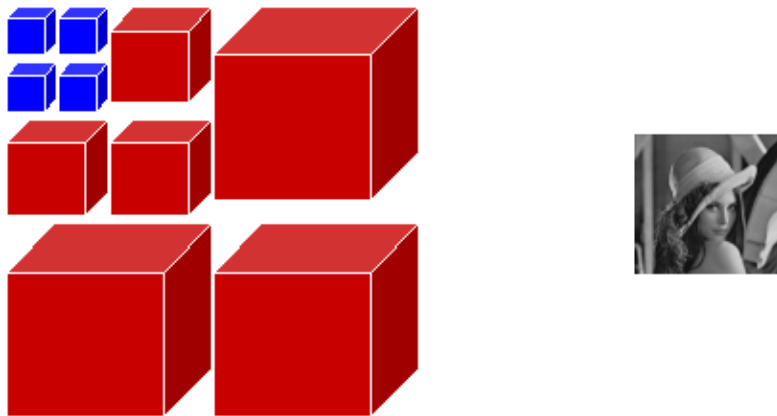


Figure 2: Resolution with three blocks decoded

### 2.3 Consequences for the Required Storage Capacity

Based on various test sets<sup>18</sup>, it appears that JPEG 2000 in lossless mode can yield a benefit of about 50% compared to an uncompressed file.

Based on the test material, it appears that the gain that can be achieved with JPEG 2000 part 1 *lossy* compression can vary – assuming the Lead Photoshop plugin compression ratio settings between 10 and 50 – between 91% and 98%<sup>19</sup>.

### 2.4 Consequences for the Image Quality

The lossless mode has no consequences for the image quality.

Lossy versions:

The quantity of compressions does degrade the image. Five versions were tested by means of the Lead JPEG 2000 Photoshop plugin (compression ratio): 25, 75, 100 and 500.

*Detailed Loss – MTF*

Original TIFF (QA-62 SFR and OECF test chart): MTF 5.91/5.91. File size 4.7 MB

Compression Ratio	MTF Horizontal and Vertical (three RGB channels on average)	File Size

<sup>18</sup> See Appendix 4.

<sup>19</sup> The Lead Photoshop plugin might not give optimal compression results. An alternative test with the native JPEG2000 plugin proved to give no great differences though. Lossless compression of the Photoshop plugin proved to be slightly less successful than that of the LEAD plugin: 53% for the former, versus 52% for the latter. Further testing of alternatives (like the Lurawave clt command compression tool <http://www.luratech.com/products/lurawave/jp2/clt/>) will have to be performed.

25	5.8 / 5.8	83 KB
75	5.8 / 5.8	62 KB
100	5.8 / 5.8	47 KB
500	3.9 / 3.1	10 KB

### *Greyscale and Colour Loss*

There is no measurable loss of greyscale in Kodak Greyscale.

Delta E values remain the same with various compression values and no extra colour shift occurs.

### *Artefacts*

In JPEG 2000 files, three clearly visible artefacts appear when the compression increases (tested based on various types of materials):

1. Posterizing or banding (coarse transitions in colour or grey tones). Clearly visible starting at an approximate compression ratio of 75 for text materials. For continuous tone images artefacts are becoming slightly visible at compression ratio 100, and well visible at 200..
2. Tiling effect: The tiles only become visible with extreme compression (compression ratio 200). This could be prevented by choosing a tile that has the same dimension of the image itself.
3. Woolly effect around elements rich in contrast. Visible starting at an approximate compression ratio of 75.

The last effect is particularly visible in text (around the letters). Continuous tone originals such as photos and paintings appear to be more suitable for strong JPEG 2000 compression than text materials do (or other materials with high-contrast transitions such as line drawings).

### *PSNR*

Topic of investigation.

## **2.5 Consequences for the Long-Term Sustainability**

In order to be able to make an accurate comparison between the JP2 format and the other formats that are either lossless compressed (“PNG 1.2” and “TIFF 6.0 with LZW compression”) or lossy compressed (“basic JFIF (JPEG) 1.02”), we divide the JP2 format into “JP2 (JPEG 2000 Part 1) lossless” and “JP2 (JPEG 2000 Part 1) lossy.”

The application of the “File Format Assessment Method” to the “JP2 (JPEG 2000 Part 1) lossless” format results in a score of 74,7 on a scale of 0-100. For the lossy compressed version, the method results in a score of 66,1. When the four formats that are compared in this report are sorted from most to least suitable for long-term storage according to the named method, “JP2 (JPEG 2000 Part 1) lossless” ends up in second place with this score, right after PNG 1.2 (with a score of 78). We then come to a point where the applied method possibly comes up short. In the method, the characteristic “Usage in the cultural heritage sector as master image file” of the Adoption criterion makes a valuable contribution to the total score. However, what is not included in the method at the moment are the prospects for the future of Adoption. The expectation is that – although JPEG 2000 and PNG are currently not used as master files to a large extent – JPEG 2000 does have potential as a master file. PNG has existed since 1996 and JP2 only since 2000.

“JP2 (JPEG 2000 Part 1) lossy” ends up in third place right above “basic JFIF (JPEG) 1.02” with almost the same score. For both the lossless as well as the lossy versions of JPEG 2000,

the score is primarily low due to the low adoption rate of the format. Adoption is a very important factor in the method. Despite the almost equal scores of “basic JFIF (JPEG) 1.02” and the lossy version of JPEG 2000, there is still a preference for “basic JFIF (JPEG) 1.02” due to the more certain future of this file. A report on the usage of JPEG 2000 as a Practical Digital Preservation Standard has recently been published on the DPC website<sup>20</sup>.

## 2.6 Consequences for the Functionality

- Options for including bibliographic and technical (EXIF) metadata
  - Bibliographic metadata: It is possible to add metadata in three boxes: One for XML data, a limited IPR (Intellectual Property Rights) box<sup>21</sup> and UUID (Universal Unique Identifier) based on ISO 11578:1996.
  - Technical metadata: There is as yet no standard manner for storing EXIF metadata in the JPEG 2000 header. Suggestions have been made to do this in an UUID box.<sup>22</sup>
- Suitability of the format for offering it as a high-resolution access master
  - Browser support: very limited (only by Apple’s Safari browser)<sup>23</sup>.
  - High-resolution image access: Because browsers do not yet support JPEG 2000, a JPEG generated on-the-fly is typically used as an intermediary image.
- Maximum size

Image dimensions width and height can be up to  $(2^{32})-1$ . File size can be unlimited with a special setup (take code stream box to the end of the file and signal “unknown” length). File format boxes can signal a length up to  $2^{64}-1$  bytes = 16 million TB. These are of course theoretical file sizes as no existing program will support them<sup>24</sup>.

### LOC Quality and Functionality Factors: <sup>25</sup>

- Normal display
  - Screen display: Yes
  - Printable: Yes
  - Zoomable: Yes
- Clarity
  - High-resolution options: Yes. A lot of compression can damage detailing (see section 3.4 above).
  - Bit depths: The JPEG 2000 Part 1 core file can vary from 1 bit to 38 bits.<sup>26</sup>

---

<sup>20</sup> Robert Buckley, JPEG 2000 – a Practical Digital Preservation Standard?, a DPC Technology Watch Series Report 08-01, February 2008: <http://www.dpconline.org/graphics/reports/index.html#jpeg2000>

<sup>21</sup> This option is greatly expanded in Part 8 of JPEG 2000 standard.

<sup>22</sup> Wikipedia, JPEG 2000, [http://en.wikipedia.org/wiki/JPEG\\_2000](http://en.wikipedia.org/wiki/JPEG_2000). The Adobe XML based XMP standard – which makes use of UUID box - seems to provide a standard way of storing EXIF information in the header. <http://www.pctoday.com/editorial/article.asp?article=articles%2F2005%2F0304%2F44t04%2F44t04web.asp>  
[http://en.wikipedia.org/wiki/Extensible\\_Metadata\\_Platform](http://en.wikipedia.org/wiki/Extensible_Metadata_Platform)

<sup>23</sup> <http://echoone.com/filejuicer/formats/jp2>

<sup>24</sup> Klaus Jung, email 13 february 2008 to Judith Rog

<sup>25</sup> [http://www.digitalpreservation.gov/formats/content/still\\_quality.shtml](http://www.digitalpreservation.gov/formats/content/still_quality.shtml).

<sup>26</sup> Buckley, *JPEG 2000 Image Archiving*, p. 45, slide 90. In Part 4, three different compliance classes can be indicated. Class 2 limits these options to 16 bits.



- Colour maintenance
  - Support of various colour spaces: Yes (though not via ICC profile).
  - Option for including gamma correction: No.
  - Options for including ICC colour profiles: JPEG 2000 part 1 offers the standard option of sRGB, greyscale and YCC. As an alternative, a limited form of ICC colour profiles <sup>27</sup> can be provided.<sup>28</sup>
- Support of graphic effects and typography.
  - Vector image options: No.
  - Transparency information: Yes.
  - Option to specify fonts and styles: No.
- Functionality beyond normal display
  - Animation: No (this option is offered in JPEG 2000 Part 3 and 12).
  - Multipage support: No (this option is offered in JPEG 2000 Part 6).
  - Multiresolution: Yes. There is also an option to construct the image proceeding from colour, quality or position.

## 2.7 Conclusion

### Format Description

- Standardization: JPEG 2000 Part 1 has been standardized since 2000 ISO/IEC. Other parts were standardized later or are not yet fully ISO standardized.
- Objective: Offer alternatives for the limited JPEG/JFIF format by using more efficient compression techniques, an option for lossless compression and multiresolution.
- Structure: The basis is a box structure which stores both the header as well as image information.
- Encoding: A six-step process. The most conspicuous is wavelet transformation (step 3) and packetizing (step 6) whereby the codestream is divided into packets and is sorted by resolution, quality, colour or position.

### Consequences for Storage Capacity

- Lossless: Storage gain is approximately 50%.
- Lossy: Storage gain is variable between 91% and 98%.

### Consequences for Image Quality

- Lossless: None.
- Lossy:
  - Some loss of details while using strong compression.
  - No loss of greyscale/colour.
  - Artefacts: Posterizing, pixelation, woolly effect around elements that are rich in contrast with a large amount of compression.
  - PSNR: Currently being investigated.

---

<sup>27</sup> Definition according to the ICC Profile Format Specification ICC.1:1998-09

<sup>28</sup> Florida Digital Archive description of JP 2 – section 1.8:

[http://www.fcla.edu/digitalArchive/pdfs/action\\_plan\\_bgrounds/jp2\\_bg.pdf](http://www.fcla.edu/digitalArchive/pdfs/action_plan_bgrounds/jp2_bg.pdf)

## **Consequences for the Long-Term Sustainability**

- Lossless: File Format Assessment Method score 74,7.
- Lossy: File Format Assessment Method score 66,1.
- Main problem: Low adoption rate.

## **Consequences for the Functionality**

The most important advantages:

- Possibility of lossless and variable lossy compression.
- Very effective wavelet compression.
- Very comprehensive multiresolution options: It is possible to create the image based on quality, resolution, colour and position.
- Comprehensive metadata possibilities.
- Options for very diverse bit depths (1 to 38 bits).

The most important disadvantages:

- Low adoption rate on the consumer market.
- Low adoption rate in software support (image editing and viewing software).
- No browser support (other than through server side software that generates JPEG files on-the-fly JPEG).
- Compressing and decompressing takes a relatively large amount of computing power.
- No standard option for adding EXIF metadata.

## **Recommendation**

### *Reason 1: Substitution*

JPEG 2000 Part 1 lossless is a good alternative from the perspective of long-term sustainability. The most effective lossless compression (50%), no loss of image quality and the flexible nature of the file format (particularly due to the wealth of multiresolution options) are an extra argument that speaks in favour of JPEG 2000 lossless. The only real long-term worry is the low rate of adoption.

Due to the irreversible loss of image information, JPEG 2000 Part 1 lossy is a much less obvious choice for substitution. The creation of visual lossless images might be considered (i.e., images that cannot visually be differentiated from the original uncompressed file) (storage gain of approximately 90%). In the latter case, it must be understood that visual lossless is a relative term – it is based on the current generation of monitors and the subjective experience of individual viewers.

### *Reason 2: Redigitisation Is Not Desirable*

In this case JPEG 2000 Part 1 lossy, in the visual lossless mode, is a viable option. The small amount of information loss can be defended more easily in this case because there is no substitution.

### *Reason 3: Master file is access file*

In this case JPEG 2000 Part 1 lossy with a larger degree of compression is self-evident. The advanced JPEG 2000 compression technique enables more storage reduction without much loss of quality (superior to JPEG). When selecting the amount of compression, the type of material must be taken into account. Compression artefacts will be more visible in text files than in continuous tone originals such as photos, for example. However, the question is whether the more efficient compression and extra functionality options of JPEG 2000 outweighs the JPEG format for this purpose, which is comprehensively supported by software (including browsers) and is widely distributed.



## 3 PNG

### 3.1 What Is PNG?

PNG (Portable Network Graphics) is a datastream and an associated file format for a lossless compressed, portable, individual raster image<sup>29</sup> which was initially developed for transmission via the Internet. A large group of developers (PNG development group) began developing the format in 1995 under the supervision of the World Wide Web Consortium (W3C) as an alternative for the then-patented GIF format and associated LZW compression. The first official version, 1.0, came into existence in 1997 as a W3C Recommendation. PNG version 1.2 was revealed in 1999, and this version has been ISO standardized (ISO/IEC 15948:2003) since 2003, with the specifications being freely available via the W3C: <http://www.w3.org/TR/PNG/><sup>30</sup>.

The objectives of the developers of PNG were as follows<sup>31</sup>:

- a. Portability: Encoding, decoding, and transmission should be software and hardware platform independent.
- b. Completeness: It should be possible to represent true colour, indexed-colour, and greyscale images, in each case with the option of transparency, colour space information, and ancillary information such as textual comments.
- c. Serial encode and decode: It should be possible for datastreams to be generated serially and read serially, allowing the datastream format to be used for on-the-fly generation and display of images across a serial communication channel.
- d. Progressive presentation: It should be possible to transmit datastreams so that an approximation of the whole image can be presented initially, and progressively enhanced as the datastream is received.
- e. Robustness to transmission errors: It should be possible to detect datastream transmission errors reliably.
- f. Losslessness: Filtering and compression should preserve all information.
- g. Performance: Any filtering, compression, and progressive image presentation should be aimed at efficient decoding and presentation. Fast encoding is a less important goal than fast decoding. Decoding speed may be achieved at the expense of encoding speed.
- h. Compression: Images should be compressed effectively, consistent with the other design goals.
- i. Simplicity: Developers should be able to implement the standard easily.
- j. Interoperability: Any standard-conforming PNG decoder shall be capable of reading all conforming PNG datastreams.
- k. Flexibility: Future extensions and private additions should be allowed for without compromising the interoperability of standard PNG datastreams.
- l. Freedom from legal restrictions: No algorithms should be used that are not freely available.

---

<sup>29</sup> In contrast to the GIF format, PNG does not offer any animation options (animated GIF). The separate MNG format was created for animation objectives: <http://www.libpng.org/pub/mng/>.

<sup>30</sup> What is strange is that the ISO itself mentions the year 2004:

[http://www.iso.org/iso/iso\\_catalogue/catalogue\\_tc/catalogue\\_detail.htm?csnumber=29581](http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=29581).

<sup>31</sup> These objectives are listed in the W3C specifications under the “Introduction” header: <http://www.w3.org/TR/PNG/> henceforth: *PNG specs*.

These objectives have been achieved in the last PNG standard.

## 3.2 How does it work?

### 3.2.1 Structure

The PNG datastream consists of a PNG signature<sup>32</sup> which indicates that it is a PNG datastream, followed by a sequence of chunks (meaning a “component”). Each chunk has a chunk type that specifies the goal. A certain number of chunks is mandatory (*critical*), and a large part is inessential (*ancillary*). This chunk structure was developed with the idea of keeping the format expandable while simultaneously being backwards compatible.

### 3.2.2 Encoding and Decoding/Filtering and Compression

Encoding takes place in six steps<sup>33</sup>:

1. Pass extraction: To allow for progressive display, the PNG image pixels can be rearranged to form several smaller images called reduced images or passes.
2. Scanline serialization: The image is serialized one scanline at a time. Pixels are ordered left to right in a scanline and scanlines are ordered top to bottom.
3. Filtering: Each scanline is transformed into a filtered scanline using one of the defined filter types to prepare the scanline for image compression.
4. Compression: Occurs on all the filtered scanlines in the image.
5. Chunking: The compressed image is divided into conveniently sized chunks. An error detection code is added to each chunk.
6. Datastream construction: The chunks are inserted into the datastream.

Only filtering and compression are described below.

Prior to compression, compression filters are used that order the bytes per scanline. A different filter can be used per scanline. This greatly increases the success of the compression. The PNG compression algorithm uses the lossless, unpatented inflate/deflate method (zlib/gzlib).<sup>34</sup>

The success of the compression depends on the correct and complete implementation of the PNG encoding options. It can be useful to research software tools for limiting PNG file sizes.<sup>35</sup>

## 3.3 Consequences for the Required Storage Capacity

Based on test sets, it appears that PNG in lossless mode can yield a benefit of about 40% compared to an uncompressed file. Further tests with more refined compression options must prove whether or not this can still be optimized.

---

<sup>32</sup> See section 5.2 of the *PNG specs* and Wikipedia:

[http://en.wikipedia.org/wiki/Portable\\_Network\\_Graphics#File\\_header](http://en.wikipedia.org/wiki/Portable_Network_Graphics#File_header).

<sup>33</sup> *PNG specs*, section 4.5.1: <http://www.w3.org/TR/PNG/#4Concepts.EncodingIntro>.

<sup>34</sup> See chapters 9 and 10 of the *PNG specs* and Wikipedia:

[http://en.wikipedia.org/wiki/Portable\\_Network\\_Graphics#Compression](http://en.wikipedia.org/wiki/Portable_Network_Graphics#Compression).

<sup>35</sup> Wikipedia PNG lemma gives recommendations for various tools

[http://en.wikipedia.org/wiki/Portable\\_Network\\_Graphics#File\\_size\\_and\\_optimization\\_software](http://en.wikipedia.org/wiki/Portable_Network_Graphics#File_size_and_optimization_software).

### 3.4 Consequences for the Image Quality

Because PNG filtering and compression is lossless there is no degradation of the image quality. However, the assumption is that the bit depth remains the same as that of the source file. Decrease of the bit depth – an option that the PNG format offers – must be viewed as a form of *lossy* compression.

### 3.5 Consequences for the Long-Term Sustainability

Applying the “File Format Assessment Method” to the “PNG 1.2” format results in a score 78 on a scale of 0-100. When the four formats that are compared in this report are sorted from most to least suitable for long-term storage according to the named method, “PNG 1.2” ends up in first place with this score, directly ahead of “JP2 (JPEG 2000 Part 1) lossless.” In the PNG case, too, the low adoption rate of the format has a negative effect on the final score. As mentioned earlier in section 1.5 (“Consequences for the Long-Term Sustainability” for JPEG 2000), despite the fact that PNG scores four point higher than JP2 (JPEG 2000 Part 1) lossless” we still prefer the latter on account of this format's better future outlook as regards adoption.

### 3.6 Consequences for the Functionality

- Options for including bibliographic and technical (EXIF) metadata
  - Bibliographic metadata: PNG offers the option of including content metadata in both ASCII as well as UTF-8 and offers a number of standard options (Title, Author, Description, Copyright, Creation Time, Software, Disclaimer, Warning, Source, Comment). It is possible to expand this set according to your own wishes.<sup>36</sup>
  - Technical metadata: PNG does not (yet) support EXIF information (technical metadata that provides information about the camera and camera settings).
- Suitability of the format for offering it as a high-resolution access master
  - Browser support: Yes.
  - High-resolution image access: In theory, yes. Through lossless compression, PNG remains relatively large for this objective.
- Maximum size
  - Topic of investigation.

#### LOC Quality and Functionality Factors:<sup>37</sup>

- Normal display
  - Screen display
  - Printable: Yes.
  - Zoomable: Yes.
- Clarity
  - High-resolution options: Yes.
  - Bit depths: Can vary from 1 to 16 bits per channel.
- Colour maintenance

---

<sup>36</sup> See section 11.3.4.2 of the *PNG specs*.

<sup>37</sup> [http://www.digitalpreservation.gov/formats/content/still\\_quality.shtml](http://www.digitalpreservation.gov/formats/content/still_quality.shtml).

- Support of various colour spaces: Yes (though not via ICC profile).
- Option for including gamma (brightness) correction: Yes (also chroma – colour saturation - correction).
- Options for including ICC colour profiles: PNG offers the option of using the sRGB colour space and including ICC colour profiles.<sup>38</sup>
- Support of graphic effects and typography.
  - Vector image options: No.
  - Transparency information: Yes.
  - Option to specify fonts and styles: No.
- Functionality beyond normal display
  - Animation: No<sup>39</sup>
  - Multipage support: No.
  - Multiresolution: No.

### 3.7 Conclusion

#### Format Description

- Standardization: PNG 1.2 has been ISO/IEC standardized since 2003.
- Objective: Follow up of the patented and limited GIF format, with a wealth of options as regards progressive structure, transparency, lossless compression and expansion of the standard.
- Structure: Chunks are the basis, which store both the header as well as image information.
- Encoding: A six-step process. What is notable is the option to apply separate filtering per scanline (thus increasing the effectiveness of the compression).

#### Consequences for Storage Capacity

- Storage gain is approximately 40%.

#### Consequences for Image Quality

- Lossless, so none.

#### Consequences for the Long-Term Sustainability

- File Format Assessment Method score 78.
- Main problem: Low adoption rate.

#### Consequences for the Functionality

The most important advantages:

- Lossless compression.
- Comprehensive support by image editing and viewer software and browsers.
- Comprehensive metadata possibilities.
- Options for very diverse bit depths (1 to 16 bits per channel).
- Comprehensive options for transparency.

The most important disadvantages:

---

<sup>38</sup>See section 4.2 of the *PNG specs*.

<sup>39</sup>The related MNG format offers this option: <http://www.libpng.org/pub/mng/>



- No option for lossy compression (other than by decreasing the bit depth), so images remain relatively large.
- No multiresolution options.
- No standard option for adding EXIF metadata.

## **Recommendation**

### *Reason 1: Substitution*

PNG Part 1 lossless is a possible alternative from the perspective of long-term sustainability. Lossless compression is ideal for substitution objectives because no image information is lost. The compression is somewhat less effective than that of JPEG 2000 Part 1 lossless (40% versus 50%). The comprehensive software support is a plus but the low level of actual use (both on the consumer market as well as in the cultural heritage sector) is worrisome.

### *Reason 2: Redigitisation Is Not Desirable*

PNG is also suitable for this goal, although the less effective, lossless compression is a minus.

### *Reason 3: Master file is access file*

In this case, PNG is a less obvious choice due to the lack of a lossy compression option (and thus more storage gain).

## 4 JPEG

### 4.1 What is JPEG?

First and foremost, JPEG (Joint Photographic Experts Group) stands for the committee that was established to create a standard for the compression of continuous tone greyscale and colour images (as the name indicates).<sup>40</sup> The committee started this task in 1986, and in 1992 the first version of this standard was ready, which in 1994 was standardized as ISO 10918-1 and as ITU-T Recommendation T.81. The JPEG committee is also at the basis of the JBIG format (bitonal compression format) and the JPEG 2000 format.

The JPEG standard is more than a description of a file format: It both specifies the codec with which the images are compressed/encoded in a datastream as well as the file format that this datastream contains.

The JPEG standard consists of four parts<sup>41</sup>:

- Part 1 - The basic JPEG standard, which defines many options and alternatives for the coding of still images of photographic quality.
- Part 2 - Sets rules and checks for making sure software conforms to Part 1.
- Part 3 - Set up to add a set of extensions to improve the standard, including the SPIFF file format.
- Part 4 - Defines methods for registering some of the parameters used to extend JPEG.

The description of the JPEG file format – the *JPEG Interchange Format* – is included as annex B of the 10981-1 standard. The confusing part is that a stripped, or real-world, version<sup>42</sup> of this description, JFIF (*JPEG File Interchange Format*), has become the de-facto standard with which applications work and which is generally designated as JPEG<sup>43</sup>. JFIF simplified a number of things in the standard – among them a standard colour space – and thus made the JPEG Interchange Format usable for a range of applications and uses.

The following discusses the JFIF standard, which will be designated as JPEG.<sup>44</sup>

---

<sup>40</sup> To cite the group: “This Experts Group was formed in 1986 to establish a standard for the sequential progressive encoding of continuous tone greyscale and colour images.” CCITT T.81 *Information Technology – Digital compression and coding of continuous-tone still images – requirements and guidelines* p. 1. <http://www.w3.org/Graphics/JPEG/itu-t81.pdf>.

<sup>41</sup> <http://www.jpeg.org/jpeg/index.html>.

<sup>42</sup> <http://www.jpeg.org/jpeg/index.html>. The JFIF format was developed by Eric Hamilton of C-Cube Microsystems.

<sup>43</sup> JFIF standard: <http://www.jpeg.org/public/jfif.pdf>. The “real world” terminology was coined by the JPEG committee itself: “As well as the standard we created, nearly all of its real world applications require a file format, and example reference software to help implementors”. <http://www.jpeg.org/jpeg/index.html>.

<sup>44</sup> Five other extensions of the standard are worth mentioning:

- JPEG\_EXIF (most recent version 2.2). This is an extension of the JPEG standard (based on the baseline JPEG) that is used en masse in digital cameras. EXIF information contains technical metadata about the camera and camera settings.
- Adobe JPEG. The version of JPEG as used by Adobe applications does comply with the JPEG standard but not with JFIF. An important difference with the JFIF standard is in the fact that Adobe can save JPEG files in the CMYK colour space. This version of JPEG is *not* publicly documented. Florida Digital Archive description of JFIF (p. 5, 6): [http://www.fcla.edu/digitalArchive/pdfs/action\\_plan\\_bgrounds/jfif.pdf](http://www.fcla.edu/digitalArchive/pdfs/action_plan_bgrounds/jfif.pdf).

## 4.2 How does it work?

### 4.2.1 Structure

Topic of investigation.

### 4.2.2 Encoding and Decoding/Filtering and Compression

Encoding (assuming a 24-bit RGB file) takes place in four steps<sup>45</sup>:

1. Conversion of the RGB colour space of the source file to the YCbCr colour space (Y is the brightness component, Cb and Cr two colour or chroma components, blue and red).<sup>46</sup>
2. The resolution of the colour data is decreased (also: downsampling or chroma subsampling), mostly with a factor of two. This is based on the fact that the human eye sees more details in the brightness component Y than in the colour components, Cb and Cr. This can already yield a 33 to 50% gain compared to the source file and is a lossy process.
3. The image is divided into 8 x 8 pixels (block splitting). For each block, for each of the Y, Cb and Cr components, so-called discrete cosine transformation (DCT) is applied (a break/conversion of the pixel values per component to a frequency-domain representation).
4. The amplitudes of the frequency components are quantified. Because the eye is less sensitive to high-frequency variations in brightness (than for small changes in colour or brightness in large/wide areas), high-frequency components are therefore stored with a lower degree of accuracy. The quality setting of the encoder determines how much of the high-frequency information is lost. In the case of extreme compression, this information is completely left out.
5. Ultimately, the 8 x 8 blocks are compressed even more by means of a lossless algorithm (a version of Huffman encoding).

Decoding simply runs in the opposite direction.

The most noticeable JPEG artefact – the pixelation – occurs in the quantifying step.

- 
- SPIFF based on Annex F of the ISO10918-1 standard. This is an extension of the JPEG standard (Part 3) and is intended as the successor of the limited JFIF standard (JFIF inventor Eric Hamilton played an important role in the development of SPIFF) which can contain both the JPEG-DCT compression as well as JBIG bitonal compression. However, this appears to be rarely used. See: <http://www.fileformat.info/format/spiff/egff.htm> en <http://www.digitalpreservation.gov/formats/fdd/fdd000019.shtml>.
  - Lossless JPEG. Two formats of the lossless JPEG have been developed:
    - Lossless-JPEG (1993) as an extension of the JPEG standard, which uses a completely different compression technique. This has not gained popularity, other than in some medical applications.
    - JPEG-LS (1999). A “near-lossless” format that offers better compression than the lossless JPEG format and is much less complex than the lossless JPEG 2000 version. This also appears to have hardly gained any popularity.

Wikipedia: [http://en.wikipedia.org/wiki/Lossless\\_JPEG](http://en.wikipedia.org/wiki/Lossless_JPEG). JPEG-LS homepage: <http://www.jpeg.org/jpeg/jpegls.html>.

<sup>45</sup> Wikipedia: <http://en.wikipedia.org/wiki/JPEG#Encoding>. The encoding described here is the most used method. This is thus not the *only* method.

<sup>46</sup> This step is sometimes skipped. This is the case for a high-quality JPEG, whereby the file is stored in sRGB “where each colour plane is compressed and quantized separately with similar quality steps.” Wikipedia [http://en.wikipedia.org/wiki/JPEG#Color\\_space\\_transformation](http://en.wikipedia.org/wiki/JPEG#Color_space_transformation).

### 4.3 Consequences for the Required Storage Capacity

Based on the test material, it appears that the gain that can be achieved with JPEG compression can vary – assuming Adobe Photoshop compression values JPEG 10 to JPEG PSD 1 – between 90 and 98%.

### 4.4 Consequences for the Image Quality

Five variations of JPEG compression were tested in Photoshop (scale 0-12, designated as *PSD*): PSD 0, 3, 5, 8 and 10. 0 and 3 are extreme compression, 5 is average, 8 and 10 are slight.

#### *Detailed loss – MTF*

Original TIFF (QA-62 SFR and OECF test chart): MTF 5.91/5.91. File size 4.7 MB

<b>Compression Ratio</b>	<b>MTF Horizontal and Vertical (three RGB channels on average)</b>	<b>File Size</b>
JPEG PSD 10	5.9 / 5.8	204 KB
JPEG PSD 8	5.4 / 5.2	128 KB
JPEG PSD 5	4.9 / 4.8	84 KB
JPEG PSD 3	4.3 / 4.2	64 KB
JPEG PSD 0	3.8 / 3.5	57 KB

#### *Greyscale and Colour Loss*

No measurable loss of greyscale in Kodak Gray.

The delta E values remain the same at various compression values and no extra colour shift occurs (in the contrary = the RGB values are drawn to each other to one value)<sup>47</sup>.

#### *Artefacts*

In JPEG files, three clearly visible artefacts appear the more the compression increases (tested based on various types of materials):

1. Posterizing or banding (coarse transitions in colour or greyscale). Somewhat visible starting at JPEG PSD 7/8. Clearly visible starting approximately at JPEG PSD 5.
2. Pixelation: Visible starting approximately at JPEG PSD 2.
3. Woolly effect around elements rich in contrast. Visible starting approximately at JPEG PSD 4.

The last effect is particularly visible in text (around the letters). Continuous tone originals such as photos and paintings appear to be more suitable for strong JPEG compression than text materials do (or other materials with high-contrast transitions such as line drawings).

#### *PSNR*

Topic of investigation.

---

<sup>47</sup> This is why delta E might not be a good tool for measuring colour differences between the compressed and uncompressed file. Colour differences *do* occur in distorting subtle colour changes (see “artefacts”).

### *Consequences of Repeated Compression*

The image degrades when it is compressed several times. Tests have shown that degradation when applying JPEG PSD 10 compression doesn't really become visible until compression has been executed four times.

#### **4.5 Consequences for the Long-Term Sustainability**

Application of the "File Format Assessment Method" to the "basic JFIF (JPEG) 1.02" format results in a score of 65,4 on a scale of 0-100. When the four formats that are compared in this report are sorted from most to least suitable for long-term storage according to the named method, "basic JFIF (JPEG) 1.02" ends up in third place with this score, not much ahead of or almost equal with "TIFF 6.0 with LZW compression" with a score of 65,3 and just beneath "JP2 (JPEG 2000 Part 1) lossy," which scores 66,1 points. The lossy form of the compression and the fact that the format is little used as a master format in the cultural heritage sector both play an important role in the final score of the format. If a choice has to be made between "JP2 (JPEG 2000 Part 1) lossy" and "basic JFIF (JPEG) 1.02," preference is given to the latter due to the more certain future of this file.

#### **4.6 Consequences for the Functionality**

- Options for including bibliographic and technical (EXIF) metadata
  - Content-related metadata: Yes.
  - Technical metadata: The separate JPEG EXIF format was developed for the inclusion of EXIF information (see note 35).
- Suitability of the format for offering it as a high-resolution access master
  - Browser support: JPEG is supported by all standard browsers.
  - High-resolution image access: Often the high-resolution JPEG is used as a zoom file. This is done by creating separate resolution layers, as separate images. Sometimes these images again parted into tiles.<sup>48</sup>
- Maximum size
  - Topic of investigation.

#### **LOC Quality and Functionality Factors:** <sup>49</sup>

- Normal display
  - Screen display: Yes.
  - Printable: Yes.
  - Zoomable: Yes.
- Clarity
  - High-resolution options: Yes. A lot of compression can damage detailing (see section 3.4 above).
  - Bit depths: Limited to 8 and 24 bits.<sup>50</sup>

---

<sup>48</sup> See the Geheugen van Nederland (memory of the Netherlands) (<http://www.geheugenvannederland.nl/>) for the first and the solution by the image database of the Amsterdam City Archive (<http://beeldbank.amsterdam.nl/>) for the second.

<sup>49</sup> [http://www.digitalpreservation.gov/formats/content/still\\_quality.shtml](http://www.digitalpreservation.gov/formats/content/still_quality.shtml).

- Colour maintenance
  - Support of various colour spaces: Yes (though not via ICC profile).
  - Option for including gamma correction: No.
  - Options for including ICC colour profiles: Yes<sup>51</sup>
- Support of graphic effects and typography.
  - Vector image options: No.
  - Transparency information: Yes.
  - Option to specify fonts and styles: No.
- Functionality beyond normal display
  - Animation: No.
  - Multipage support: No.
  - Multiresolution: More or less. It is possible to store thumbnails with larger images<sup>52</sup>. However, this function is not or rarely supported by image editing and viewer software.

## 4.7 Conclusion

### Format Description

- Standardization: The JPEG standard has been ISO/IEC (10918-1) standardized since 1994. An extension of Annex B of the standard – JFIF – has become the de facto standard and is simply designated as JPEG.
- Objective: To create a standard for the compression of continuous tone greyscale and colour images.
- Structure: Topic of investigation.
- Encoding: A five-step process. Most noteworthy is the use of the DCT compression technique.

### Consequences for Storage Capacity

- Storage gain is variable between approximately 89% and 96%.

### Consequences for Image Quality

- Gradual loss of detail with increased compression.
- No measurable loss of greyscale/colour.
- Artefacts: Visible posterizing, pixelation, woolly effect around elements that are rich in contrast with a large amount of compression.
- PSNR: Topic of investigation.

### Consequences for the Long-Term Sustainability

---

<sup>50</sup> A 12-bit JPEG is used in some medical applications. The 12-bit JPEG is a part of the JPEG standard but is rarely used and supported. Wikipedia JPEG:

[http://en.wikipedia.org/wiki/JPEG#Medical\\_imaging:\\_JPEG.27s\\_12-bit\\_mode](http://en.wikipedia.org/wiki/JPEG#Medical_imaging:_JPEG.27s_12-bit_mode).

<sup>51</sup> ICC profile 4.2.0.0. LOC description.

<http://www.digitalpreservation.gov/formats/fdd/fdd000018.shtml#factors>.

<sup>52</sup> Starting with version 1.02 LOC description JFIF

<http://www.digitalpreservation.gov/formats/fdd/fdd000018.shtml#factors>.

- File Format Assessment Method score 65,4.
- Main problems: Lossy compression and slight use as a master format in the cultural heritage sector.

### **Consequences for the Functionality**

The most important advantages:

- Comprehensive support by image editing and viewer software and browsers.
- Compression and decompression requires little computing power.
- Efficient, variable DCT compression.
- Standardized method for accommodating EXIF metadata (in JPEG EXIF format).

The most important disadvantages:

- No options for lossless compression.
- Limited bit depth options (8 bits greyscale, 24 bits colour).
- No multiresolution options.

### **Recommendation**

#### *Reason 1: Substitution*

JPEG is not the most obvious file format choice for substitution purposes. In particular the irreversible loss of image information is not desirable in view of long-term storage. The relatively low File Format Assessment Method score (66) stems from this fact. One option to consider could be the creation of visual lossless images – JPEG PSD 10 and higher (storage gain approx. 89%). In the latter case, it must be understood that visual lossless is a relative term – it is based on the current generation of monitors and the subjective experience of individual viewers.

#### *Reason 2: Redigitisation Is Not Desirable*

In this case a visual lossless JPEG is a viable option. The small amount of information loss can be defended more easily in this case because there is no substitution. The comprehensive distribution and support of JPEG is an extra argument that speaks in favour of this file format.

#### *Reason 3: Master file is Acces File*

In this case JPEG with a larger degree of compression is self-evident. The JPEG compression technique enables a rather large decrease in storage without much loss of quality. When selecting the amount of compression, the type of material must be taken into account. Compression artefacts in text files will be visible before those in continuous tone originals such as photos, for example.



## 5 TIFF LZW

### 5.1 What is TIFF LZW?

Strictly speaking, TIFF LZW is not a separate file format. TIFF (Tagged Image File Format) 6.0 is the file format, LZW (Lempel-Ziv-Welch, the names of the developers) is the compression algorithm that is used *within* TIFF (in addition to LZW compression, TIFF offers the option of using ITU\_G4, JPEG and ZIP compression). The following provides a brief description of the TIFF 6.0 format, with a more detailed discussion of the LZW compression method.

The first version of the TIFF specification (developed by Microsoft and Aldus, with the last version currently being a part of Adobe) appeared in 1986 and unofficially is called version 3.0. Version 4.0 was launched in 1987 and version 5.0 in 1988. The latter offered options for limited colour space (palette colour) and LZW compression. The baseline TIFF 6.0 standard dates from 1992, which included CYMK colour definition and the use of JPEG compression, among other things. Version 6.0 was followed by various extensions (see section 4.2.1 below) – the most important ones being: TIFF/EP (2001), TIFF/IT (2004), DNG (2005) and EXIF.<sup>53</sup>

The baseline TIFF 6.0 is not ISO-IEC standardized.

The objective was to create a file format to store raster images originating from scanners and image editing software. The main objective “is to provide a rich environment within which applications can exchange image data. This richness is required to take advantage of the varying capabilities of scanners and other imaging devices”<sup>54</sup>. The standard must also be expandable based on new imaging requirements: “A high priority has been given to structuring TIFF so that future enhancements can be added without causing unnecessary hardship to developers”<sup>55</sup>. This option has been abundantly used. The disadvantage of this is that not all extensions are used by all image editing and viewer software.

The LZW compression algorithm dates from 1984 and is basically an improved version of the LZ78 algorithm from 1978. The name gives Jacob Ziv and Abraham Zempel developed the LZ78 format, and Terry Welch developed the faster, improved LZW. It was developed as a lossless data (thus not only for images) compression algorithm. In addition to being used in TIFF, LZW became famous largely due to its use in the GIF format. In addition, LZW is notorious due to the patent that Unisys claimed to have on the algorithm (via developer Terry

---

<sup>53</sup> TIFF/EP extension (ISO 12234-2) for digital photography ([http://en.wikipedia.org/wiki/ISO\\_12234-2](http://en.wikipedia.org/wiki/ISO_12234-2))  
TIFF/IT (ISO 12369) extension for prepress purposes  
(<http://www.digitalpreservation.gov/formats/fdd/fdd000072.shtml>).  
DNG Adobe TIFF UNC extension for storing RAW images  
(<http://www.digitalpreservation.gov/formats/fdd/fdd000188.shtml>).  
EXIF technical metadata of cameras and camera settings  
(<http://www.digitalpreservation.gov/formats/fdd/fdd000145.shtml>).

<sup>54</sup> TIFF Revision 6.0 June 1992. p. 4. Scope. <http://partners.adobe.com/public/developer/en/tiff/TIFF6.pdf>.

<sup>55</sup> Ibidem.

Welch<sup>56</sup>). This patent expired in 2003 (US) and 2004 (Europe and Japan), although Unisys still claims to possess certain improvements to the algorithm.<sup>57</sup>

## 5.2 How does it work?

### 5.2.1 Structure

The TIFF file begins with an 8-byte image file header (IFH) that refers to the image file directory (IFD) with the associated bitmap. The IFD contains information about the image in addition to pointers to the actual image data.<sup>58</sup>

The TIFF tags, which are contained in the header and in the IFDs, contain basic geometric information, the manner in which the image data are organized and whether a compression scheme is used, for example. An important part of the tags belongs to the so-called baseline TIFF.<sup>59</sup> All tags outside of this are extended and contain things such as alternative colour spaces (CMYK and CIE Lab) and various compression schemes.<sup>60</sup>

There are also tags called *private tags*. The TIFF 6.0 version offers users the option to use their own tags (and also to develop them through private IFDs<sup>61</sup>), and this is done quite a lot. The above-mentioned TIFF/EP, TIFF/IT make use of this option. Because the used tags are public, there is talk of open extensions. The LOC documentation contains a valuable overview<sup>62</sup> of this extension:

[http://www.digitalpreservation.gov/formats/content/tiff\\_tags.shtml](http://www.digitalpreservation.gov/formats/content/tiff_tags.shtml).

### 5.2.2 Encoding and Decoding/Filtering and Compression

Topic of investigation.

## 5.3 Consequences for the Required Storage Capacity

Based on test sets, it appears that TIFF LZW in lossless mode can yield a benefit of about 30% compared to an uncompressed file.

## 5.4 Consequences for the Image Quality

Because LZW compression is lossless there is no degradation of the image quality.

## 5.5 Consequences for the Long-Term Sustainability

Applying the “File Format Assessment Method” to the “TIFF 6.0 with LZW compression” format results in a score 65,3 on a scale of 0-100. When the four formats that are compared in this report are sorted from most to least suitable for long-term storage according to the named method, “TIFF 6.0 with LZW compression” ends up in last place with this score, not far behind or almost equal with “basic JFIF (JPEG) 1.02,” with a score of 65,4.

---

<sup>56</sup> As an employee of the Sperry Corporation, Welch developed the algorithm, and that is what the patent was initially based on. Sperry Corporation later became a part of Unisys.

[http://en.wikipedia.org/wiki/Graphics\\_Interchange\\_Format#Unisys\\_and\\_LZW\\_patent\\_enforcement](http://en.wikipedia.org/wiki/Graphics_Interchange_Format#Unisys_and_LZW_patent_enforcement).

<sup>57</sup> [http://www.unisys.com/about\\_unisys/lzw/](http://www.unisys.com/about_unisys/lzw/).

<sup>58</sup> A TIFF can contain several IFDs – this is then a multipage TIFF (not a baseline TIFF).

<sup>59</sup> Part 1 from the TIFF 6.0 specs: <http://partners.adobe.com/public/developer/en/tiff/TIFF6.pdf>.

<sup>60</sup> Part 2 from the TIFF 6.0 specs: *ibidem*.

<sup>61</sup> The EXIF extension makes use of this option:

[http://www.digitalpreservation.gov/formats/content/tiff\\_tags.shtml](http://www.digitalpreservation.gov/formats/content/tiff_tags.shtml).

<sup>62</sup> Which is strangely enough not maintained by Adobe itself.

This low score primarily stems from the possible patents that still exist on the LZW compression method (see [http://www.unisys.com/about\\_unisys/lzw/](http://www.unisys.com/about_unisys/lzw/)) and the resulting low rate of adoption of this version of TIFF as a master archive format in the cultural sector. The patents that Unisys still claims to hold are different from the ones that were often referred to in the past and expired in 2003/2004. When we used the same evaluation method to assess a baseline TIFF 6.0, we see a much higher score because LZW compression is not used in this version. Therefore, from the perspective of long-term sustainability use of TIFF 6.0 with LZW compression is discouraged

## 5.6 Consequences for the Functionality

- Options for including bibliographic and technical (EXIF) metadata
  - Content-related metadata: Yes.
  - Technical metadata (EXIF): Yes.
- Suitability of the format for offering it as a high-resolution access master
  - Browser support: No.
  - High-resolution image access: TIFF LZW is very limited when it comes to exchangeability of high-resolution images via the Web. Because the format compresses in a lossless manner the files remain relatively large. TIFF is also not supported by browsers. JPEG thus becomes the more obvious choice.
- Maximum size
  - File size: 4 GB. There are proposals to enlarge this to 20 GB (BigTIFF)<sup>63</sup>

### LOC Quality and Functionality Factors: <sup>64</sup>

- Normal display
  - Screen display: Yes.
  - Printable: Yes.
  - Zoomable: Yes.
- Clarity
  - High-resolution options: Yes.
  - Bit depths: The TIFF 6.0 standard offers the options of 1 bit, 4 bits, 8 bits, 16 bits (and theoretically even 32 bits) per channel.
- Colour maintenance
  - Support of various colour spaces: Yes (though not via ICC profile). Standard: Bitonal, greyscale, RGB, CMYK, YCbCR, CIEL\*a\*b
  - Option for including gamma correction: No.
  - Options for including ICC colour profiles: Yes. ICC colour profiles can be included, although there does not appear to be a standard way for this. The TIFF/EP and TIFF/IT standards developed private tags that can also be

---

<sup>63</sup> <http://www.awaresystems.be/imaging/tiff/big TIFF.html>

Photoshop should be possible to open the 4 GB file

<http://kb.adobe.com/selfservice/viewContent.do?externalId=320005&sliceId=1>

<sup>64</sup> [http://www.digitalpreservation.gov/formats/content/still\\_quality.shtml](http://www.digitalpreservation.gov/formats/content/still_quality.shtml)

included in regular TIFF 6.0 files. Adobe Photoshop, on the other hand, appears to use yet another method.<sup>65</sup>

- Support of graphic effects and typography.
  - Vector image options: No.
  - Transparency information: Yes (through a so-called alpha channel).
  - Option to specify fonts and styles: No.
- Functionality beyond normal display
  - Animation: No.
  - Multipage support: Yes.
  - Multiresolution: TIFF offers the option of multiresolution (Image Pyramid). It is unclear whether this a subsequent addition to the private tags.
  - In any case, it is not a part of the TIFF 6.0 1992 standard (baseline and extended). It is also unclear to which extent this functionality is supported by viewers.

## 5.7 Conclusion

### Format Description

- Standardization: The baseline TIFF 6.0 is not an ISO-IEC standard. The description of the baseline TIFF 6.0 (1992) is freely available on the Adobe website. LZW compression has been a part of the (extended) TIFF standard since version 5.0 (1988).
- Objective: Creation of a rich and extensible file format for raster images.
- Structure: The basis of the file format is formed by the so-called tags located both in the header (IFH) and in the image file directories (IFD).
- LZW encoding: Topic of investigation.

### Consequences for Storage Capacity

- Storage gain is approximately 30%.

### Consequences for Image Quality

- Lossless, so none.

### Consequences for the Long-Term Sustainability

- File Format Assessment Method (lowest score): 65,3
- Main problem: Possible patents on the LZW compression method and the resulting low rate of adoption as a master archive format in the cultural sector.

### Consequences for the Functionality

The most important advantages:

- Lossless compression
- Support of image editing and viewer software
- Comprehensive metadata possibilities
- Options for very diverse bit depths (1 to 16 bits per channel)
- Option for including EXIF information

---

<sup>65</sup> LOC TIFF docu: <http://www.digitalpreservation.gov/formats/fdd/fdd000022.shtml#factors>.

The most important disadvantages:

- No option for lossy compression, which leaves the images relatively large
- No browser software support

### **Recommendation**

#### *Reason 1: Substitution*

TIFF 6.0 LZW is the least desirable option from the perspective of long-term sustainability (the lowest score in the File Format Assessment Method). The uncertainties regarding the patents that appear to exist on the LZW compression method render the choice of TIFF LZW unwise for this objective. Lossless compression LZW is in itself ideal for substitution objectives because no image information is lost. However, the compression is much less effective (30%) than that of JPEG 2000 Part 1 lossless and PNG (50% and 40%, respectively). The comprehensive software support is a plus but the low level of actual use (by both consumers as well as the cultural heritage sector) is worrisome.

#### *Reason 2: Redigitisation Is Not Desirable*

The patents and the less effective lossless compression do not make TIFF LZW an obvious choice for this objective.

#### *Reason 3: Master File is Access File*

The lack of a lossy compression option does not make TIFF LZW an obvious choice for this objective.

## 6 Conclusion

### Description of Formats

Standardization: JPEG 2000, PNG and JPEG are ISO/IEC standardized. TIFF 6.0 is not, though the TIFF 6.0 standard is public and is made available by Adobe.

### Consequences for the Storage Capacity

On the storage test two limitations have been placed:

- Only 24 bit, RGB (8 bit per colour channel) files have been tested
- Only two sets of (about 100) originals have been tested: a set low contrast text material and a set of photographs

File Format	Storage Gain Compared to the Uncompressed TIFF File
JPEG 2000 Part 1 lossless	52%
JPEG 2000 Part 1 lossy	Variable between 91% and 98%
PNG lossless	43%
JPEG lossy	Variable between 89% and 96%
TIFF LZW lossless	30%

Between the two sets of originals no obvious differences in storage gain were found. Is it clear however that high contrast, textual material will yield higher compression profits – this is part of further, future research.

JPEG 2000 Part 1 is obviously the most effective for lossless and lossy compression.

However, JPEG is not really much inferior to lossy JPEG 2000 compression other than that compression artefacts occur earlier than with JPEG 2000 (see below).

### Consequences for Image Quality

Naturally, no loss of image quality occurs with the lossless formats JPEG 2000 Part 1 lossless, PNG and TIFF LZW.

The lossy formats JPEG 2000 Part 1 lossy and JPEG degrade when compression levels are rising.

- The sharpness of JPEG degrades gradually when compression increases. In JPEG 2000, some sharpness deterioration occurs only with extreme compression.
- No measurable loss of greyscale and colour (colour shift and Delta E) is observed for both JPEG and JPEG 2000. However, with increasing compression excessive “simplification” of the colour subtleties occurs which in the most extreme case results in unnatural tone and colour transitions (banding) (this is caused by the quantification step in the encoding process).
- The artefacts that occur with increasing compression in JPEG 2000 and JPEG resemble each other a lot. What is important to note is that the visibility of these artefacts occurs earlier in JPEG than in JPEG 2000.
  - Banding (rough colour or tone transitions)
  - Pixelation (the tiles into which the files are divided become visible)
  - Woolly effect around elements rich in contrast.

*A remaining topic of investigation is the expression of PSNR (Peak Signal-to-Noise Ratio) of the degradation that occurs during lossy compression.*

### **Consequences for the Long-Term Sustainability**

Application of the previously discussed File Format Assessment Method (see the introduction and Appendices 2 and 3) to the image formats discussed in this report, plus the uncompressed TIFF format that has been used until now for the master images, results in the following order in these formats from most to least suitable for long-term storage:

<b>Ranking</b>	<b>Format</b>	<b>Score</b>
1	Baseline TIFF 6.0 uncompressed	84,8
2	PNG 1.2	78,0
3	JP2 (JPEG 2000 Part 1) lossless	74,7
4	JP2 (JPEG 2000 Part 1) lossy	66,1
5	Basic JFIF (JPEG) 1.02	65,4
6	TIFF 6.0 with LZW compression	65,3

The main thing is that from the perspective of long-term sustainability the choice for “Baseline TIFF 6.0 uncompressed” is the safest one. In practice it appears that this is not a viable option due to the large size of the files and the associated high storage costs.

The ‘File Format Assessment Method’ is still in its infancy. Feedback is being awaited from colleague institutions regarding this method. Additionally, not much experience has been had with the application of this method in practice. Based on the experiences gained in this study it appears necessary to adapt the method. It is therefore too early to entirely ascribe the choice of a durable format to this method. The results of the method will be tested against previous knowledge and experiences.

As the above table indicates, the choice for “Baseline TIFF 6.0 uncompressed” is the safest one from the perspective of long-term sustainability. If an alternative format has to be selected, we see that “PNG 1.2” and “JP2 (JPEG 2000 Part 1) lossless” – both lossless compressed formats – are the alternatives. Here we reach a point where the applied method may fall short. In the method, the characteristic “Usage in the cultural heritage sector as master image file” of the Adoption criterion makes a valuable contribution to the total score. However, what is not included in the method at the moment are the prospects for the future of this criterion. Although neither format is currently used on a large scale as a preservation master file in the cultural sector, JPEG 2000 has more potential. PNG has been in existence since 1996 and JP2 since 2000. The preference, for lossless formats, is thus for JPEG 2000.

Another issue that is neglected by the method is the loss of image quality caused by applying lossy compression methods. Although a file that is a qualitatively worse representation of the original can also be stored in the long-term, it is important – certainly if the original cannot be rescanned – to not only consider the use of the digitalized material in the short term but also in the long term. What must be considered in this respect is that a loss of quality which may be deemed acceptable today may no longer be acceptable for future, other uses of the material. For example, you might consider the use of alternative “display” hardware with a better resolution or different scope. From a long-term sustainability perspective, the use of lossy compression algorithms is discouraged. This certainly applies when the objective of digitisation is to replace the original (objective 1, substitution). If a lossy compression method



is selected nevertheless, the use of “basic JFIF (JPEG) 1.02” is recommended due to the more certain future of this format as compared to the lossy JPEG 2000 Part 1 variant.

The ultimate advice, rendered exclusively from the perspective of long-term sustainability and the File Format Assessment Method, for an alternative image format for uncompressed TIFFs comes down to the following list, sorted from most to least suitable:

1. JP2 (JPEG 2000 Part 1) lossless
2. PNG 1.2
3. Basic JFIF (JPEG) 1.02
4. JP2 (JPEG 2000 Part 1) lossy
5. TIFF 6.0 with LZW compression

### Consequences for the Functionality

Only the most relevant functions (for master storage) are listed in the table below.

Functionality	File Format
Lossless compression option	JPEG 2000 Part 1, PNG, TIFF LZW
Lossy compression option	JPEG 2000 Part 1, JPEG
Lossy and lossless compression option	JPEG 2000 Part 1
Option to add bibliographic metadata	JPEG 2000 Part 1, PNG, JPEG, TIFF LZW
Standard way to add EXIF metadata	JPEG, TIFF LZW
Browser support	JPEG, PNG
Multiresolution options (suitability of the file as a high-resolution <i>access</i> master)	JPEG 2000 Part 1, TIFF LZW, to a very slight degree: JPEG
Maximum size	JPEG 2000 Part 1: unlimited ( $2^{64}$ ). PNG: Topic of investigation. JPEG: Topic of investigation. TIFF LZW: 4 GB
Bit depths:	JPEG 2000: 1 to max. 38 bits per channel. Compliance class 2: 16 bits per channel. PNG: 1 to 16 bits per channel. JPEG: 8 bits per channel. TIFF LZW: 1 to 16 bits per channel (theoretically to 32 bits per channel)
Standard support of colour spaces	JPEG 2000 Part 1: bitonal, greyscale, sRGB, palletized/indexed colour space PNG: bitonal, greyscale, sRGB, palletized/indexed colour space JPEG: greyscale, RGB TIFF LZW: Bitonal, greyscale, RGB, CMYK, YCbCR, CIEL*a*b
Option to use ICC profiles	JPEG2000 Part 1, PNG, JPEG, TIFF LZW (although not in a standard manner)
Multipage support	TIFF LZW

## Summary

The table below summarizes all the above information in a matrix. The figures only indicate the order of success in the various parts.

	JPEG 2000 part 1 lossless	JPEG 2000 part1 lossy	PNG lossless	JPEG/JFIF lossy	TIFF LZW lossless
<b>Standardization</b>	5	5	5	5	5
<b>Storage Savings</b>	3	5	2	4	1
<b>Image Quality</b>	5	4	5	3	5
<b>Long-term Sustainability</b>	5	2	4	3	1
<b>Functionality</b>	5	5	4	3	4
<b>Score</b>	<b>23</b>	<b>21</b>	<b>20</b>	<b>18</b>	<b>16</b>

It is noteworthy that JPEG 2000 comes out on top in both the lossless as well as the lossy versions.

The table above does not make a distinction between the three reasons for the long-term storage of master files as mentioned in the introduction. Some of the criteria on the left hand side of the table are less relevant depending on these reasons. In the recommendations below the importance of each of the five criteria are taken into account.

## Recommendations

### *Reason 1: Substitution*

The criteria “Long-term sustainability”, “Standardisation” and “Image Quality” are considered the most important when substitution of the original is the main reason for the long-term storage of the master file. JPEG 2000 Part 1 lossless, closely followed by PNG, are the most obvious choices from the perspective of long-term sustainability. When the storage savings (PNG 40%, JPEG 2000 lossless 53%) and the functionality are factored in, the scale tips in favour of JPEG 2000 lossless. The lossless TIFF LZW is not a viable option due to the slight storage gain (30%) and the low score in the File Format Assessment Method (especially due to patents, resulting in a low score on the “Restrictions on the interpretation of the file format” characteristic).

Due to the irreversible loss of image information, lossy compression is a much less obvious choice for this objective.

The creation of visual lossless images might be considered though. Both JPEG 2000 Part 1 (compression ratio 10, storage gain about 90%) and JPEG (PSD10 and higher, storage savings about 89%) offer options in this respect. In the latter case, it must be understood that visual lossless is a relative term – it is based on the current generation of monitors and the subjective experience of individual viewers. A big advantage of the JPEG file format is the enormous distribution and the comprehensive software support, including browsers.

### *Reason 2: Redigitisation Is Not Desirable*

The criteria “Storage savings” and “Image Quality” are considered the most important when the main reason for the long-term storage of the master files is not wanting to do redigitisation. In this case lossy compression, in the visual lossless mode, is a more viable option. The small amount of information loss can be defended more easily in this case

because there is no substitution. The above mentioned JPEG 2000 lossy and JPEG visual lossless versions are the obvious choices.

However, if absolutely no image information may be lost, then the above-mentioned JPEG 2000 lossless and PNG formats are the two recommended options.

*Reason 3: Master File is Access File*

The criteria “Storage savings” and “Functionality” are considered the most important when using the master file as access file is the main reason for the long-term storage of the master file. In this case a larger degree of lossy compression is self-evident. The two options are then JPEG 2000 Part 1 lossy and JPEG with a higher level of compression. The advanced JPEG 2000 compression technique enables more storage reduction without much loss of quality (superior to JPEG). When selecting the amount of compression, the type of material must be taken into account. Compression artefacts will be more visible in text files than in continuous tone originals such as photos, for example. However, the question is whether the more efficient compression and extra options of JPEG 2000 outweighs the JPEG format for this purpose, which is comprehensively supported by software (including browsers) and is widely distributed.

## Appendix 1: Use of Alternative File Formats

The below list is by no means complete. Its sole objective is to provide an idea of the distribution of the various file formats.

### JPEG 2000

Although there are many institutions that are using JPEG 2000 files as “access copies” and many institutions are investigating the use of JPEG 2000 as an archival format, only one cultural institution has been found to date that has definitively chosen JPEG 2000 as its sole archival format. A topic of investigation is the use of JPEG 2000 in the medical field.

Examples of institutions and companies that use JPEG 2000:

- The British Library is the only institution that has chosen JPEG 2000 as one of its archival format (TIFF is still uses as well): “The DPT have taken the view that since the budget for hard drive storage for this project has already been allocated, it would be impractical to recommend a change in the specifics as far as file format is concerned for this project. As such, we recommend retaining the formats originally agreed in MLB\_v2.doc. These are:
  - Linearized PDF 1.6 files for access, with the “first page” being either the table of contents, or the first page of chapter one, depending on the specifics of the book being scanned.
  - JPEG 2000 files compressed to 70 dB PSNR for the preservation copy.”
  - METS/ALTO3 XML for metadata.

The JP2 files fulfil the role of master file but a lack of industry take-up is a slight concern from a preservation viewpoint. However, the format is well defined and documented and poses no immediate risk.”<sup>66</sup>

The risk of “lack of industry take-up” is thus recognized but is not considered as a large enough threat to prevent a choice for JPEG 2000.

- Library of Congress: Uses JPEG 2000 accessing the American Memory website (<http://memory.loc.gov/ammem/index.html>).
- National Digital Newspaper program (NDNP) (<http://www.loc.gov/ndnp/>) uses uncompressed TIFF 6.0 as master and JPEG 2000 for all derivatives.
- At the National Archives of Japan you can choose between JPEG and JPEG 2000 for accessing objects in the Digital Gallery ([http://jpimg.digital.archives.go.jp/kouseisai/index\\_e.html](http://jpimg.digital.archives.go.jp/kouseisai/index_e.html)). The format in which the master images are stored is unclear.
- Google uses JPEG 2000 in Google Earth and Google Print.
- Second Life uses JPEG 2000.

---

<sup>66</sup> <http://www.bl.uk/aboutus/stratpolprog/ccare/introduction/digital/digpresmicro.pdf>.

- Motion JPEG 2000 (MJ2) is used by the members of Digital Cinema Initiatives (DCI) as a standard for digital cinema. Some of the members of DCI are:
  - Buena Vista Group (Disney)
  - 20th Century Fox
  - Metro-Goldwyn-Mayer
  - Paramount Pictures
  - Sony Pictures Entertainment
  - Universal Studios
  - Warner Bros. Pictures
- The medical arena uses JPEG 2000 quite a lot - see DICOM (<http://medical.nema.org/>).
- Biometrics: e.g. the new German passport contains a chip with biometric data and an image in JPEG 2000.
- Video Surveillance Applications
- The Library and Archives Canada (LAC) conducted a feasibility study regarding the use of JPEG 2000 (<http://www.archimuse.com/mw2007/papers/desrochers/desrochers.html>). Up until now however, a copy in TIFF is being archived as well. This is done as an extra safety net.
- Internet Archive.
- University of Connecticut ([http://charlesolson.uconn.edu/Works\\_in\\_the\\_Collection/Melville\\_Project/index.htm](http://charlesolson.uconn.edu/Works_in_the_Collection/Melville_Project/index.htm)).
- University of Utah (<http://www.lib.utah.edu/digital/collections/sanborn/>).
- Smithsonian Libraries.
- J. Paul Getty.

## **PNG**

- The National Archives of Australia uses PNG as archival format.
- No further cultural heritage institutions were found that use the PNG format as an archival master.

## **JPEG**

- The masters of the newspapers of the Leids Archief (Archive in Leiden) are stored as JPEGs.
- The National Library of the Czech Republic uses high-quality JPEG (PSD 12) files as masters for the Memoria and Kramerius projects.  
<http://www.ncd.matf.bg.ac.yu/casopis/05/Knoll/Knoll.pdf>.

**TIFF LZW**

- The National Archives and Records Administration (NARA) in the U.S. uses TIFF LZW as archival master for their internal digitisation projects.
- No other examples were found.

## Appendix 2: File Format Assessment Method – Output

See Appendix 3 for a description of the method. The criteria, characteristics and weighing factors are not exactly the same as in the IPRES paper in Appendix 3. This is because after presenting the paper at IPRES and gaining more experience in applying the method, it has already appeared necessary to adapt the method.

Raster Images	Weight <sup>67</sup>	Baseline TIFF 6.0 uncompressed		basic JFIF (JPEG) 1.02		JP2 (JPEG-2000 Part 1) lossy compressed		JP2 (JPEG-2000 Part 1) lossless compressed		PNG 1.2		TIFF_LZW 6.0	
		Score	Total	Score	Total	Score	Total	Score	Total	Score	Total	Score	Total
Openness	3												
Standardization	9	1	3	1.5	4.5	2	6	2	6	2	6	1	3
Restrictions on the interpretation of the file format	9	2	6	1	3	1	3	1	3	2	6	1	3
Reader with freely available source	7	2	4.66667 <sup>68</sup>	2	4.66667	2	4.66667	2	4.66667	2	4.66667	2	4.66667
Adoption	2												
World wide usage	4	1	2	2	4	1	2	1	2	1	2	1	2
Usage in the cultural heritage sector as archival format	7	2	7	0	0	0	0	1	3.5	1	3.5	0	0
Complexity	3												
Human readability	3	0	0	0	0	0	0	0	0	0	0	0	0
Compression	6	2	4	0	0	0	0	1	2	1	2	1	2
Variety of features	3	1	1	1	1	1	1	1	1	1	1	1	1
Technical Protection Mechanism (DRM)	5												
Password protection	3	2	1.2	2	1.2	2	1.2	2	1.2	2	1.2	2	1.2
Copy protection	3	2	1.2	2	1.2	2	1.2	2	1.2	2	1.2	2	1.2
Digital signature	3	2	1.2	2	1.2	2	1.2	2	1.2	2	1.2	2	1.2
Printing protection	3	2	1.2	2	1.2	2	1.2	2	1.2	2	1.2	2	1.2

<sup>67</sup> The weights that are assigned to the criteria and their characteristics are not fixed. They depend on the local policy of an institution. The weights that are used in the examples in this paper are the weights as assigned by the KB based on its local policy, general digital preservation literature and common sense.

<sup>68</sup> 4,6667 = 2 (score) \* 7 (weight for the characteristic) / 3 (normalisation factor because there are 3 sub-characteristics for the openness criterion)



Content extraction protection	3	2	1.2	2	1.2	2	1.2	2	1.2	2	1.2	2	1.2
Self-documentation	2												
Metadata	1	2	1	1	0.5	2	1	2	1	1	0.5	2	1
Technical description of format embedded	1	1	0.5	0	0	0	0	0	0	0	0	1	0.5
Robustness	5												
Format should be robust against single point of failure	2	1	0.4	1	0.4	2	0.8	2	0.8	1	0.4	0	0
Support for file corruption detection	2	0	0	0	0	0	0	0	0	0	0	0	0
File format stability	2	2	0.8	2	0.8	2	0.8	2	0.8	2	0.8	2	0.8
Backward compatibility	2	2	0.8	2	0.8	2	0.8	2	0.8	2	0.8	2	0.8
Forward compatibility	2	2	0.8	0	0	0	0	0	0	0	0	2	0.8
Dependencies	4												
Not dependent on specific hardware	8	2	4	2	4	2	4	2	4	2	4	2	4
Not dependent on specific operating systems	8	2	4	2	4	2	4	2	4	2	4	2	4
Not dependent on one specific reader	8	2	4	2	4	2	4	2	4	2	4	2	4
Not dependent on other external resources (font + codecs)	8	2	4	2	4	2	4	2	4	2	4	2	4
Maximum score = 63,667		<b>53.9667</b>		<b>41.6667</b>		<b>42.0667</b>		<b>47.5667</b>		<b>49.6667</b>		<b>41.5667</b>	
Perct of 100		84.7644		65.445		66.0733		74.712		78.0104		65.2879	

## Appendix 3 File Format Assessment Method – Explained

*The following paper, concerning the File Format Assessment Method, was presented, in a slightly different form, on the IPRES Conference 2007 (<http://ipres.las.ac.cn/>) but is not yet published. Some changes have been made in the definitions of the criteria and characteristics after gaining more experience with applying the method and receiving feedback from others.*

# Evaluating File Formats for Long-term Preservation

Judith Rog, Caroline van Wijk  
National Library of the Netherlands; The Hague, The Netherlands  
[judith.rog@kb.nl](mailto:judith.rog@kb.nl), [caroline.vanwijk@kb.nl](mailto:caroline.vanwijk@kb.nl)

## Abstract

*National and international publishers have been depositing digital publications at the National Library of the Netherlands (KB) since 2003. Until recently, most of these publications were deposited in the Portable Document Format. New projects, for example the web archiving project, force the KB to handle more heterogeneous material. Therefore, the KB has developed a quantifiable file format risk assessment method. This method can be used to define digital preservation strategies for specific file formats. The choice for a specific file format at creation time or later in the life cycle of a digital object influences the long-term access to the digital object. The evaluation method contains seven sustainability criteria for file formats that are weighed for importance. There seems to be consensus on the sustainability criteria. However, as the weighing of these criteria is connected to an institution's policy, the KB wonders whether agreement on the relative importance of the criteria can be reached at all. With this paper, the KB hopes to inspire other cultural heritage institutions to define their own quantifiable file format evaluation method.*

## Introduction

Over more than a decade, the Koninklijke Bibliotheek (KB) has been involved with the preservation of digital publications. In 1996, the first agreements were signed with Elsevier Science and Kluwer Academic, international publishers of Dutch origin, on the long-term preservation of their e-journals. In 2002 it was decided that the scope of the e-Depot would be broadened to cover the whole spectrum of international scientific publishing. The e-Depot, the electronic archive the KB uses for the long-term storage and preservation of these journals, became operational in 2003 (National Library of the Netherlands, 2007a). At this moment, the e-Depot holds over 10 million international e-publications. Up until now, the vast majority of the publications in the e-Depot consist of articles from e-journals. For all but a few of these articles the format in which they are published is the Portable Document Format (PDF), ranging from PDF version 1.0 to 1.6. For this reason, the research the KB has done to keep the articles preserved and accessible for future use, focused mainly on PDF. At this moment, however, the scope of the e-Depot is broadened. Apart from the ongoing ingestion of the electronic publications, in the coming five years, data resulting from ongoing projects such as web archiving (Digital Preservation Department KB, 2007b), DARE (Digital Preservation Department KB, 2007c), national e-Depot (KB, 2007d) and several digitisation projects (KB, 2007e) will be ingested in the e-Depot as well. The content from these projects

is very heterogeneous concerning file formats. Even the 'traditional' publications that the publishers are providing are getting more and more diverse. Articles can be accompanied by multi media files or databases that illustrate the research.

This more diverse content forces the KB to reconsider its digital preservation strategy. At the foundation of each strategy is the basic principle that the KB will always keep the original publication. The digital preservation strategy describes what actions (e.g. migration or emulation) the KB undertakes to ensure that these publications are preserved and remain accessible for future use. The strategy also describes which choices to make for specific formats during creation, ingest or at a later stage because choices at each of these stages can influence the sustainability of the file. The current strategy is mainly focused on preserving PDF files, but our strategy will need to cover a much wider variety of formats from now on. Whether preservation actions are needed and which actions are needed, depends among other things on the long-term sustainability of the file format of the publication. But what makes a file format suitable for long-term preservation? The criteria for evaluating file formats have been described by several authors (Folk & Barkstrom, 2002; Christensen, 2004; Brown, 2003; Arms & Fleischhauer, 2005; Library of Congress, 2007). But only very rarely though are these criteria applied to a practical assessment of the file formats (Anderson, Frost, Hoebelheinrich & Johnson, 2005). To apply the sustainability criteria we need to know whether all criteria are equally important or whether some are more important than others. And how do you measure whether, and to what degree the format meets the criteria? The application of the criteria should be quantifiable to be able to compare file formats and to give more insight into the preference for certain file formats for long-term preservation.

The KB has started to develop such a quantifiable file format risk assessment. The file format risk assessment facilitates choosing file formats that are suitable for long-term preservation. This paper describes the file format assessment method that the KB has developed and how it is applied in the preservation strategies at the KB. The KB invites the digital preservation community to start a discussion on sustainability criteria and the importance of each criterion by presenting its file format evaluation method.

## File Format Assessment for Long-term Preservation

### *Methodology*

The general preservation criteria used in the KB's method originate from the aforementioned digital preservation literature. The KB's assessment method does not take into account quality and functionality criteria such as clarity or functionality beyond normal rendering as defined in Arms & Fleischhauer (2005). The KB archives publications which are end products that for example do not need editing functionality after publishing. Also the KB archives the publications for long-term preservation purposes and is not the main point of distribution for these publications. Regular access to and distribution of publications is offered by publisher's websites and university repositories etc. This reasoning might be very specific to the KB and it explains the choice for only applying sustainability criteria in the risk assessment method. In the next sections, the criteria, the weighing of the criteria and an example of the application of the method will be described.

The criteria on which classifications of suitability of file formats from the view point of digital preservation will be based are described below. The criteria form measurable standards by which the suitability of file formats can be assigned. The criteria are broken

down into several characteristics that can be applied to all file formats. Values are assigned to each characteristic. The values that are given differ among file formats. The sustainability criteria and characteristics will be weighed, as the KB does not attribute the same importance for digital preservation planning to all characteristics. The weights that are assigned to the criteria and their characteristics are not fixed. They depend on the local policy of an institution. The weights that are used in the examples in this paper are the weights as assigned by the KB based on its local policy, general digital preservation literature and common sense. The range of values that can be assigned to the characteristics are fixed.

The weighing scale runs from zero to seven. These extremes are arbitrary. Seven is the weight that is assigned to very important criteria from the point of view of digital preservation and zero is the score assigned to criteria that are to be disregarded. The values that are assigned to the characteristics range from zero to two. The lowest numerical value is assigned to the characteristic value that is seen as most threatening to digital preservation and long-term accessibility. This value is zero. The highest numerical value is assigned to the characteristic value that is most important for digital preservation and long-term accessibility. This value is two. The scale from zero to two is arbitrary. The criteria do not all have the same number of characteristics. The total score that is assigned to all characteristics is therefore normalised by dividing the score by the number of characteristics.

By applying the file format assessment method to a file format, the format receives a score that reflects its suitability for long-term preservation on a scale from zero to hundred. The higher the score, the more suitable the format is for long-term preservation. The score a format receives can vary over time. A criterion such as *Adoption* for example is very likely to change over time as a format gets more popular or becomes obsolete.

### *Criteria defined*

The criteria that are used in this methodology are *Openness, Adoption, Complexity, Technical Protection Mechanism (DRM), Self-documentation, Robustness and Dependencies*.

#### Openness

The criterion *Openness* of a file format is broken down into the characteristics *Standardisation, Restrictions on the interpretation of the file format, Reader with freely available source*. These characteristics indicate the relative ease of accumulating knowledge about the file format structure. Knowledge about a file format will enhance the chance of successful digital preservation planning.

#### Adoption

The criterion *Adoption* of a file format has two characteristics: *World wide usage* and *Usage in the cultural heritage sector as archival format*. These characteristics indicate the popularity and ubiquity of a file format. When a specific file format is used by a critical mass, software developers (commercial, non commercial) have an incentive to sustain support for a file format by developing software for the specific file format such as readers and writers. However, as a cultural heritage institution, it is not only important to consider usage in general, but also, and more importantly even, the usage by other cultural heritage institutions that share the same goal of preserving the documents for the long-term.

#### Complexity

The characteristic *Complexity* of a file format is broken down into the characteristics *Human readability*, *Compression*, *Variety of features*. These characteristics indicate how complicated a file format can be to decipher. If a lot of effort has to be put into deciphering a format, and with the chance it will not completely be understood, the format can represent a danger to digital preservation and long-term accessibility.

#### Technical Protection Mechanism (DRM)

The characteristic *Technical Protection Mechanism* of a file format is broken down into the characteristics *Password protection*, *Copy protection*, *Digital signature*, *Printing protection* and *Content extraction protection*. These characteristics indicate the possibilities in a file format to restrict access (in a broad sense) to content. Restricted access to content could be a problem when the digital preservation strategy migration is necessary to provide permanent access to the digital object.

#### Self-documentation

The characteristic *Self-documentation* of a file format is broken down into the characteristics *Metadata* and *Technical description of format embedded*. These characteristics indicate the format possibilities concerning encapsulation of metadata. This metadata can be object specific or format specific. When a format facilitates the encapsulation of object specific information (such as author, description etc.) or format specific information in the header on how to read the format for example, the format supports the preservation of information without references to other sources. The more that is known about a digital object, the better it can be understood in the future.

#### Robustness

The characteristic *Robustness* of a file format is broken down into the characteristics *Robust against single point of failure*, *Support for file corruption detection*, *File format stability*, *Backward compatibility* and *Forward compatibility*. These characteristics indicate the extend to which the format changes over time and the extend to which successive generations differ from each other. Also, this characteristic provides information on the ways the file format is protected against file corruption. A frequently changing format could threaten continuity in accessibility for the long term. Large differences among generations of a file format could endanger this continuity equally. The values for *file format stability* ‘rare release of newer versions’, ‘limited release of newer versions’ and ‘frequent release of newer versions’ correspond to ‘release once in ten years’, ‘release once in five years’ and ‘release once a year’ respectively.

#### Dependencies

The characteristic *Dependencies* of a file format is broken down into the characteristics *Not dependent on specific hardware*, *Not dependent on specific operating systems*, *Not dependent on one specific reader* and *Not dependent on other external resources*. These characteristics indicate the dependency on a specific environment or other resources such as fonts and codecs. A high dependency on a specific environment or on external resources provides a risk for digital preservation and long-term accessibility. External resources could be lost over time and difficult to retain and a high dependency on a specific environment strongly ties the format to a specific time and space.

The full list of criteria, the weights as assigned by the KB, the criteria and their possible values can be found in Appendix I. An example of the file format assessment method applied to MS Word 97-2003 and PDF/A-1 can be found in Appendix II

## Application of File Format Assessments

The KB has defined a digital preservation policy for the content of the e-Depot. This policy is the starting point for digital preservation strategies for the digital objects stored in the e-Depot. A digital preservation strategy starts at creation time of a digital object and defines preservation actions on the object at a later stage in the object's life cycle. The KB will not restrict the use of specific file formats for deposit. Any format in general use can be offered. However, KB does give out recommendations and uses the file format assessment method to define strategies.

During the last decade the KB has carried out many digitisation projects. The development of digitisation guidelines has been part of these projects. These guidelines not only make sure that specific image quality requirements are met. They also ensure that the created master files meet the requirements that the digital preservation department has set for metadata and technical matters such as the use of specific file formats and the use of compression (no compression or lossless compression). A file format evaluation method is essential for making well thought-out choices for specific file formats at creation time of digital objects.

The KB has had a lot of influence on the creation process as the owner of the digitisation master files. However, this is not the case for millions of digital publications that have been and will be deposited by international publishers. The KB does have deposit contracts that contain several technical agreements (e.g. file format in which the publisher chooses to submit the publications). Also, as most publications are deposited in PDF, guidelines for the creation of publications in PDF (Rog, 2007) have been created. The PDF guidelines are related to the standard archiving format PDF/A, but are easier to read for non-technical persons. They contain ten 'rules' for PDF functionality that describe best practices at creation.

As was mentioned before, the deposited publications have been quite homogenous concerning file formats. Most publications have been deposited in PDF version 1.0 to 1.6. The file format assessment method has been used to assess this main format stored for its digital preservation suitability. However, new projects will make the digital content of the archive more heterogeneous in the near future. This will require more elaborated file format evaluations.

One example of the use of file format evaluations for new e-Depot content is the evaluation of formats that are harvested for the DARE project. DARE publications are harvested from scientific repositories such as the Dutch university repositories. Most harvested publications are PDFs, however a small part of the articles are harvested in MS Office document formats such as MS Word and MS PowerPoint and in the WordPerfect format. The concrete result of the use of file format risk assessment at the KB is the decision to normalise MS Office documents and WordPerfect documents to a standard archiving format: PDF/A. MS Word documents score 22% if assessed by the assessment method. PDF/A's assessment score amounts to 89%. The main difference between the formats can be found in the criteria *Openness*, *Adoption* and *Dependencies*. For these three criteria, MS Word does have a considerably lower score than PDF/A-1 has. In accordance with the preservation policy both original and normalised files are stored for long-term preservation purposes.

Interestingly enough, an archival institution that is partner in the National Digital Preservation Coalition (NCDD), does not consider PDF/A suitable for archiving its digital data for the long term. One of its valid arguments for not using PDF/A is that PDF/A does not offer the same editing functionality that is available in datasheets. It would be very interesting to compare the differences among cultural heritage institutions concerning the sustainability criteria and the importance of these criteria. This will be much easier if institutions make their file format evaluation quantifiable.

The biggest challenge for the application of the file format risk assessment in the near future will be the web archiving project. As websites contain many different file formats, this new type of content for the e-Depot will require quite different preservation strategies and plans from the current ones.

## Conclusion and Discussion

This paper describes the file format assessment that was developed by the KB to assess the suitability of file formats for long-term preservation. The suitability is made quantifiable and results in a score on a scale from zero to hundred that reflects the suitability of the format for long-term preservation. Formats can easily be compared to each other. The criteria, characteristics and scores that the formats receive are transparent.

The KB hopes to receive feedback on the methodology from other institutions that have to differentiate between formats to decide which format is most suitable for long-term preservation. There seems to be consensus on the sustainability criteria. However, the KB would like to know whether these criteria are the right ones and whether the possible scores a format can receive on a characteristic offer practical options to choose from. The weighing that can be applied to a criterion is not fixed in the methodology. The weighing can be adjusted to the local policy. Therefore, the KB would like to invite other cultural heritage institutions for a discussion about and preferably a comparison of quantifiable file format risk assessments.

## References

- National Library of the Netherlands (KB). (2007a). *The archiving system for electronic publications: The e-Depot*. Retrieved August 20, 2007, from: <http://www.kb.nl/dnp/e-depot/dm/dm-en.html>
- Digital Preservation Department National Library of the Netherlands (KB). (2007b). *Web archiving*. Retrieved August 20, 2007, from [http://www.kb.nl/hrd/dd/dd\\_projecten/projecten\\_webarchivering.html](http://www.kb.nl/hrd/dd/dd_projecten/projecten_webarchivering.html);
- Digital Preservation Department National Library of the Netherlands (KB). (2007c). *DARE: Digital Academic Repositories*. Retrieved August 20, 2007, from [http://www.kb.nl/hrd/dd/dd\\_projecten/projecten\\_dare-en.html](http://www.kb.nl/hrd/dd/dd_projecten/projecten_dare-en.html)
- National Library of the Netherlands (KB). (2007d). *Online deposit of electronic publications*. Retrieved August 20, 2007, from <http://www.kb.nl/dnp/e-depot/loket/index-en.html>
- National Library of the Netherlands (KB). (2007e) *Digitisation programmes & projects*. Retrieved August 20, 2007, from <http://www.kb.nl/hrd/digi/digdoc-en.html>



- Folk, M. & Barkstrom, B. R. (2002). *Attributes of File Formats for Long-Term Preservation of Scientific and Engineering Data in Digital Libraries*. Retrieved August 20, 2007, from [http://www.ncsa.uiuc.edu/NARA/Sci\\_Formats\\_and\\_Archiving.doc](http://www.ncsa.uiuc.edu/NARA/Sci_Formats_and_Archiving.doc)
- Christensen, S.S. (2004). *Archival Data Format Requirements*. Retrieved August 20, 2007, from [http://netarchive.dk/publikationer/Archival\\_format\\_requirements-2004.pdf](http://netarchive.dk/publikationer/Archival_format_requirements-2004.pdf)
- Brown, A. (2003). *Digital Preservation Guidance Note 1: Selecting File Formats for Long-Term Preservation*. Retrieved August 20, 2007, from [http://www.nationalarchives.gov.uk/documents/selecting\\_file\\_formats.pdf](http://www.nationalarchives.gov.uk/documents/selecting_file_formats.pdf)
- Arms, C. & Fleischhauer, C. (2005). Digital formats: Factors for sustainability, functionality and quality. In *Proceedings Society for Imaging Science and Technology (IS&T) Archiving 2005* (pp. 222-227).
- Library of Congress. (2007). *Sustainability of Digital Formats Planning for Library of Congress Collections*. Retrieved August 20, 2007, from <http://www.digitalpreservation.gov/formats/sustain/sustain.shtml>
- Anderson, R., Frost, H., Hoebelheinrich, N. & Johnson, K. (2005) The AIHT at Stanford University, *D-Lib Magazine* (11), 12. Retrieved 20 August, 2007, from <http://www.dlib.org/dlib/december05/johnson/12johnson.html>
- Rog, J. (2007). *PDF Guidelines: Recommendations for the creation of PDF files for long-term preservation and access*, Retrieved from [http://www.kb.nl/hrd/dd/dd\\_links\\_en\\_publicaties/PDF\\_Guidelines.pdf](http://www.kb.nl/hrd/dd/dd_links_en_publicaties/PDF_Guidelines.pdf)

## Author Biography

Caroline van Wijk (1973) has a BA degree in Art and an MA in Political Science. She finished a Java software engineer training in 2000. Directly after, she had been working at a number of web development companies for well over four years before she joined the KB in 2004. At the KB, she had worked on the pilot project Tiff-archive as the technical project staff member until December 2005. Since 2006, she leads the migration implementation project and takes part in the European project Planets as a digital preservation researcher and work package leader.

Judith Rog (1976) completed her MA in Phonetics/Speech Technology in 1999. After working on language technology at a Dutch Dictionary Publisher she was employed at the National Library of the Netherlands/Koninklijke Bibliotheek (KB) in 2001. She first worked in the IT department of the KB for four years before joining the Digital Preservation Department in 2005. Within the Digital Preservation Department she participates in several projects in which her main focus is on file format research.

## Appendix I

**Table 1: All criteria, weighting factors, characteristics and values that can be applied**

Criterion	Characteristic (weighing factor)	values	
Openness			
	Standardisation (9)		
		2	De jure standard
		1,5	De facto standard, specifications made available by independent organisation
		1	De facto standard, specifications made available by manufacturer only
		0,5	De facto standard, closed specifications
		0	No standard
	Restrictions on the interpretation of the file format (9)		
		2	No restrictions
		1	Partially restricted
		0	Heavily restricted
	Reader with freely available source (7)		
		2	Freely available open source reader
		1	Freely available reader, but not open source
		0	No freely available reader
Adoption			
	World wide usage (4)		
		2	Widely used
		1	Used on a small scale
		0	Rarely used
	Usage in the cultural heritage sector as archival format (7)		
		2	Widely used
		1	Used on a small scale
		0	Rarely used
Complexity			
	Human readability (3)		
		2	Structure and content readable
		1	Structure readable
		0	Not readable
	Compression (6)		
		2	No compression
		1	lossless compression
		0	lossy compressed
	Variety of features (3)		
		2	Small variety of features
		1	Some variety of features
		0	Wide variety of features
Technical Protection Mechanism (DRM)			
	Password protection (3)		
		2	Not possible
		1	Optional
		0	Mandatory
	Copy protection (3)		

Criterion	Characteristic (weighing factor)	values	
		2	Not possible
		1	Optional
		0	Mandatory
	Digital signature (3)		
		2	Not possible
		1	Optional
		0	Mandatory
	Printing protection (3)		
		2	Not possible
		1	Optional
		0	Mandatory
	Content extraction protection (3)		
		2	Not possible
		1	Optional
		0	Mandatory
Self-documentation			
	Metadata (1)		
		2	Possibility to encapsulate user-defined metadata
		1	Possibility to encapsulate a limited set of metadata
		0	No metadata encapsulation
	Technical description of format embedded (1)		
		2	Fully self-describing
		1	Partially self-describing
		0	No description
Robustness			
	Format should be robust against single point of failure (2)		
		2	Not vulnerable
		1	Vulnerable
		0	Highly vulnerable
	Support for file corruption detection (2)		
		2	Available
		0	Not available
	File format stability (2)		
		2	Rare release of new versions
		1	Limited release of new versions
		0	Frequent release of new versions
	Backward compatibility (2)		
		2	Large support
		1	Medium support
		0	No support
	Forward compatibility (2)		
		2	Large support
		1	Medium support
		0	No support
Dependencies			
	Not dependent on specific hardware (8)		
		2	No dependency
		1	Low dependency
		0	High dependency
	Not dependent on specific operating systems (8)		

Criterion	Characteristic (weighing factor)	values	
		2	No dependency
		1	Low dependency
		0	High dependency
	Not dependent on one specific reader (8)		
		2	No dependency
		1	Low dependency
		0	High dependency
	Not dependent on other external resources (7)		
		2	No dependency
		1	Low dependency
		0	High dependency

## Appendix II

**Table 2: Example application of the file format assessment method to MS Word 97-2003 and PDF/A-1**

Criteria	Characteristics	Weight	PDF/A-1		MS Word 97-2003	
			Score	Total	Score	Total
Openness		3				
	Standardisation	9	2	6	0,5	1,5
	Restrictions on the interpretation of the file format	9	2	6	0	0
	Reader with freely available source	7	2	4,666666667 <sup>69</sup>	0	0
Adoption		2				
	World wide usage	4	2	4	2	4
	Usage in the cultural heritage sector as archival format	7	2	7	0	0
Complexity		3				
	Human readability	3	1	1	0	0
	Compression	6	1	2	0	0
	Variety of features	3	1	1	0	0
Technical Protection Mechanism (DRM)		5				
	Password protection	3	2	1,2	1	0,6
	Copy protection	3	2	1,2	1	0,6
	Digital signature	3	2	1,2	1	0,6
	Printing protection	3	2	1,2	2	1,2
	Content extraction protection	3	2	1,2	2	1,2
Self-documentation		2				
	Metadata	1	2	1	2	1
	Technical description of format embedded	1	0	0	0	0
Robustness		7				
	Format should be robust against single point of failure	2	0	0	0	0
	Support for file corruption detection	2	0	0	0	0
	File format stability	2	2	0,8	1	0,4
	Backward compatibility	2	2	0,8	2	0,8
	Forward compatibility	2	1	0,4	0	0
Dependencies		4				
	Not dependent on specific hardware	8	2	4	0	0
	Not dependent on specific operating systems	8	2	4	0	0
	Not dependent on one specific reader	8	2	4	0	0
	Not dependent on other external resources	8	2	4	1	2
<b>Total score</b>				56,66666667		13,9
	<b>Normalised to percentage of 100<sup>70</sup></b>			89,01 %		21,83 %

<sup>69</sup>  $4,6667 = 2 \text{ (score)} * 7 \text{ (weight for the characteristic)} / 3 \text{ (normalisation factor because there are 3 sub-characteristics for the openness criterion)}$

<sup>70</sup> The maximum score a format can receive is 63,667. By multiplying the total score by 100 and dividing it by 63,667 it is normalised to a scale from 0-100.



## Appendix 4: Storage Tests

As said in the introduction tow limitations were places upon this test:

- Only 24 bit, RGB (8 bit per colour channel) files have been tested
- Only two sets of originals have been tested: a set low contrast text material and a set of photographs

The test images on which the below data are based are 94, 300 ppi, 24 bits RGB low contrast scans of popular ballads.<sup>71</sup> The originals vary in format between slightly larger than A4 to smaller than A5.

File Format and Compression	File Size of Test Batch	Average File Size <sup>72</sup>	Storage Savings Compare to Uncompressed TIFF <sup>73</sup>	Storage Interpolated for 500.000 Files <sup>74</sup>
Uncompressed TIFF	623 MB	6.6 MB		3.1 TB
TIFF LZW lossless	428 MB	4.6 MB	31%	2.2TB
JPEG 10 <sup>75</sup>	66 MB	0.7 MB	89%	343 GB
JPEG 8	35 MB	0.4 MB	94%	195 GB
JPEG 6	26 MB	0.3 MB	96%	146 GB
JPEG 1	10 MB	0,1 MB	98%	49 GB
PNG lossless	355 MB	4 MB	43%	2 TB
JPEG2000 lossless <sup>76</sup>	298 MB	3,2 MB	52%	1.5 TB
JPEG2000 compression ratio10	54 MB	0.6 MB	91%	280 GB
JPEG2000 compression ratio 25	25 MB	0.3 MB	96%	146 GB
JPEG2000 compression ratio 50	13 MB	0,1 MB	98%	68 GB

In addition to this set of popular ballads, a test was conducted on 104 scans from photo prints, scanned in RGB. The results were almost identical.

<sup>71</sup> Scanned within the scope of the Geheugen van Nederland (Memory of the Netherlands) project <http://www.geheugenvannederland.nl/straatliederen>.

<sup>72</sup> The number of files – 94 – divided by the file size of all files together.

<sup>73</sup> Percentage of total storage of 94 uncompressed TIFFs (RGB 653 GB and grey 218 GB ) compared to total storage of 94 compressed files.

<sup>74</sup> Average file size multiplied by 500.000.

<sup>75</sup> JPEG Adobe Photoshop scale quality 10.

<sup>76</sup> Lead JPEG2000 plugin for Photoshop is used, whereby the amount of compression is set by means of the compression ratio. Compression ratio 10 is minimum compression and is qualitatively comparable to JPEG10. Compression ratio 25 is average compression and is qualitatively comparable to JPEG6. Compression ratio 50 is strong compression and is qualitatively comparable to JPEG1.

Additional testing was performed with the Photoshop native plugin. Lossless compression proved to be slightly less successful than that of the LEAD plugin: 53% for the Lead plugin, versus 52% for the Photoshop plugin.

Additional testing with other converters, like the Lurawave tool (<http://www.luratech.com/products/lurawave/jp2/clt/>), is necessary.

## Bibliography

### General

- LOC Sustainability of Digital Formats - Planning for Library of Congress Collections - Still Images Quality and Functionality Factors  
[http://www.digitalpreservation.gov/formats/content/still\\_quality.shtml](http://www.digitalpreservation.gov/formats/content/still_quality.shtml).
- *Sustainability of Digital Formats Planning for Library of Congress Collections, Format Description for Still Images*  
[http://www.digitalpreservation.gov/formats/fdd/still\\_fdd.shtml](http://www.digitalpreservation.gov/formats/fdd/still_fdd.shtml).
- Florida Digital Archive Format Information  
<http://www.fcla.edu/digitalArchive/formatInfo.htm>.
- Roberto Bourgonjen, Marc Holtman, Ellen Fleurbaay, *Digitalisering ontrafeld. Technische aspecten van digitale reproductie van archiefstukken* (Digitisation Unraveled. Technical Aspects of Digital Reproduction of Archive Pieces).  
[http://stadsarchief.amsterdam.nl/stadsarchief/over\\_ons/projecten\\_en\\_jaarverslagen/digitalisering\\_ontrafeld\\_web.pdf](http://stadsarchief.amsterdam.nl/stadsarchief/over_ons/projecten_en_jaarverslagen/digitalisering_ontrafeld_web.pdf).

### JPEG2000

- Judith Rog, *Notitie over JPEG 2000 voor de KB* (Note regarding JPEG 2000 for the RL), version 0.2 (August 2007).
- JPEG2000 homepage: <http://www.jpeg.org/jpeg2000/>.
- Wikipedia JPEG 2000 lemma: [http://en.wikipedia.org/wiki/JPEG\\_2000](http://en.wikipedia.org/wiki/JPEG_2000).
- Robert Buckley, *JPEG2000 for Image Archiving, with Discussion of Other Popular Image Formats*. Tutorial IS&T Archiving 2007 Conference.
- Robert Buckley, *JPEG 2000 – a Practical Digital Preservation Standard?*, a DPC Technology Watch Series Report 08-01, February 2008:  
<http://www.dpconline.org/graphics/reports/index.html#jpeg2000>
- Florida Digital Archive description of JPEG 2000 part 1:  
[http://www.fcla.edu/digitalArchive/pdfs/action\\_plan\\_bgrounds/jp2\\_bg.pdf](http://www.fcla.edu/digitalArchive/pdfs/action_plan_bgrounds/jp2_bg.pdf).
- *Sustainability of Digital Formats Planning for Library of Congress Collections, JPEG 2000 Part 1, Core Coding System*  
<http://www.digitalpreservation.gov/formats/fdd/fdd000138.shtml>.

### PNG

- Portable Network Graphics (PNG) Specification (Second Edition)  
Information technology — Computer graphics and image processing — Portable Network Graphics (PNG): Functional specification. ISO/IEC 15948:2003 (E), W3C Recommendation 10 November 2003 <http://www.w3.org/TR/PNG/>.
- Wikipedia PNG lemma: [http://nl.wikipedia.org/wiki/Portable\\_Network\\_Graphics](http://nl.wikipedia.org/wiki/Portable_Network_Graphics).
- *Sustainability of Digital Formats Planning for Library of Congress Collections, PNG, Portable Network Graphics*  
<http://www.digitalpreservation.gov/formats/fdd/fdd000153.shtml>.
- Greg Roelofs, *A Basic Introduction to PNG Features*  
<http://www.libpng.org/pub/png/pngintro.html>.

### JPEG

- JPEG standard: <http://www.w3.org/Graphics/JPEG/itu-t81.pdf>.



- JPEG homepage: <http://www.jpeg.org/jpeg/index.html>.
- JFIF standard: <http://www.jpeg.org/public/jfif.pdf>.
- Florida Digital Archive description of JFIF 1.02:  
[http://www.fcla.edu/digitalArchive/pdfs/action\\_plan\\_bgrounds/jfif.pdf](http://www.fcla.edu/digitalArchive/pdfs/action_plan_bgrounds/jfif.pdf).
- *Sustainability of Digital Formats Planning for Library of Congress Collections, JFIF JPEG File Interchange Format*  
<http://www.digitalpreservation.gov/formats/fdd/fdd000018.shtml>.
- Wikipedia JPEG lemma: <http://en.wikipedia.org/wiki/JPEG>.
- JPEG-LS homepage: <http://www.jpeg.org/jpeg/jpegls.html>.
- Wikipedia JPEG-LS lemma: <http://www.jpeg.org/jpeg/jpegls.html>.

## **TIFF LZW**

- TIFF 6.0 standard: <http://partners.adobe.com/public/developer/en/tiff/TIFF6.pdf>.
- TIFF/EP extension ISO 12234-2 [http://en.wikipedia.org/wiki/ISO\\_12234-2](http://en.wikipedia.org/wiki/ISO_12234-2).
- TIFF/IT (ISO 12369) extension for prepress purposes  
<http://www.digitalpreservation.gov/formats/fdd/fdd000072.shtml>.
- DNG Adobe TIFF UNC extension for storing RAW images  
<http://www.digitalpreservation.gov/formats/fdd/fdd000188.shtml>.
- EXIF technical metadata of cameras and camera settings  
(<http://www.digitalpreservation.gov/formats/fdd/fdd000145.shtml>).
- LOC overview of TIFF extension  
[http://www.digitalpreservation.gov/formats/content/tiff\\_tags.shtml](http://www.digitalpreservation.gov/formats/content/tiff_tags.shtml).
- Unisys patent LZW [http://www.unisys.com/about\\_unisys/lzw/](http://www.unisys.com/about_unisys/lzw/).
- *Sustainability of Digital Formats Planning for Library of Congress Collections, TIFF, Revision 6.0* <http://www.digitalpreservation.gov/formats/fdd/fdd000022.shtml>.
- *Sustainability of Digital Formats Planning for Library of Congress Collections, TIFF, Bitmap with LZW Compression*  
<http://www.digitalpreservation.gov/formats/fdd/fdd000074.shtml>.

## Post scriptum to report ‘Alternative File Formats for Storing Master Images of Digitisation Projects’

The report ‘Alternative File Formats for Storing Master Images of Digitisation Projects’ was written in early 2008, and as such it reflects the KB’s knowledge of JPEG 2000 Part 1 (effectively the JP2 format) at that time.

We would like to emphasise that the scope of this study is limited to a largely *theoretical* take on the suitability of JP2 (and the other discussed formats) for long-term preservation. It does not address any of the more *practical* issues that one may encounter when actually using JPEG 2000.

In addition, the report’s coverage of JP2’s colour space support is incomplete, which has implications for some of the conclusions on quality, long-term sustainability and functionality.

We are currently preparing a publication in which we will explain these issues in greater detail, as well as suggesting some solutions.

These observations do not change the KB’s position on using JP2 as a master image format. However, we would merely like to point out that the report only presents a partial view, and we do not advocate its use as the sole basis for deciding upon JP2 (or *any* format, for that matter).

For more information please contact Johan van der Knijff ([johan.vanderknijff@kb.nl](mailto:johan.vanderknijff@kb.nl)) or Barbara Sierman ([barbara.sierman@kb.nl](mailto:barbara.sierman@kb.nl)).

# Library of Congress

## News Releases

[The Library of Congress](#) > [News Releases](#) > Library of Congress and DuraCloud Launch Pilot Program

[Print](#) [Subscribe](#) [Share/Save](#)

### News from the Library of Congress

**Contact:** Erin Allen, Library of Congress (202) 707-7302

**Contact:** Carol Minton Morris, DuraSpace, [carol@duraspace.org](mailto:carol@duraspace.org)

**July 14, 2009**

#### **Library of Congress and DuraCloud Launch Pilot Program Using Cloud Technologies to Test Perpetual Access to Digital Content**

##### **Service is Part of National Digital Information Infrastructure and Preservation Program**

How long is long enough for our collective national digital heritage to be available and accessible? The Library of Congress National Digital Information Infrastructure and Preservation Program (NDIIPP) and DuraSpace have announced that they will launch a one-year pilot program to test the use of cloud technologies to enable perpetual access to digital content. The pilot will focus on a new cloud-based service, DuraCloud, developed and hosted by the DuraSpace organization. Among the NDIIPP partners participating in the DuraCloud pilot program are the New York Public Library and the Biodiversity Heritage Library.

Cloud technologies use remote computers to provide local services through the Internet. Duracloud will let an institution provide data storage and access without having to maintain its own dedicated technical infrastructure.

For NDIIPP partners, it is not enough to preserve digital materials without also having strategies in place to make that content accessible. NDIIPP is concerned with many types of digital content, including geospatial, audiovisual, images and text. The NDIIPP partners will focus on deploying access-oriented services that make it easier to share important cultural, historical and scientific materials with the world. To ensure perpetual access, valuable digital materials must be stored in a durable manner. DuraCloud will provide both storage and access services, including content replication and monitoring services that span multiple cloud-storage providers.

Martha Anderson, director of NDIIPP Program Management said "Broad online public access to significant scientific and cultural collections depends on providing the communities who are responsible for curating these materials with affordable access to preservation services. The NDIIPP DuraCloud pilot project with the DuraSpace organization is an opportunity to demonstrate affordable preservation and access solutions for communities of users who need this kind of help."

The New York Public Library offers a set of scholarly research collections with an intellectual and cultural range that is both global and local. The DuraCloud pilot program at the library will replicate large collections of digital images from a Fedora repository into DuraCloud. The New York Public Library plans to convert the images from the TIFF format to JPEG 2000 and to serve these images using a powerful JPEG 2000 image engine within DuraCloud.

The Biodiversity Heritage Library provides access to historical journal literature in biodiversity in collaboration with global partners, including the Smithsonian Institution, the Missouri Botanical Gardens and the Woods Hole Marine Biology Lab. Their DuraCloud pilot will focus on replication of digital content to provide protection for valuable biodiversity resources. The pilot will demonstrate bi-directional replication of content among partners in the United States and Europe. The library will use the cloud-computing capabilities offered by DuraCloud to analyze biodiversity texts to extract key information such as species-related words. The institution will also deploy a JPEG2000 image engine via DuraCloud to process and serve digital images.

DuraCloud is a cloud-based service developed and hosted by the nonprofit organization DuraSpace. DuraCloud was developed with support from the Gordon and Betty Moore Foundation and the Andrew W. Mellon Foundation. The DuraCloud pilot program is being launched with support from NDIIPP. DuraCloud is aimed at helping institutions and individuals take advantage of cloud technologies in providing access to their digital materials. DuraCloud is focused on providing trusted solutions for organizations such as universities, libraries, cultural heritage organizations, research centers, and others who are concerned with ensuring perpetual access to their digital content.

DuraSpace provides leadership and innovation in open-source technologies for global communities who manage, preserve and provide access to digital content. DuraSpace was established by merging Fedora Commons and the DSpace Foundation, two of the largest providers of open-source repository software worldwide. DuraSpace serves more than 750 institutions that are committed to the use of open-source software solutions for the dissemination and preservation of academic, scientific, and cultural digital assets.

The mission of the National Digital Information Infrastructure and Preservation Program is to develop a national strategy to collect, preserve and make available digital content, especially materials that are created only in digital formats, for current and future generations.

# # #

PR 09-140

07/14/09

ISSN 0731-3527

[Back to top](#)

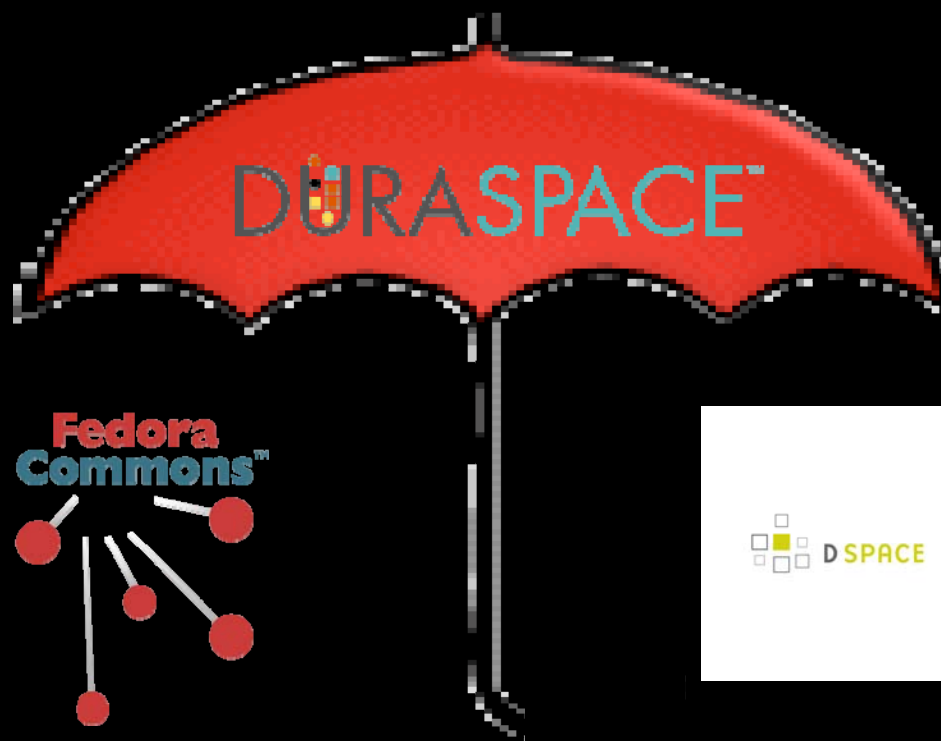
# DuraCloud

*Open technologies and services for  
managing durable data in the cloud*

The logo for DuraSpace, featuring the word "DURA" in a grey, sans-serif font above the word "SPACE" in a teal, sans-serif font. A small graphic of colored squares (orange, yellow, blue, black) is positioned to the left of the "A" in "DURA".

Michele Kimpton, CBO DuraSpace  
“Repositories in the Cloud” Seminar, Feb 2010  
[mkimpton@duraspace.org](mailto:mkimpton@duraspace.org)

## Open Source Portfolio



## Implications for our future work



more distributed

more collaborative

more web-oriented

more open

more interoperable

## Challenges (from survey 1/22/2010)

*Preservation support is hard to implement consistently*

*“Our preservation support is collection based where we have had grants or specific initiatives. There is no system effort.”*

*“Where it is prioritized as mission critical, it is being done well. It is not being done well where it is not mission critical.”*

*“We have not invested enough to make it a service of which we are proud...”*

*“Collection development and storage are more important than computing.”*



# Key Advantages

completed 1/22/2010

145 participants higher ed

<b>Most Impactful Advantages Electronic Survey</b>	<b>Responses</b>
Scalability	79
Remote, Off Campus Storage of Digital Assets	64
Ease of Implementation	54
Flexibility	53
Don't Have to Staff Locally	39
Cost	33
Elasticity	26
Pay for Use	14
Other	5

# Key Challenges

completed 1/22/2010

145 participants higher ed

<b>Key Challenges Electronic Survey</b>	<b>Responses</b>
Trusting Third Party to Manage Critical Assets	64
Long-term Reliability of Solution	52
Data Security	51
Performance and Bandwidth Concerns	37
Loss of Control	34
Administrative Burden of SLAs	17
Transparency of Solution	16
Concerns about Data Lock-in	16
Less Customizable	10
Other	12

# Likely to use cloud services in next 12 months

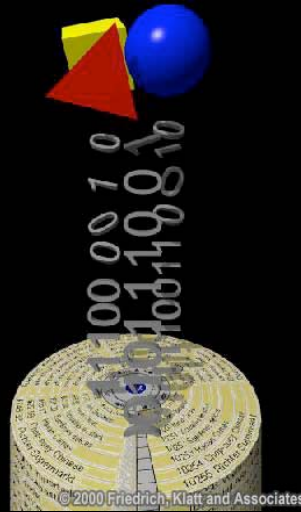
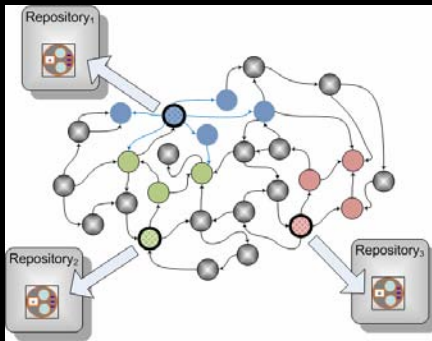
Percentage of electronic survey respondents noting it is “very likely” or “likely” they will use cloud compute or cloud storage services to manage, store or provide access to digital collections in the next twelve months.

Category		Subcategory	Percentage
Non-US			47.7%
US Institutions	Institution Size	Large, very large	47.2%
		Medium	68.8%
		Small, very small	42.9%
	Enrollment Profile	RU/VH	52.1%
		RU/H, DRU	50.0%
		Master's S, M and L	46.2%
		Bac and Assoc	57.1%
	Public/Private	Public	46.9%
		Private	59.3%

# Institutional needs: managing digital collections

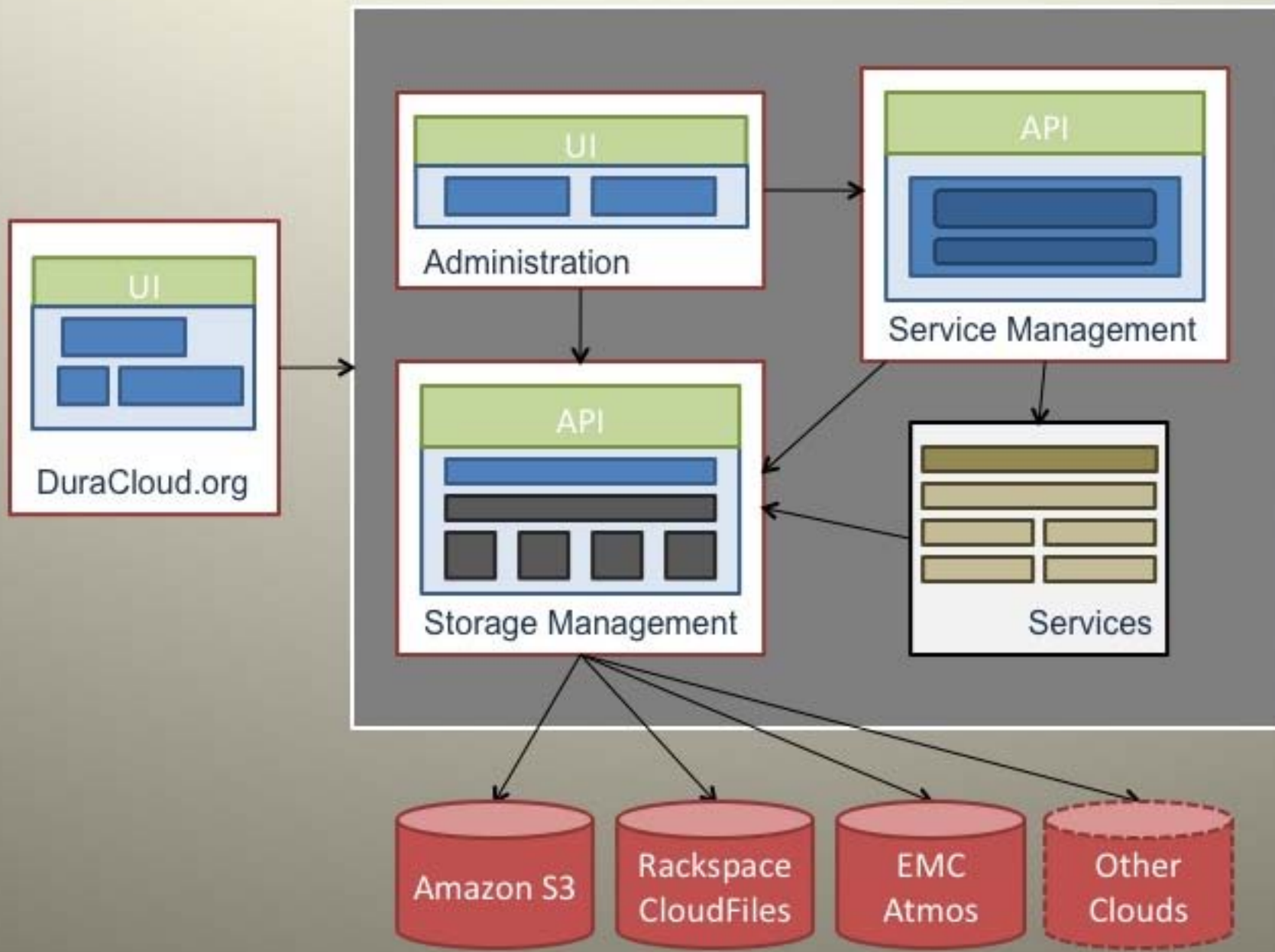
<b>Service Area</b>	<b>Importance</b>	<b>Extent Need is Met</b>	<b>Difference</b>	<b><i>Likelihood to Use Cloud Services</i></b>
Remote secondary storage of digital collections	3.54	2.60	0.94	3.09
Preservation support	3.35	2.17	1.18	2.88
Intra-institution shared collections	3.11	2.47	0.64	2.69
Inter-institution shared collections	2.72	2.07	0.65	2.67
Compute services	2.80	2.25	0.55	2.54
<i>Online primary storage</i>	<i>3.51</i>	<i>2.97</i>	<i>0.53</i>	<i>2.29</i>

# Services in the cloud for durable digital content



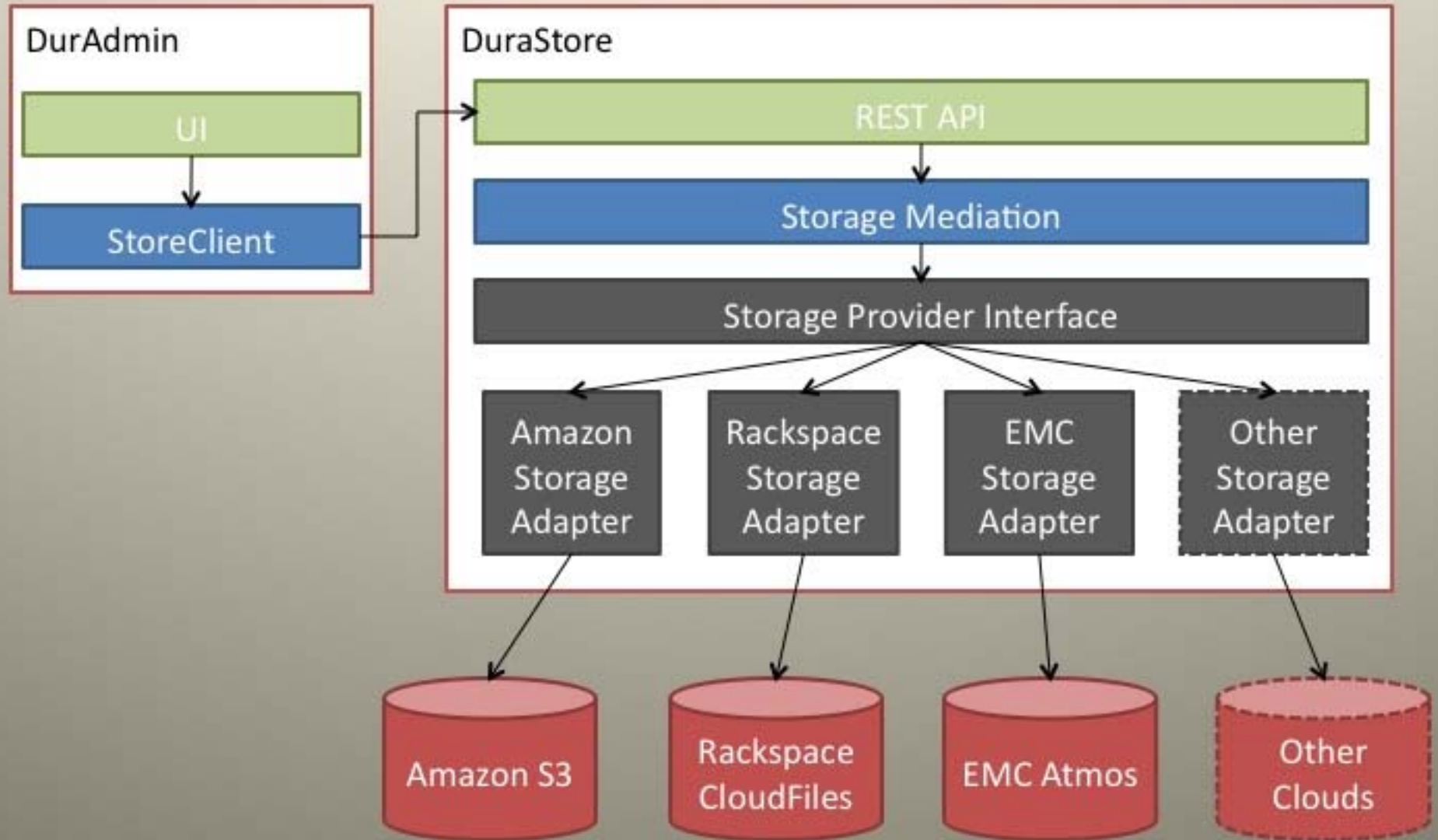
**DuraCloud Platform: Allow organizations to utilize cloud infrastructure easily offering data storage, data replication, preservation support and access services**

# DuraCloud High Level Interaction

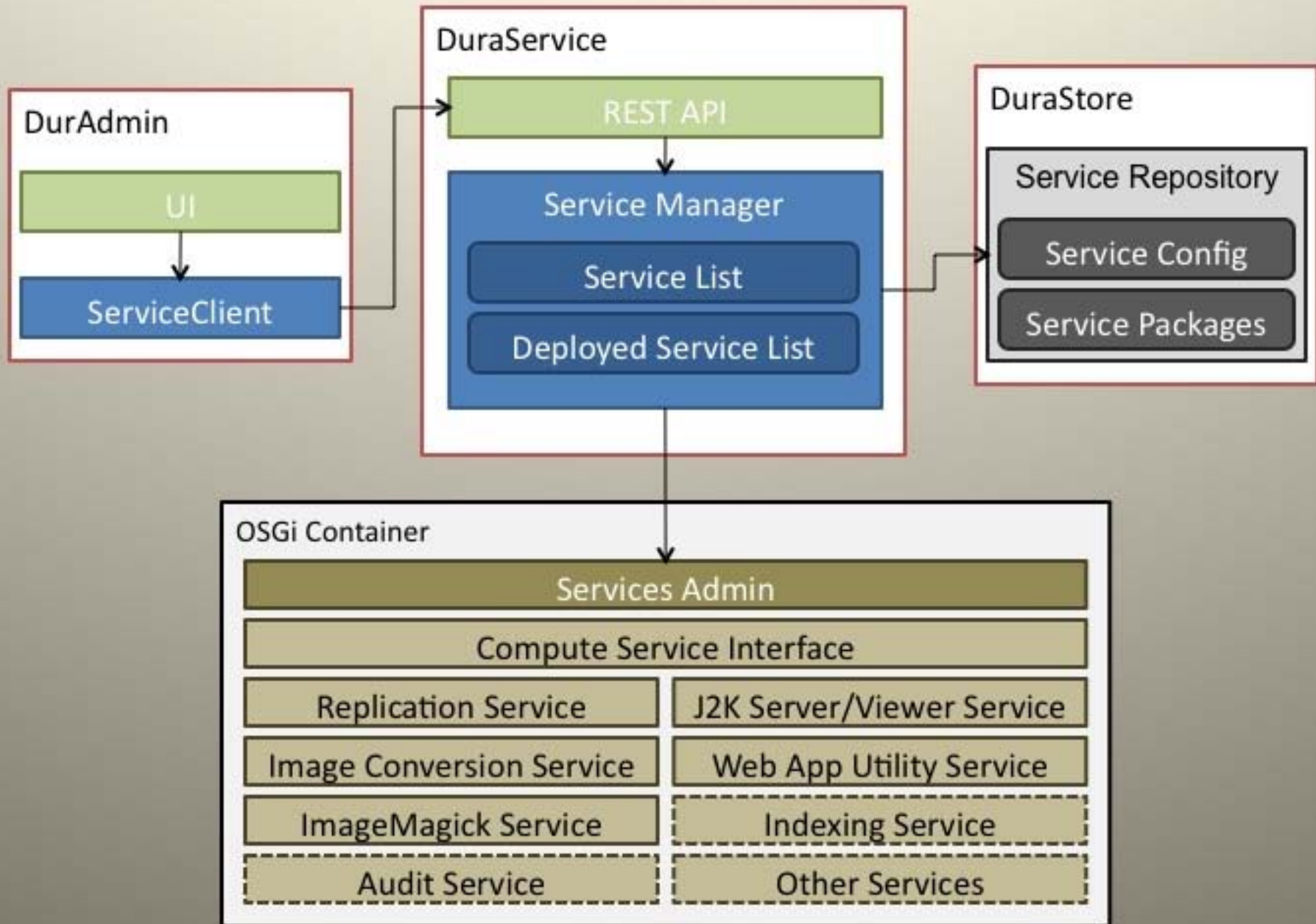




# DuraCloud Storage Architecture



# DuraCloud Service Architecture





# Preservation Services

- ability to replicate content to multiple providers and locations*
- ability to synchronize backup with primary store or repository system*
- access to content through web based interface*
- ability to do bit integrity checking*
- ability to do file format transformations*

# Partners and Pilots

- Selected initial cloud providers



- Selected 3 initial pilot partners



# NYPL pilot

## *Digital Gallery Collection*

Use case: back up online preservation copy to Fedora, file format transformation

-back up copy all TIFF images (10 TB data)

-transformation from Tiff to JPEG 2000 using Imagemagick


-run J2k image server in cloud

-Push JPEG 2000 back into Fedora Repository

Explore NYPL's new online exhibition: [Treasures of the New York Public Library](#)

The New York Public Library

**Digital Gallery**



Lispenard Street between Church Street and Broadway ... Source: Drugstore Photographs, Or, A Trip Along the Yangtze River, 1999, by Dylan Stone. Digital ID: 518688.

**NYPL Digital Gallery** provides free and open access to over 700,000 images digitized from the The New York Public Library's vast collections, including illuminated [manuscripts](#), historical [maps](#), vintage [posters](#), rare [prints](#), [photographs](#) and [more](#).


**SEARCH**  [More Options](#)

**BROWSE**

- Subjects
- Names
- Library Divisions
- My Selections

Arts & Literature  
Cities & Buildings  
Culture & Society  
History & Geography  
Industry & Technology  
Nature & Science  
Printing & Graphics

**DIGITAL GALLERY PICKS**



# BHL pilot

## BioDiversity Heritage Library

Use case: Find the best cost competitive solution for keeping multiple copies in multiple geographies, easily accessible.

-back up copy entire corpus (40 TB data-  
JPEG, Tiff

-have multiple copies including Europe  
-Run J2K image server in cloud

The screenshot shows the Biodiversity Heritage Library website. At the top, it says "News and Updates from the Biodiversity Heritage Library" and "available online at http://www.biodiversitylibrary.org". Below this, it states "Over 250 years of biodiversity literature to support the work of scientists, researchers, students and you." There are several small illustrations of various organisms: a jellyfish, a squid, a dragonfly, a butterfly, a lizard, a plant, a whale, and a crab. The main content area is titled "TUESDAY, AUGUST 25, 2009" and "Book of the Week: BHL, EOL, and Marine Life". It features a large image of two orcas. To the left, there are sections for "BHL RECENT ADDITIONS" listing various publications and their dates. To the right, there is an "ABOUT BHL" section describing the library's mission and a "LABELS" section with a "News (10)" link.

## WGBH Media Library and Archives

Use case: Provide backup preservation for video files from repository and other sources, and create derivative files for access and streaming.

- Archive large video files
- Provide public access to streaming versions
- Transcode files in cloud
- Edit files where appropriate to sell clips
- Give third party access to cloud store for processing and access



# Challenges

- Provisioning bandwidth at local institution to transfer data
- Transferring large files over the wire ( over 5 GB is rejected, found issues in transfer over 1 GB)
- Consistency of operation of 2<sup>nd</sup> tier providers (EMC, RackSpace)
- Enabling others to easily build on platform
- Best process for integration of 3<sup>rd</sup> party applications into hosting service
- Cost effective bit integrity checking
- Balancing ease of use and more sophisticated functionality

# Advantages of hosted platform

## Strategic partnerships with cloud providers

Better pricing

Transparency

Early notification

## Ease of implementation for end user

Multiple copies in multiple geographies/administrations through one interface

✓ Access to broad number of services relevant to the repository community

# Timeline

- Begin pilots— September 2009
- DuraCloud Alpha Pilot release- Oct 2009
- Pilot data loading and testing – Fall 2009
- **Beta for repository community – Q2 2010**
- **Pilot testing with software services Q2 2010**
- Cloud partner evaluations complete-Q3 2010
- Hosting service pricing and SLA's complete-Q3 2010
- Report pilot results – Q3 2010
- Code available open source-Q3 2010
- Launch production service Q4 2010



## Next Steps(Feb-April)

- V.2 release complete
  - Replication, web access and viewing, file format conversion, J2K image server, bit integrity checking
- Launch Fedora and DSpace plug ins
- V.3 release primary features
  - Synchronization with local repository( Fedora and DSpace)
- Expand pilot in April to include 15 new users, to connect with current repositories
- Continue to test robustness and performance of commercial cloud partners

# Thank You



For more information:

DuraSpace Organization: <http://duraspace.org>

Wiki: <http://www.fedora-commons.org/confluence/display/duracloudpilot>

/

DuraCloud project page: <http://duracloud.org>

[Mkimpton@duraspace.org](mailto:Mkimpton@duraspace.org)

[Back to Results](#)**Web of Knowledge**  
**Page 1 (Articles 1 -- 1)**[Print This Page](#)

◀ [ 1 ] ▶

**Record 1 of 1****Accession Number:** 12301330**Document Type:** Conference Paper**Title:** Perceptual quantitative quality assessment of JPEG2000 compressed ct images with various slice thicknesses**Author(s):** Pambrun, J.-F.; Noumeir, R.**Source:** Proceedings of the 2011 IEEE International Conference on Multimedia and Expo (ICME 2011) Pages: 6 pp. Published: 2011**Meeting Information:** 2011 IEEE International Conference on Multimedia and Expo (ICME 2011)

Barcelona, Spain, 11-15 July 2011

**Language:** English**Treatment:** Practical

**Abstract:** Modern medical equipments produce huge amounts of data that need to be archived for long periods and efficiently transferred over networks. Data compression plays an essential role in reducing the amount of medical imaging data. Medical images can usually be compressed by a factor of three before any degradation appears. Higher compression levels are desirable but can only be achieved with lossy compression, thus scarifying image quality. The diagnosis value of compressed medical images has been studied and recommendations about maximum acceptable compression ratios have been provided based on qualitative visual analysis. It has been suggested, without further investigation, that CT images, with thicknesses below five mm, cannot undergo lossy compression if diagnostic value needed to be preserved. In this pa per, we present an objective quantitative quality assessment of compressed CT images using Visual Signal to Noise Ratio. Our results show that visual fidelity can be significantly affected by two factors, slice thickness and exposure time, for images compressed using the same compression ratio.

**Controlled Indexing:** biomedical equipment; computerised tomography; data compression; image coding; medical image processing**Uncontrolled Indexing:** perceptual quantitative quality assessment; JPEG2000 compressed CT images; slice thicknesses; medical equipments; data compression; medical imaging data; lossy compression; image quality; qualitative visual analysis; visual signal to noise ratio; computerized tomography**Classification Codes:** A8760J X-rays and particle beams (medical uses); A8770E Patient diagnostic methods and instrumentation; B7510P X-ray techniques: radiography and computed tomography (biomedical imaging/measurement); B6135C Image and video coding; C7330 Biology and medical computing; C5260B Computer vision and image processing techniques**International Patent Classification:** A61B6/03; G06F19/00; G06T; G06T9/00**Author Address:** Pambrun, J.-F.; Noumeir, R.; Dept. of Electr. Eng., Univ. of Quebec, Montreal, QC, Canada.**Publisher:** IEEE, Piscataway, NJ, USA**Number of References:** 32**U.S. Copyright Clearance Center Code:** 978-1-61284-350-6/11/\$26.00**Standard Book Number:** 978-1-61284-349-0[Back to Results](#)**Web of Knowledge**  
**Page 1 (Articles 1 -- 1)**[Print This Page](#)

◀ [ 1 ] ▶

[Back to Results](#)Web of Knowledge  
Page 1 (Articles 1 -- 1)[Print This Page](#)

◀ [ 1 ] ▶

**Record 1 of 1****Title:** CCD noise influence on JPEG2000 compression of astronomical images**Author(s):** Pata, P (Pata, Petr)**Editor(s):** Tomanek P; Senderakova D; Pata P**Source:** PHOTONICS, DEVICES, AND SYSTEMS V **Book Series:** Proceedings of SPIE **Volume:** 8306 **Article Number:** 83061P **DOI:** 10.1117/12.912322 **Published:** 2011**Times Cited in Web of Science:** 0**Total Times Cited:** 0**Cited Reference Count:** 8

**Abstract:** Compression of astronomical images is still current task. In most applications, lossless approaches are used that do no damage to the compressed data. These algorithms, however, have lower compression ratios and are not as effective. It is therefore important to deal with more efficient lossy compression techniques. For them it is necessary to define quality criteria and level of acceptable distortion of image data. The usual multimedia approach is not possible to use for the scientific image data. They are optimized for human vision. This work deals with the influence of noise generated in the CCD structure on the defined quality criteria. It will also be shown the impact of the lossy standard JPEG2000 on quality of image data in astronomy.

**Accession Number:** WOS:000297582500061**Language:** English**Document Type:** Proceedings Paper**Conference Title:** Conference on Photonics, Devices, and Systems V**Conference Date:** AUG 24-26, 2011**Conference Location:** Prague, CZECH REPUBLIC**Conference Sponsor(s):** European Opt Soc (EOS), Czech & Slovak Soc Photon (CSSF), Act M Agcy**Author Keywords:** CCD noise; ACC; astrometry; astronomical image compression; JPEG2000**Addresses:** Czech Tech Univ, Fac Elect Engn, Dept Radioelect, CR-16635 Prague, Czech Republic**Reprint Address:** Pata, P (reprint author), Czech Tech Univ, Fac Elect Engn, Dept Radioelect, CR-16635 Prague, Czech Republic**E-mail Address:** pata@fel.cvut.cz, pata@fel.cvut.cz**Publisher:** SPIE-INT SOC OPTICAL ENGINEERING**Publisher Address:** 1000 20TH ST, PO BOX 10, BELLINGHAM, WA 98227-0010 USA**Web of Science Category:** Optics**Subject Category:** Optics**IDS Number:** BXX98**ISSN:** 0277-786X**ISBN:** 978-0-81948-953-1**29-char Source Abbrev.:** PROC SPIE**Source Item Page Count:** 6[Back to Results](#)Web of Knowledge  
Page 1 (Articles 1 -- 1)[Print This Page](#)

◀ [ 1 ] ▶

5-7-2009

# A Status Report on JPEG 2000 Implementation for Still Images: The UConn Survey

David Lowe

*University of Connecticut - Storrs*, david.lowe@uconn.edu

Michael J. Bennett

*University of Connecticut - Storrs*, michael.bennett@uconn.edu

---

## Recommended Citation

Lowe, David and Bennett, Michael J., "A Status Report on JPEG 2000 Implementation for Still Images: The UConn Survey" (2009).  
*UConn Libraries Published Works*. Paper 19.  
[http://digitalcommons.uconn.edu/libr\\_pubs/19](http://digitalcommons.uconn.edu/libr_pubs/19)

This Conference Proceeding is brought to you for free and open access by the University of Connecticut Libraries at DigitalCommons@UConn. It has been accepted for inclusion in UConn Libraries Published Works by an authorized administrator of DigitalCommons@UConn. For more information, please contact [digitalcommons@uconn.edu](mailto:digitalcommons@uconn.edu).

# A Status Report on JPEG 2000 Implementation for Still Images: The UConn Survey

David B. Lowe and Michael J. Bennett; University of Connecticut Libraries; Storrs, CT/USA

## Abstract

*JPEG 2000 is the product of thorough efforts toward an open standard by experts in the imaging field. With its key components for still images published officially by the ISO/IEC by 2002, it has been solidly stable for several years now, yet its adoption has been considered tenuous enough to cause imaging software developers to question the need for continued support. Digital archiving and preservation professionals must rely on solid standards, so in the fall of 2008 the authors undertook a survey among implementers (and potential implementers) to capture a snapshot of JPEG 2000's status, with an eye toward gauging its perception within this community.*

*The survey results revealed several key areas that JPEG 2000's user community will need to have addressed in order to further enhance adoption of the standard, including perspectives from cultural institutions that have adopted it already, as well as insights from institutions that do not have it in their workflows to date. Current users were concerned about limited compatible software capabilities with an eye toward needed enhancements. They realized also that there is much room for improvement in the area of educating and informing the cultural heritage community about the advantages of JPEG 2000. A small set of users, in addition, perceived problems of cross-codec consistency and future file migration issues.*

*Responses from non-users disclosed that there were lingering questions surrounding the format and its stability and permanence. This was stoked largely by a dearth of currently available software functionality, from the point of initial capture and manipulation on through to delivery to online users.*

## Background

In the fall of 2008, the authors surveyed the status of JPEG 2000 implementation as a still image format among cultural heritage institutions involved in digitization. This sample was taken from August 27, 2008 through October 31, 2008. Respondents totaled 161, with the overwhelming majority coming from academic research libraries [1]. The following focuses primarily on the major issues broached by respondents, examines current use and perceived barriers to the standard's adoption, and proposes recommendations towards JPEG 2000's greater utilization within the cultural heritage community.

## Migration Concerns

### Codec Inconsistency

An interesting opinion among respondents focused on perceived codec inconsistencies among software vendors. Coupled with this were migration concerns based upon such

inconsistencies and also the general nature of JPEG 2000's currently limited adoption, and future migration toolkits:

"Lack of consistency across codecs (e.g. Aware, Kakadu) for creating JPEG 2000s." [written in response to the question of drawbacks to JPEG 2000 implementation]

"JPEG 2000 is a great format, but the main problem resides in acceptance not only in the repository level but also commercially. To have a fully robust digital archival format we will require good migration software for when it becomes obsolete. If it becomes commonly used (such as TIFF) migration software will work smoother with less errors as they will not have to necessarily be homegrown."

"It's a new format with an unproven history or migration."

Codec concerns may be ameliorated to some extent when put into the larger context of the standard itself. JPEG 2000 is a fully documented and open standard and as such is available for software developers of all types (vendors and freeware authors) to write encoders for. Much like capture hardware's vagaries of unique sensor filters and device-specific profiles, software encoders are similarly geared around their creator's best perceptions of fidelity in the production of these files.

Perhaps the most important residual of the standard's openness in this regard, however, is the fact that decoding of valid JPEG 2000 files remains transparent regardless of the encoder used. In this way, migration concerns may be mitigated to a degree as developers today and into the future can be assured access to the standard in order to write such applications. Yet, the questions of future prevalence and quality of software toolkits for JPEG 2000 mass migration remain foremost in many practitioners' minds.

### Visually Lossless, Mathematically Lossless

The possibility of visually lossless (mathematically lossy) JPEG 2000 compression as an archival storage option has recently begun to gain traction particularly in areas of large scale TIFF migration at both Harvard and the Library of Congress [2][3]. Mass digitization projects such as the Open Content Alliance have also adopted visually lossless JPEG 2000 as their archival standard [4]. Among survey respondents there was a divergence of opinion on the idea as many felt that possible future migration costs for moving out of the standard may not make up for the real benefits of storage efficiencies realized today:

"I have some concerns that once we start going down a slope of compromising images what the potential of it being accentuated after multiple migrations possibly with different

lossy compression schemes. Considering the relative cost of space I don't think it is a worthwhile risk."

"...visually lossless but technically lossy compression is not a good basis for later format migrations."

Here, notions of fidelity between hardware and software, born digital vs. converted digital were used to try to strike a balance in the decision making process:

"...visually lossless is fine. The only reason to use mathematically lossless would be in conversion of born digital materials where there hasn't already been loss due to analog to digital conversion. For analog materials, the loss inherent in lossy jp2 is minimal compared to sensor and sampling error from the original scanning."

Lossless JPEG 2000 compression, though in fact lossless at the bit-level upon decoding, still elicited its own migration anxieties:

"I have little concern over lossless compression other than prominence and easy migration. It adds another level of encoding which could very well complicate future migrations (especially if one is missed) unless it is common and well documented. Again the availability of good migration software is useful."

Finally, as one respondent phrased it, nothing may ever be perfect:

"One problem with the widespread acceptance of .jp2 is the fear of future migration. However, I have heard that migration projects of tiff formats haven't gone smoothly either."

### **Current Use Scenarios Point to Advantages**

JPEG 2000's support of both true lossless and a wide array of visually lossless (lossy) compression enjoyed broad use at many of the responding institutions. Scalable storage savings through the standard's comparative file size economy to TIFF and JPEG 2000's flexible individual file rendering on the web were focal reason for its favor. A sample of the more intricate use scenarios included the following responses:

"We produce TIFF files for our new photography, and for some projects we produce lossless JP2 files that we class as "master". In these cases we discard the original TIFF and the losslessly-encode file serves as master and delivery image. For some projects we save uncompressed-TIFF files, classes as "masters", and also produce a lossy compressed JP2 file for delivery purposes. A third common workflow produces TIFF files that are used to produce conservatively, but lossy-compressed JP2 files for delivery. The TIFF files are not saved and the lossy JP2 files serve as masters and as delivery images."

"Yes, we have started to make lossless JPEG 2000 images for some collections where we would have previously saved (LONG term) uncompressed TIFF and lossy JP2 for delivery.

We like keeping a single file that can be used as master and deliverable, and the fidelity is equivalent to an uncompressed TIFF."

"For our large-scale book scanning projects (published materials from circulating collections) we are saving conservatively, but lossy compressed color JP2 images. This is a high quality, but lower cost and high volume service and we need to take advantage of the power of lossy compression to reduce our file storage costs."

### **Misperceived Disadvantages Affect Adoption**

Incorrect assumptions on the standard were, however, common throughout the survey and revealed a real need for better education and understanding. Common threads included a lack of trust in JPEG 2000's lossless compression as being truly lossless. Others believed that such lossless compression did not confer significant file size savings in comparison to uncompressed TIFF or that JPEG 2000 did not support higher bit depth images. A small number of respondents continued to make the unfortunate association of JPEG 2000 with JPEG as two lossy-only standards, a belief that has hounded JPEG 2000 in particular since its inception. Also false notions of JPEG 2000 as being proprietary in nature and not fully documented lead some to believe that software tools would forever be scarce, expensive, and never open source.

### **Compression Choices in the Context of other Standards and Best Practices**

As part of the "visually lossless" (that is, slightly lossy) vs. mathematically lossless compression debate, it is important to emphasize that, although any whiff of lossiness may raise eyebrows among some colleagues, there may be perfectly reasonable situations for choosing the visually lossless route. The major case in point is in mass digitization efforts, converting print pages to digital images.

Consider the Digital Library Federation's (DLF) "Benchmark for Faithful Digital Reproductions of Monographs and Serials," which specifies 600dpi TIFF, compressed losslessly, but bitonal for text or line art, which represents the bulk of historical print materials, but far from everything. For more complex print situations, such as grayscale photos or color illustrations, the Benchmark recommends (but does not require) progressing up to grayscale and color as needed, albeit at only 300dpi [5]. Genealogically, this print capture standard essentially developed from the joint Michigan and Cornell Making of America Project (Phase I, circa 1996) and it has been the one implemented in major book digitization projects among DLF members and beyond [6].

More recently, the Internet Archive has developed its visually lossless JPEG 2000-based benchmark in concert with partner institutions, who helped settle on an all-color alternative, which eliminates the human factor of bit-level decisions from the moment of capture [7]. Surely anyone familiar with an activity even remotely as repetitive as scanning a book page-by-page can appreciate the fact that, in a three-level system, many color or grayscale pages of material will often slip through as the default bitonal by mistake. Thus, considering the choice of having either a visually lossless color JPEG 2000 or a losslessly compressed TIFF bitonal image of a page that should have been in color in

order to render features faithfully, then clearly the visually lossless compromise comes out ahead. This is not to say that in certain situations some may not still opt for bitonal, on the low end, or, on the high end, for mathematically lossless throughout for a particular object or set of objects. In a nutshell, the visually lossless color is at least as satisfactory as bitonal in the grand scheme of things, where realistically files need to fit on the servers allotted.

One of the primary goals driving the development of the JPEG 2000 standard was the unfilled need for scalability that would extend from a high resolution archival master to a lower resolution, web-deliverable browser image. The bifurcating path of preserving a bulky, high resolution TIFF for the master, then running processes to extract derivative files for user access is inherently inefficient. The survey results were striking in that, of the survey respondents who considered themselves implementers, the majority reported that they use JPEG 2000 to provide web access to material, while only a minority used it for archiving images. This was despite the fact that one of the main problems to date with the standard is the lack of browser support, while one of the chief advantages is its more efficient file size for high resolution mastering. A more efficient model than the TIFF/derivative method would be for the JPEG 2000 format to do both the heavy lifting of high resolution archiving as well as the delivery to the user.

## **Current Tools & Browser Support**

Adobe Photoshop with its free optional plug-in proved to be the most utilized JPEG 2000 file creation tool among practitioners. Feelings expressed on this score were that the plug-in was easy to use, could be integrated into batch processing, but could also be slow.

Interestingly however, beginning in 2007, Adobe themselves have questioned their own continued development of the plug-in in light of cameras not entering the market with native JPEG 2000 support, coupled with the standard's assumed minimal adoption among Photoshop users [8]. To date, Adobe plans to keep shipping the plug-in with its newest Photoshop versions, but will do so most likely as an optional installation (personal communication with John Nack of Adobe, February 19, 2009).

The digital collection management software, CONTENTdm, with its built-in JPEG 2000 converter was also a popular utility. In this case the tool's primary reported use, the ingest and subsequent conversion of pre-created high quality JPEG or TIFF archivals into access JPEG 2000 files, pointed to the fact that much of JPEG 2000's use at least within this community was as an access format.

Frustration was expressed on the current lack of native browser support for the standard. This focused primarily on the resulting server-side requirements that are needed in order to take advantage of the standard's flexible, zoom and panning capabilities from single JPEG 2000 files. In most cases, this dedicated server layer interprets a zoom scale request from the browser, then converts the stored JPEG 2000 into a format like JPEG or BMP that the browser can support and finally render to users. Respondent's comments included:

“Currently, very little client software and very few repositories seem to take advantage of the jpip protocol. This means that jpeg2000 images either need to be transformed on the server (for different regions, resolutions, etc.) to jpeg for example, or the whole image downloaded by the client before displaying native jpeg2000. It also means that features such as quality layers and region of interest are less likely to be taken advantage of as this information should be client/user preference and is difficult to efficiently communicate without a dedicated client/server protocol.”

Yet, among some there was confidence that the browsers would eventually come around. Indeed, though native support is currently absent from Internet Explorer, Mozilla/Firefox, Safari, and Chrome, the QuickTime plug-in for each can render JPEG 2000, though only at one zoom level. The authors feel that it is imperative for browser developers to bake into their code native JPEG 2000 support that includes the full range of image manipulations that the standard enables, such as broad panning and deep zooming. Since part of the design aspect of wavelet compression schemes, like that of JPEG 2000, involves pushing more computing to the user's viewing device and its software, and since the major developers involved tend to want their browsers' codebases to travel light as a competitive advantage, there exists a threshold for implementation that has not been kind to the image standard. Setting aside for a moment the unhandy option of dedicated JPEG 2000 servers, the extra code required in the browser has relegated JPEG 2000 to the realm of extensions and add-ons, most of which, like QuickTime, serve a much broader audience and do not take the potential functionality much beyond a zoom level or two.

## **IPR Barriers: Genuine Paper Chase Threats vs. Paper Tigers**

To a limited extent, the UConn survey revealed that patent claims surrounding JPEG 2000 remain a concern to some in our community. The blogosphere, not surprisingly, can go even farther, for example: “JPEG 2000 is doomed to failure because of the patent issue [9].” Yet even professionals much closer to the inner workings of the standard have viewed the issue as a significant barrier as recently as 2005 [10]. It is important to keep in mind several points while considering the legal implications of choosing JPEG 2000.

First, the Intellectual Property Rights (IPR) landscape of networked computing is fraught with or at least constantly borders potentially litigious issues with practically any conceivable standard. For still images alone, the earlier JPEG and GIF file formats have been no stranger to legal entanglements. In the case of GIF, what had been a free and open format became litigious when the patent holders changed their minds about that formerly open model [11]. Technically, it was the Lempel-Ziv-Welch (LZW) compression algorithm that was the specific patent card played, but the format, including its LZW compression scheme, had been freely open since 1987 when the patent holders suddenly exerted their rights in 1993, resulting in what is referred to as a “submarine patent claim.” In the case of JPEG, in 2002, a company claimed patent rights despite the existence of “prior art,” or public evidence that the company's claim was not in fact



original. Since then, in 2007, a patent mill has sought to squeeze the last drops of revenue from the final remaining patent recognized by the U.S. Patent Office from the mostly expired JPEG chest, and again prior art appears to be rendering this last claim invalid as well. In the final analysis, unpredictable human behavior will always threaten progress, and the best an organization can do is to take prudent steps to minimize the threat.

Fortunately, the JPEG Committee has indeed taken such proactive measures by having all contributors to the JPEG 2000 standard itself sign...

agreements by which they provide free use of their patented technology for JPEG 2000 Part 1 applications. During the standardization process, some technologies were even removed from Part 1 because of unclear implications in this regard. Although it is never possible to guarantee that no other company has some patent on some technology, even in the case of JPEG, unencumbered implementations of JPEG 2000 should be possible [12].

However, as it so happened, one of the signers did file a suit against a competitor, claiming patent infringement, but the District Court judge in the case ruled the patent claim invalid [13]. JPEG 2000, then, not only has the benefit of a foundation cleared for patent issues by its designers, but it has also thwarted an offensive by one of those designers—one who had been most likely to succeed in the crucial prior art category—and now enjoys a record of patent claims resistance. Moreover, the JPEG committee remains vigilant, seeking to identify any IPR claims regarding JPEG 2000, and regularly solicits information toward this end at each triannual meeting in their ongoing standards work. By documenting these claims (or more accurately the lack thereof) via regular updates, the case for future GIF-like submarine patent claims is severely curtailed, if not nullified.

## Recommendations

1. Compile an implementation registry, which would include contact information, of JPEG 2000 related digital projects in cultural heritage institutions (similar to the current METS and PREMIS implementation registries at the Library of Congress) [14][15].

2. Suggest the creation of a new set of JPEG 2000 benchmarks (e.g. NDNP's profile for newspapers) that could be referenced in collaborative projects, vendor RFPs, and grant applications [16]. Outline the standard's appropriate use as an archival and access solution by format, including:

- a. General collections books (e.g. Internet Archive: lossy color) [17]
- b. Special collections books (e.g. lossless color)
- c. Photographs
- d. Maps
- e. Image files migrated from other raster formats

3. Vet the above suggested benchmarks through a competent collaborative body, such as DLF, and pursue its stamp of approval.

4. Have the collaborative body identify and empower a liaison from among imaging experts to serve as an advocate for

JPEG 2000 to browser developers and imaging software developers.

5. Better educate the cultural heritage community about the soundness and advantages of JPEG 2000 in the context of possible format benchmarks.

## References

- [1] D. Lowe, and M. Bennett, "Digital Project Staff Survey of JPEG 2000 Implementation in Libraries" (2009), (retrieved 3.20.09). *UConn Libraries Published Works*. Paper 16. [http://digitalcommons.uconn.edu/libr\\_pubs/16](http://digitalcommons.uconn.edu/libr_pubs/16)
- [2] S. Abrams, S. Chapman, D. Flecker, S. Kreigsman, J. Marinus, G. McGath, and R. Wendler, "Harvard's Perspective on the Archive Ingest and Handling Test," *DLib Magazine*, 11, 12 (2005), (retrieved 3.18.09) <http://dlib.org/dlib/december05/abrams/12abrams.html>
- [3] "Preserving Digital Images (January-February 2008)" - Library of Congress Information Bulletin. (retrieved 3.17.2009) <http://www.loc.gov/loc/lcib/08012/preserve.html>
- [4] R. Miller, "Internet Archive (IA): Book Digitization and Quality Assurance Processes," confidential document for IA partners; permission to reference from R. Miller 3/16/2009.
- [5] Digital Library Federation, "Benchmark for Faithful Digital Reproductions of Monographs and Serials," (retrieved 3.18.2009), <http://purl.oclc.org/DLF/benchrepro0212>
- [6] E. Shaw and S. Blumson, "Making of America: Online Searching and Page Presentation at the University of Michigan," *D-Lib Magazine*, July/August (1997), (retrieved 3.18.2009), <http://www.dlib.org/dlib/july97/america/07shaw.html>
- [7] *ibid.*, R. Miller.
- [8] "John Nack on Adobe: JPEG 2000 - Do you use it?" (retrieved 3.17.2009), [http://blogs.adobe.com/jnack/2007/04/jpeg\\_2000\\_do\\_yo.html](http://blogs.adobe.com/jnack/2007/04/jpeg_2000_do_yo.html)
- [9] B. Dowling on February 18, 2007 04:07 AM on J. Atwood's blog "Coding Horror: programming and human factors," (retrieved 3.18.2009), <http://www.codinghorror.com/blog/archives/000794.html>
- [10] Y. Nomizu, "Current status of JPEG 2000 Encoder Standard," (retrieved 3.18.2009), [http://www.itscj.ipsj.or.jp/forum/forum20050907/5Current\\_status\\_of\\_JPEG%202000\\_Encoder\\_standard.pdf](http://www.itscj.ipsj.or.jp/forum/forum20050907/5Current_status_of_JPEG%202000_Encoder_standard.pdf)
- [11] M. C. Battilana, "The GIF Controversy: A Software Developer's Perspective," (retrieved 3.18.2009), <http://www.cloanto.com/users/mcb/19950127giflwz.html>
- [12] D. Santa Cruz, T. Ebrahimi, and C. Christopoulos, "The JPEG 2000 Image Coding Standard," (retrieved 3.18.2009), <http://www.ddj.com/184404561>
- [13] "Gray Cary Successfully Represents Earth Resource Mapping in Patent Dispute," from <http://www.businesswire.com> dated March 27, 2004, (accessed 3.18.2009 via LexisNexis: <http://www.lexisnexis.com/>)
- [14] "METS Implementation Registry: Metadata Encoding and Transmission Standard (METS) Official Web Site," (retrieved 3.23.09), <http://www.loc.gov/standards/mets/mets-registry.html>
- [15] "PREMIS Implementation Registry - PREMIS: Preservation Metadata Maintenance Activity (Library of Congress)," (retrieved 3.23.09), <http://www.loc.gov/standards/premis/premis-registry.php>
- [16] "NDNP JPEG 2000 Profile," (retrieved 3.23.09), <http://www.loc.gov/ndnp/pdf/JPEG2kSpecs.pdf>
- [17] *ibid.*, R. Miller.

## **Authors Biographies**

*David Lowe serves as Preservation Librarian at the University of Connecticut. A 1996 graduate of the School of Information at the University of Michigan, Lowe has worked with both analog and digital media as a Preservation staff member at the University of Michigan (1994-1997) and at Columbia University (1997-2005), before coming to UConn (2005- ), where his primary duties now involve preservation issues surrounding local digital collections.*

*Michael J. Bennett is Digital Projects Librarian & Institutional Repository Coordinator at the University of Connecticut. There he manages digital reformatting operations while overseeing the University's institutional repository. Previously he has served as project manager of Digital Treasures, a digital repository of the cultural history of Central and Western Massachusetts and as executive committee member for Massachusetts' Digital Commonwealth portal. He holds a BA from Connecticut College and an MLIS from the University of Rhode Island.*

# WHERE WE ARE TODAY: AN UPDATE TO THE UCONN SURVEY ON JPEG 2000 IMPLEMENTATION FOR STILL IMAGES

JPEG 2000 SUMMIT  
MAY 12, 2011  
LIBRARY OF CONGRESS  
WASHINGTON, D.C., USA

David Lowe & Michael J. Bennett  
University of Connecticut Libraries, Storrs, CT, USA

# Agenda

---

- The Survey
  - Background & Results
  - Interim Developments
  - Survey Takeaways
- Technical Considerations
  - Tools
  - Format Landscape

# Survey Background

- 2003-4: Aware-partnered JPEG2000 project and conference at UConn
- 2007: [Adobe product manager's blog](#) queries community re: its JPEG2000 support in Photoshop
- 2008: We develop and post [our survey](#) to gauge JPEG2000 acceptance status among cultural heritage institutions

# Survey Results

## 175 responses:

1. **How would you classify your institution?**
  - a. **77% Academic/Research Libraries**
  - b. **16% Public Libraries**
2. **Breakdown of “other”**
  - a. **Corporate**
  - b. **Service Bureaus**
  - c. **etc.**

# Survey Results

- 3. Do you use the JPEG2000 file format at all?**
  - a. 60% Yes (103 responses)**
  - b. 40% No (70 responses)**
  
- 4. If you do NOT use the JPEG 2000 file format at all, why not? (59 responses)**
  - a. Software shortcomings**
    - a. Creation/Manipulation**
    - b. Access (Browser Support)**
  - b. Lack of staff expertise**
  - c. Patent issues**

# Survey Results

5. **Do you use JPEG 2000 as an archival format for new collections?** (142 responses)
  - a. **20% Yes**
  - b. **80% No**
  
6. **Do you use JPEG 2000 as an archival format for images converted from legacy formats?** (141)
  - a. **16% Yes**
  - b. **84% No**



# Survey Results

- 7. Do you use JPEG 2000 to provide online access images for new collections?**
  - a. 53.5% Yes**
  - b. 46.5% No**
  
- 8. Do you use JPEG 2000 to provide online access for images converted from legacy formats?**
  - a. 46% Yes**
  - b. 54% No**

# Survey Results

- 9. What tools do you use in your JPEG 2000 workflows? (Please indicate all that apply)**
  - a. 53% Photoshop**
  - b. 37% CONTENTdm**
  - c. 19% IrfanView**
  - d. 18% Aware**
  - e. 17% Kakadu**
- 10. Note any tools not listed above.**
  - a. (No dark horses.)

# Survey Results

- 11. If you have migrated sets of files to JPEG2000 from legacy formats, what tools did you use?**
  - a. Photoshop
  - b. CONTENTdm
- 12. What do you see as the strengths of the available tools?**
  - a. Ease of use
- 13. What do you see as the weaknesses?**
  - a. Slowness

# Survey Results

- 14. What do you see as the weaknesses of the available tools?**
  - a. Browser support**
- 15. What do you see as viable, lasting alternatives to JPEG 2000 for archival master and/or access derivative copies?**
  - a. TIFF for archival**
  - b. JPG for access derivative**

# Survey Results

- 16. Do you use or would you consider *mathematically lossless* JPEG 2000 compression?**
  - a. Strong yes for archival**
  
- 17. Do you use or would you consider *visually lossless* JPEG 2000 compression for archival master purposes?**
  - a. Near half yes for archival**
  - b. No and undecided split the rest**

# Interim Developments

---

- Inline Browser support effort
- Djatoka Image Server
  - ▣ Local testing promising

# Survey Takeaways: Misconceptions

- **Lack of trust in JPEG2000 lossless compression**
  - *IS truly lossless*
- **File size savings not seen as significant vs. TIFF**
  - *Average of 1:2 in size savings vs. TIFF*
- **Lack of awareness of higher bit depth range**
  - *Includes 48 bit support*
- **Seen as lossy-only and proprietary**
  - *Neither is the case*
- **Note preference of JPEG2000 as access format**
  - *Designed to scale from archival through access derivatives*

# JPEG 2000: Some Hang-ups Persist

- Among general lack of software support, also notable is the lack of Adobe Lightroom support beyond one known plugin which spoofs the program.
- Limited native DAM software support “out of the box.” Dodgy performance once implemented.
- Damaging PR: Nov/Dec 2009 DIB article
- Lossy JPEG 2000 files prone to encoding errors in Adobe Photoshop when created in large batches.



# Lightroom JPEG 2000 Plugin

- <http://www.lightroom-plugins.com/JP2index.php>
- Plugin works by making Lightroom think that it is reading a TIFF instead of a JPEG2000 file.
- Renaming JPEG 2000 filenames using Lightroom subsequently doesn't work entirely correctly.
- Editing embedded metadata first requires the creation of a backup file.
- It is good, however, to see that clever work is being considered in this area.

# Limited DAM Support

- Most DAM software packages leverage ImageMagick, <http://www.imagemagick.org/script/index.php>, to do their heavy image manipulation lifting (see later slides for more on the problems of this with JPEG 2000)
- Many DAMs don't configure JPEG 2000 support out of the box.
- Many DAMs now have browser-based interfaces (both front and back end) and have a hard time displaying JPEG 2000 images as a result of browser JPEG 2000-rendering limitations.

# Damaging PR: DLib Article

- <http://www.dlib.org/dlib/november09/kulovits/11kulovits.html>
- Much-cited by JPEG 2000 skeptics.
- Article states its conclusions on JPEG 2000's weaknesses based primarily upon the performance of two open source tools...

ImageMagick TIFF to JP2000	Version: 6.5.3-5 2009-06-11 Q16	-compress Lossless -quality 100
GraphicsMagick TIFF to JP2000	Version: 1.3.6 2009-07-25 Q8	-compress Lossless -quality 100

# Damaging DLib Article

- Concludes that, “The direct pixels comparison, using both GraphicsMagick's and ImageMagick's compare functionality, indicated that pixels had been changed during migration from TIFF to JPEG 2000...”
- It is worth noting that van der Knijff (2010)\* has recently cited that among the tools he tested, ImageMagick was particularly poor in its ability to accurately interpret JPEG 2000 conversions based upon its JasPer JPEG2000 library dependencies.\*\*

\* <http://www.udfr.org/jp2kwiki/images/4/4f/Jp2kMigrationCharacterisationKBExternal.pdf>

\*\* <http://studio.imagemagick.org/discourse-server/viewtopic.php?f=1&t=15807>

# Damaging DLib Article

- van der Knijff goes on to state, “Of all tools, ImageMagick’s ‘Identify’ tool shows the poorest performance: the information it provides on resolution is erroneous and incomplete. It only detects ICC profiles that are of the ‘restricted’ type. Moreover, it reports non-existent ICC profiles when colour is defined using an enumerated (e.g. sRGB) colour space. Because of this, I would advise against the use of ImageMagick for the characterisation of JPEG2000 files.”\*

\* <http://www.udfr.org/jp2kwiki/images/4/4f/Jp2kMigrationCharacterisationKBExternal.pdf>

# Damaging DLib Article

- On the other hand, independent testing at UConn of TIFF > lossless JPEG 2000 conversion using the Photoshop CS4 JPEG 2000 plugin, confirms that JPEG 2000's lossless compression is **truly lossless at the pixel level** (using stacked TIFF & JPX layers of same image > toggling difference blending mode > histogram check).
- Murray (2007) has also previously done similar direct testing of Kakadu and also reports similar lossless JPEG 2000 compression results.\*

\* <http://dltj.org/article/lossless-jpeg2000/>

# Damaging DLib Article

---

















*So, resist the urge to completely judge the specification's attributes based solely on the performance of one of its more inconsistent tools.*

# JPEG 2000: Leveraging Lossless Compression

- A solid substitute for uncompressed TIFF archival files (for those who need rendered archival files and want to save storage space)
- In so doing, makes it easier to also archive raw DNG “safety masters” along with a rendered format (JPEG 2000). For a given image, storage footprint results in something smaller than a single uncompressed TIFF.



# Archival Storage Considerations

Name	Size	Date modified	Type	
 002.CR2	26,644 KB	5/3/2011 2:11 PM	CR2 File	Camera raws
 003.CR2	26,170 KB	5/3/2011 2:12 PM	CR2 File	
 004.CR2	26,724 KB	5/3/2011 2:13 PM	CR2 File	
 005.CR2	26,813 KB	5/3/2011 2:13 PM	CR2 File	
 002.dng	22,687 KB	5/5/2011 2:15 PM	DNG File	DNG raws
 003.dng	22,720 KB	5/5/2011 2:15 PM	DNG File	
 004.dng	23,652 KB	5/5/2011 2:15 PM	DNG File	
 005.dng	23,686 KB	5/5/2011 2:15 PM	DNG File	
 002.jpf	24,790 KB	5/5/2011 2:22 PM	JPF File	Lossless JP2000
 003.jpf	23,992 KB	5/5/2011 2:21 PM	JPF File	
 004.jpf	21,699 KB	5/5/2011 2:21 PM	JPF File	
 005.jpf	21,536 KB	5/5/2011 2:20 PM	JPF File	
 002.tif	61,621 KB	5/5/2011 2:24 PM	TIF File	Uncompressed TIFF
 003.tif	61,620 KB	5/5/2011 2:24 PM	TIF File	
 004.tif	61,620 KB	5/5/2011 2:25 PM	TIF File	
 005.tif	61,620 KB	5/5/2011 2:25 PM	TIF File	

# Archival Storage Considerations:

getting richer data preservation bang for your storage footprint buck

 002.dng

22,687 KB

 002.jpf

24,790 KB

*You can archive both the original latent raw image data & a losslessly rendered format...*

 002.tif

61,621 KB

*...all while using less storage space than a single uncompressed TIFF*

*47,477KB (DNG + JPF) vs. 61,621KB (TIF)*

# In Turn @ UConn Libraries...

For special collection printed & illustrated texts, and maps we reformat and archive in this manner:

- 1) *DNG raw “safety masters” (converted either from camera raws or native from scanners running VueScan)\**
- 2) *Lossless JPEG 2000 “archival masters” (reversible JPX, Photoshop)*
- 3) *Lossy JPEG 2000 “processed masters” (irreversible JP2, Photoshop)*

*\*For additional background on DNG as an archival format, see: [http://digitalcommons.uconn.edu/libr\\_pubs/23/](http://digitalcommons.uconn.edu/libr_pubs/23/)*

# Contact

---

David B. Lowe

Preservation Librarian

University of Connecticut Libraries

[david.lowe@uconn.edu](mailto:david.lowe@uconn.edu)

Michael J. Bennett

Digital Projects Librarian & Institutional Repository Manager

University of Connecticut Libraries

[michael.bennett@uconn.edu](mailto:michael.bennett@uconn.edu)

---

## D-Lib Magazine

May/June 2011

Volume 17, Number 5/6

---

### JPEG 2000 for Long-term Preservation: JP2 as a Preservation Format

Johan van der Knijff

KB / National Library of the Netherlands

[johan.vanderknijff@kb.nl](mailto:johan.vanderknijff@kb.nl)

doi:10.1045/may2011-vanderknijff

---

#### Abstract

Despite the increasing popularity of JPEG 2000 in the archival community, the suitability of the JP2 format for long-term preservation has been poorly addressed by existing literature. This paper demonstrates how some parts of the JP2 file specification (related to ICC profiles and grid resolution) contain ambiguous information, leading to a situation where different software vendors are interpreting the standard in slightly different ways. This results in a number of risks for preservation. These risks could be reduced by applying some minor changes to the format specification, in combination with the adherence to the updated standard by software vendors.

---

#### Introduction

The last few years have seen a marked rise in the use of JPEG 2000 in the cultural heritage sector. Several institutions are now using JPEG 2000 Part 1 (the JP2 format) as a preferred archival and access format for digital imagery. Examples include (but are not limited to) the National Library of the Netherlands ([Gillesse et al.](#), 2008), the British Library ([McLeod & Wheatley](#), 2007), the Wellcome Library ([Henshaw](#), 2010a), Library of Congress ([Buckley & Sam](#), 2006), the National Library of Norway ([National Library of Norway](#), 2007), and the National Library of the Czech Republic ([Vychodil](#), 2010). A number of other institutions are currently investigating the feasibility of using JP2 as a replacement of uncompressed TIFF, which is still the most widely used still image format for long-term archiving and preservation. In spite of the wide interest in JPEG 2000 from the archival community, the existing literature is surprisingly sparse on the actual suitability of the standard for long-term preservation. If preservation is addressed at all, what's often lacking is a specification of *what* information inside an image is worth preserving in the first place. Moreover, such discussions are often limited to largely theoretical considerations (e.g. features of the JP2 format), without going into the more practical aspects (e.g. to what extent do existing software tools actually follow the features that are defined by the format specification).

However, without taking such factors into account, can we say anything meaningful about how an image that is created using today's software will be rendered in, say, 30 years time? Also, at some point in the future it may be necessary to migrate today's images to a new format. How confident can we be about not losing any important information in this process? Alternatively, if we opt for emulation as a preservation strategy, how will the images behave in an emulated environment?

The above questions are central to this paper. There are many aspects to assessing the suitability of a file format for a particular preservation aim (see e.g. [LoC](#), 2007 and [Brown](#), 2008). In this paper I limit myself to

addressing two areas where the JP2 format specification can be interpreted in more than one way: support of ICC profiles and the definition of grid resolution. I demonstrate how these ambiguities have led to divergent interpretations of the format by different software vendors, and how this introduces risks for long-term preservation. I also present some possible solutions. Finally, I provide a number of practical recommendations that may help institutions to mitigate the risks for their existing collections.

Unless stated otherwise, the observations in this paper only apply to the JP2 file format, which is defined by JPEG 2000 Part 1 ([ISO/IEC, 2004a](#)).

---

## Colour management in JP2: restricted ICC profiles

Section I.3 of the JP2 format specification ([ISO/IEC, 2004a](#)) describes the methods that can be used to define the colour space of an image. The most flexible method uses ICC profiles, and is based on version ICC.1:1989-09 of the ICC specification ([ICC, 1998](#)). JP2 supports the use of ICC profiles for monochrome and three-component colour spaces (such as greyscale and RGB). However, JP2 does not support all features of the ICC standard. Instead, it uses the concept of a "Restricted ICC profile", which is defined as follows:

"This profile shall specify the transformation needed to convert the decompressed image data into the  $PCS_{XYZ}$ , and shall conform to either the Monochrome Input or Three-Component Matrix-Based Input profile class, and contain all the required tags specified therein, as defined in ICC.1:1998-09." ([ISO/IEC 2004a](#), Table I.9).

To appreciate what this actually means, it is helpful to give some additional information on the ICC standard. First of all, the ICC specification distinguishes 7 separate ICC profile classes. The most commonly used ones are the "Input Device" (or simply "Input"), "Display Device" ("Display") and "Output Device" ("Output") classes. Another one that is relevant in this context is the "ColorSpace Conversion" class. Second, it is important to know how colour transformations can be defined within the ICC standard. For monochrome images, the colour transformation is always described using a gray tone reproduction curve (TRC), which is simply a one-dimensional table. For RGB spaces, two methods are available. The first one is based on a three-component matrix multiplication. The second (N-component LUT-based method) uses an algorithm that includes a set of tone reproduction curves, a multidimensional lookup table and a set of linearisation curves ([ICC, 1998](#)).

Going back to the JP2 specification, the restrictions in the "Restricted ICC profile" class are:

1. For RGB colour spaces, N-component LUT-based profiles are not allowed (only three-component matrix-based profiles).
2. Only (device) *input* profiles are allowed (for both monochrome and RGB spaces)

The first restriction makes sense, since N-component LUT-based profiles are more complex than three-component matrix-based ones, and thus more difficult to implement. The logic behind the restriction of allowing only *input* profiles is more difficult to understand, since it prohibits the use of all other ICC profile classes. According to the ICC specification, the "input" class represents input devices such as cameras and scanners. However, widely used working colour spaces such as Adobe RGB 1998 ([Adobe, 2005](#)) and eciRGB v2 ([ECI, 2007](#)) are defined using profiles that belong to the *display* profile class. As a result, they are not allowed in JP2, even though both the Adobe RGB 1998 and eciRGB v2 profiles use the three-component matrix-based transformation method. Since there is no obvious reason for prohibiting such profiles, it would appear that the restriction to "input" profiles may be nothing more than an unintended error in the file specification. This impression is reinforced by the fact that the file specification of the JPX format (which is defined by JPEG 2000 Part 2) also consistently uses the phrase "*input* ICC profiles" in the definition of its "Any ICC profile" method (which doesn't have any restrictions on the use of N-component LUT-based profiles) ([ISO/IEC, 2004b](#)).

A major consequence of the "input" restriction is that a literal interpretation of the format specification limits

the use of ICC profiles to such a degree that any serious colour management becomes impossible in JP2. For colour imagery, the only colour space that can be handled without using ICC profiles is sRGB. Full-colour printed materials often contain colours that cannot be represented in the sRGB colour space. If such materials need to be digitised with minimal loss of colour fidelity, a colour space with a wider gamut (such as Adobe RGB or eciRGB) is needed, and this requires the use of ICC profiles. Since the format specification prohibits this, this means that – in its current form – the JP2 format is unsuitable for applications that require colour support beyond sRGB.

## Handling of ICC profiles by different encoders

In order to test how the most widely-used JPEG 2000 encoders handle ICC profiles in practice, I took a number of TIFF images that contain embedded ICC profiles, and tried to convert them to JP2 with the most widely used JPEG 2000 encoders. The ICC profiles in all experiments were display device profiles for Adobe RGB 1998 and eciRGB v2 working colour spaces, which both use the three-component matrix-based transformation method. I subsequently analysed all generated images using ExifTool 8.12 ([Harvey](#)) and JHOVE 1.4 ([JHOVE](#)). Table 1 summarises the results.

Software	Result
Luratech command line tool version 2.1.20.10 (jp2clt.exe)	Upon detection of non-"input" profile, output is automatically written in JPX format. ICC profile embedded using "Any ICC" method
Luratech command line tool version 2.1.22.0 (jp2clt.exe)	JP2 file; ICC profile embedded using "Restricted" method. The profile class of the original profile ("display device") is changed to "input device"
Adobe Photoshop CS4 using Adobe JPEG2000 (version: 2.0, 2007) plugin; JP2 compatible option <i>not</i> checked	JPX (JPEG 2000 Part 2) file; ICC profile embedded using "Any ICC" method
Adobe Photoshop CS4 using Adobe JPEG2000 (version: 2.0, 2007) plugin; JP2 compatible option checked	JPX file that contains two versions of the ICC profile: <ul style="list-style-type: none"> <li>○ Original profile using "Any ICC" method</li> <li>○ Modified version using "Restricted" method. The profile class of the original profile ("display device") is changed to "input device", and a "modified" prefix is added to the profile description field (e.g. "Modified eciRGB v2")</li> </ul>
Kakadu 6.3 (kdu_compress.exe)	JP2 file; ICC profile not embedded
ImageMagick 6.6.1-2 (convert.exe)	JP2 file; ICC profile not embedded
Aware JPEG 2000 SDK 3.18 (j2kdriver.exe)	JP2 file; ICC profile embedded using "Restricted" method

*Table 1: Preservation of ICC profiles in TIFF to JPEG 2000 migration using different encoders.*

We can make a couple of interesting observations from these results. First, 3 out of the 7 experiments resulted in a JPX file. JPX is an extension of the JP2 format that is defined by JPEG 2000 Part 2 ([ISO/IEC, 2004b](#)). Most JPX files can be read by JP2 decoders, which will simply ignore any features that are not permitted within JP2. It also contains a separate "Any ICC" method that – unlike JP2 – supports the use of N-component LUT-based

ICC profiles. Decoders that do *not* include JPX support will simply ignore ICC profiles that are defined using this method. At present very few decoders include support for JPX, and the adoption of the format is negligible. Because of this, the format is not well suited for preservation.

With this in mind, the behaviour of version 2.1.20.10 of the Luratech software (which was reported also by [Henshaw](#), 2010b) is somewhat odd. Depending on the characteristics of the input image, the encoder may decide to use the JPX format without any explicit instruction from the user to do so. Even worse, users may be completely unaware of this. Since the ICC profiles in all test images use the three-component matrix-based transformation, the only reason for not allowing them in JP2 would be the fact that they are not "input" profiles. However, since the "Any ICC" method in the format specification of JPX contains the very same "input" restriction, switching to JPX doesn't solve this problem. This behaviour has been corrected in more recent versions of Luratech's software. If version 2.1.22.0 of the encoder encounters a "display" profile in the input image, it writes a JP2 file, but it changes the "display" profile class value of the original profile to "input" in the resulting image [1].

Adobe's JPEG 2000 plugin for Photoshop only encodes to JPX format [2]. However, it has an option to create JPX files that are "JP2 compatible". When this option is activated, in addition to the original profile, it adds a modified version to the image, where the "display" class is simply changed to "input". So, these images contain two different versions of the same profile.

ICC profiles are lost altogether in the Kakadu and ImageMagick migrations. This is consistent with earlier results by [Kulovits et al.](#) (2009). I should add here that Kakadu *does* actually support the use of ICC profiles, but in an indirect way that requires the user to specify a profile's parameters on the command line.

Only the Aware encoder managed to create JP2 images that include embedded "display" ICC profiles without altering them in any way during the migration. So, only Aware and recent versions of the Luratech encoder currently permit basic colour management in the JP2 format. Aware achieves this by deviating from the JP2 format specification, whereas Luratech simply changes the profile class fields.

---

## Resolution headers

Most still image formats use straightforward, fixed header fields for describing the grid resolution of the image data. For JP2 (and the other JPEG 2000 formats) the situation is somewhat more complex, because it distinguishes two separate resolution types. They are both optional, and an image may contain any, both or neither.

First, there is a "capture resolution", which is defined as *"the grid resolution at which the source was digitized to create the image samples specified by the codestream"*. Two examples are given: the resolution of the flatbed scanner that captured a page from a book, or the resolution of an aerial digital camera or satellite camera ([ISO/IEC 2004a](#), Section I.5.3.7.1).

Second, there is a "default display resolution", which is defined as *"a desired display grid resolution"*. The specification states that *"this may be used to determine the size of the image on a page when the image is placed in a page-layout program"*. It then continues by warning that *"this value is only a default"*, and that *"each application must determine an appropriate display size for that application"* ([ISO/IEC 2004a](#), Section I.5.3.7.2).

The definition of these resolution types is problematic for a number of reasons. First of all, the use of the word "digitized" in the definition of "capture resolution" implies that it only covers analog-to-digital capture processes, such as the scanning of a printed photograph. However, in the case of born-digital materials there is no such analog-to-digital capture process, so the definition does not apply. A similar situation arises if we scan a photograph at, say, 300 ppi, and subsequently resample the resulting image to 150 ppi. Obviously the original image has a capture resolution of 300 ppi, but it is less clear where we should store the grid resolution of the



resampled image. One possibility would be to use the "default display" fields. However, the definition of "default display resolution" is rather vague, and it is difficult to understand what it means at all (e.g. what is "desired", and if this value is "only a default", what is this "default" based on?). My interpretation is that it is basically intended to allow reader applications to establish some sensible (but arbitrary) default zoom level upon opening the image. If this is correct, its value may be quite different from the grid resolution of the (either resampled or born-digital) image.

Semantic issues aside, the use of two separate sets of resolution fields also creates practical problems. First of all, it complicates the process of establishing the grid resolution of an image, since the location of this information ("capture" or "default display" fields) would become dependent on its creation history. Second, in the case of format migrations that may be part of imaging workflows as well as (future) preservation actions, there is no obvious mapping between the resolution fields of JP2 and other formats. Figure 1 illustrates this. Just as an example, most digitisation workflows still use TIFF for capture and intermediate processing, and the conversion to JP2 is only done as a final step. Since a TIFF image only has one set of resolution fields, to which JP2 fields should we map these values (taking into account that the TIFF may or may not have been resampled after capture)? Finally, there is the observation that, to the best of my knowledge, there is not a single example of a JPEG 2000 encoder that uses JP2's resolution fields in a manner that is consistent with the format specification. I will illustrate this in the next section.

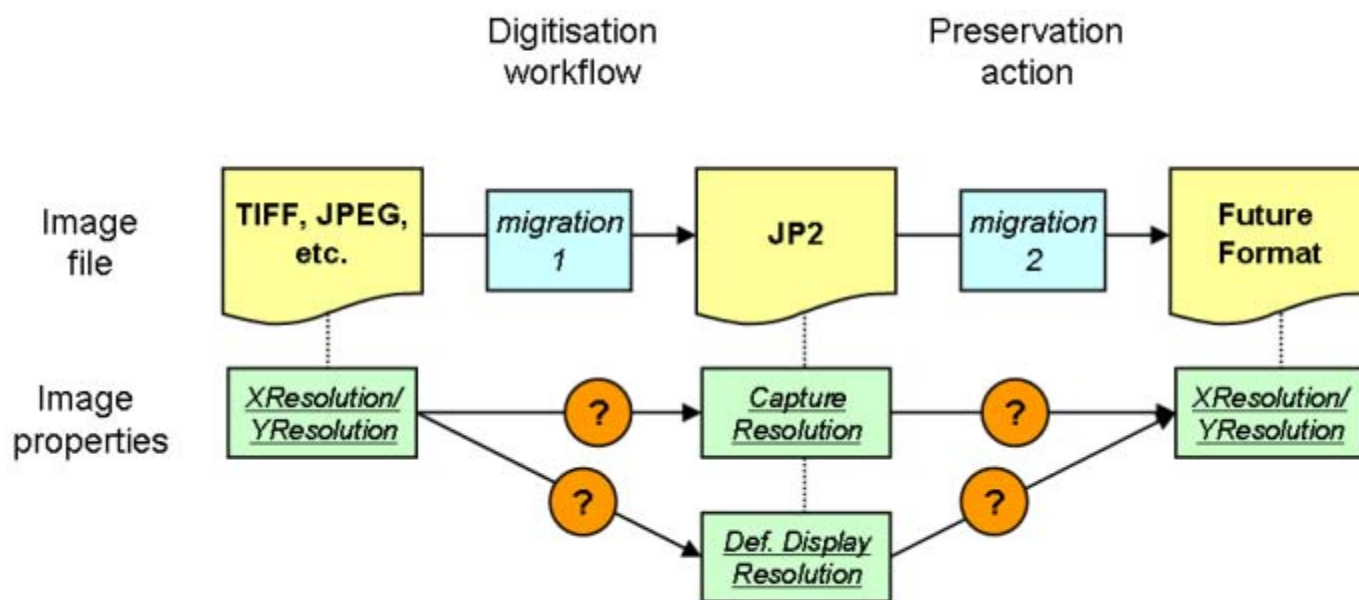


Figure 1: Mapping of resolution fields in migrations to and from JPEG 2000. Migration 1 is a typical TIFF to JP2 migration in a digitisation workflow; migration 2 represents a preservation action that involves a migration from JP2 to some future image format. In both cases, the mapping of the resolution fields before and after the migration is not clearly defined.

## Handling of resolution headers by different encoders

In order to find out how current encoders are handling the resolution fields in practice, I analysed how grid resolution is stored in the output images of the aforementioned TIFF to JPEG 2000 migration experiment. Table 2 shows the results.

Software	Capture resolution	Display resolution
Luratech command line tool version 2.1.20.10 (jp2clt.exe)	✓	—
Luratech command line tool version 2.1.22.0 (jp2clt.exe)	✓	—
Adobe Photoshop CS4 using Adobe JPEG2000 (version: 2.0, 2007) plugin	✓	—
Kakadu 6.3 (kdu_compress.exe)	—	✓
ImageMagick 6.6.1-2 (convert.exe)	—	—
Aware JPEG 2000 SDK 3.18 (j2kdriver.exe)	✓	—

*Table 2: Header fields used for storing grid resolution after TIFF to JPEG 2000 migration using different encoders.*

Luratech, Adobe and Aware always map the TIFF resolution fields to "capture resolution" in JPEG 2000. The ImageMagick files do not contain any resolution information at all. Only Kakadu always uses the "default display" fields. On a side note, Accusoft ImageGear, which uses the Kakadu libraries for writing JP2, also uses the "display" fields. This may apply to other Kakadu-based products as well. Crucially, none of these encoders use "capture resolution" in the way it is described in the format specification.

What these results show is that establishing the grid resolution of a JP2 image is not straightforward, because the location of this information is not well defined. It also shows that most encoders ignore the literal meaning of "capture resolution" in the JP2 format specification, and simply use these fields in a manner that is analogous to the TIFF resolution fields.

---

## Implications for preservation

### *ICC profiles*

In the previous sections I explained how the JP2 file specification appears to be unnecessarily restrictive with respect to embedded ICC profiles, and I demonstrated that different software vendors are handling these restrictions in a variety of ways. From a preservation point of view, the central issue here (as already stated in the introduction to this paper) is what may be the impact of this on rendering existing images in the future, and the preservation of information in any future migration to some new format. There are several problems here. First of all, a strict adherence to the format specification would simply rule out the use of ICC profiles in most cases. This would make the format unsuitable for any applications that require a colour gamut beyond sRGB space. The Aware encoder permits the use of JP2 for such applications by ignoring the "input" profile restriction. However, by doing so, such files no longer adhere to the format specification. Recent versions of Luratech's encoder *do* stick to the format specification, but enable the use of "display" ICC profiles by changing the profile class fields. The impact on future migrations, or the use of such files in an emulated environment will most likely be minor in both cases. An "input" profile defines a transformation from a device-dependent colour space to a universal profile connection space (PCS), whereas a "display" profile simply describes the reverse pathway (from the PCS to a device-dependent space). Technically, both are identical, and the colour transformation will be performed correctly even if the profile class label doesn't match the actual use. However, as for Aware's solution, one cannot completely rule out that future decoders *may* ignore embedded "display" profiles, which is a potential risk for future migrations. Luratech's current solution is also somewhat unsatisfactory, as it achieves adherence to the format specification by modifying (if only slightly) the original data.

Earlier versions of the Luratech encoder produce a JPX file if they encounter an ICC profile that doesn't adhere to the "restricted ICC" definition. As software support for JPX is so poor, there is a real risk that the ICC profiles will get lost in a future migration (even though the image data will most likely be preserved). Moreover, since the JPX file specification also limits the use of ICC profiles to the "input class", such files do not adhere to the JPX file specification either. The same applies to Adobe's implementation, although the risks are even greater for these files because of the use of an erroneous file type header field, which makes the handling of these files by current and future decoders largely unpredictable.

### *Resolution header fields*

Grid resolution does not directly affect the rendering of an image (unlike ICC profiles). Nevertheless, it is an important image property: for digitised imagery, resolution enables us to establish the dimensions of the digitised object. From a preservation point of view, the main risk that results from the current situation with JP2's resolution header fields is that resolution information may be lost in future migrations (see also Figure 1). For instance, a (future) decoder that expects grid resolution to be stored in the "capture" fields and ignores the "default display" fields will not be able to establish any meaningful resolution information from images that were created using current versions of Kakadu. Some tools will internally substitute the missing resolution fields with default values. For instance, if Adobe Photoshop cannot find the "capture resolution" fields, it assumes a default value of 72 ppi. If such files are subsequently re-saved, it will actually write this (entirely fictional) value to the resolution fields of the created file. Other tools may behave in a similar way, which introduces the risk that resolution information may change after a migration. Also, none of the existing encoders appear to follow the (strict) definitions of these fields in the file specification. The file specification allows the use of *both* sets of fields in one file. Although I am not aware of any existing applications that actually do this, the correct interpretation of the resolution information would get very confusing in that case.

---

## Way forward for ICC profile and resolution issues

Although the issues I reported here are relatively minor, they can have major consequences within a preservation context. However, both the ICC and the resolution issues could be largely fixed by making some small changes to the JP2 file specification. Regarding the ICC issue, the JPEG committee is already working on a proposal for extending the support of ICC profiles in JP2, and bringing it in line with the latest ICC specification. This would involve removing the "input" restriction in the "Restricted ICC" method, which would allow the use of "Display Device" profiles (Robert Buckley, personal communication). (The "Output Device" class would still be prohibited in that case, since it always uses N-component LUT-based profiles.)

As for the resolution issue, the solution may be as simple as slightly expanding the definition of "capture resolution". As explained before, the current definition only covers analog-to-digital capture processes. However, both the rasterisation of a vector drawing (born-digital material) and the resampling of an existing image can be seen as digital-to-digital capture processes. Hence, a possible solution would be to include such cases in the definition of "capture resolution", which could then be generalised as *"the grid resolution at which the source was captured to create the image samples specified by the codestream"*. This updated definition should then be illustrated using examples of both analog-to-digital and digital-to-digital capture processes. This would make these fields consistent with their *de facto* use by most existing encoders (as shown by Table 2). It would also ensure backward compatibility for existing files as they are produced by most encoders (except Kakadu, and some products that are based on the Kakadu libraries). The definition of "default display resolution" could either be made more specific, or, alternatively, these fields could be deprecated altogether.

In addition to these changes in the file specification, software vendors should be encouraged to produce encoders that are compliant with the (corrected) standard. The cultural heritage community could play an important role here by insisting on using software that is standards-compliant.

---

## Interim recommendations for existing collections

In the previous section I suggested a way forward, which requires actions from the standards body and the software industry. In the meantime, institutions that are currently using JP2 as a preservation format may take a number of steps to mitigate any future risks. For existing collections, it is essential that any features that may imply a risk are both known and documented. This documentation should at least answer the following questions:

- What is the file format (JP2, JPX)?
- Do the images contain ICC profiles?
- What are the underlying characteristics of the ICC profiles (profile description, matrix coefficients, and so on)?
- Are ICC profiles embedded using the "Restricted" or "Any ICC" method?
- Do the images contain multiple versions of the ICC profile?
- Which fields (if any) are used for storing the image's grid resolution?
- Which software was used to create the images?

Apart from the last one, all the above questions can be answered using freely available software tools. Particularly useful in this respect are ExifTool ([Harvey](#)) and JHOVE ([JHOVE](#)). Both tools are capable of giving the required information on file format and resolution fields. JHOVE does not give any direct information about embedded ICC profiles; however, it will tell whether an image makes use of the "Restricted" or "Any ICC" methods. On the other hand, ExifTool provides detailed information about embedded ICC profiles, but it doesn't tell what method was used. So both tools complement each other here.

The resulting documentation will be helpful for making a realistic assessment of long-term risks. It may also be a starting point for planning a medium-term preservation action, such as the normalisation to standards-compliant (i.e. compliant to an updated version of the standard) JP2 images. However, in the latter case one should be aware that such a normalisation procedure by itself introduces further risks of information loss. If done thoughtlessly, the long-term outcome may be worse than doing nothing at all.

For new and ongoing digitisation projects, the most sensible interim recommendations would be to stick to the JP2 format whenever possible, avoid JPX, embed ICC profiles using the "Restricted" method, and avoid multiple ICC profile versions. In addition, the aforementioned recommendations for existing collections all apply here as well.

---

## Conclusions

In this paper I showed that the current JP2 format specification leaves room for multiple interpretations when it comes to the support of ICC profiles, and the handling of grid resolution information. This has led to a situation where different software vendors are implementing these features in different ways. In the case of ICC profiles, a strict interpretation of the standard even completely prohibits the use of ICC profiles for defining working colour spaces, which would make the format unsuitable for any applications that require colour support beyond the sRGB colour space. For preservation, this results in a number of risks, because images may not be rendered properly by future viewers, and colour space and resolution information may be lost in future migrations.

These issues could be remedied by some small adjustments of JP2's format specification, which would create minimal backward compatibility problems, if any at all. For the ICC profile issue, a proposal for such an adjustment is already under way from the JPEG committee, and I have suggested a possible solution for the resolution issue here. In addition, it would be necessary that software vendors adhere to the modified standard. Small as they may be, such changes could significantly improve the suitability and acceptance of JP2 as a preservation format.

---

## Acknowledgements

I would like to thank Hans van Dormolen (KB) for sharing his observations on various problems related to the handling of ICC profiles and grid resolution. This ultimately served as the impetus for much of the research presented here. Thanks are also due to Christy Henshaw, Laurie Auchterlonie and Ben Gilbert (Wellcome Library) for providing the Luratech 2.1.22 test images. Wouter Kool (KB) is thanked for providing the Luratech 2.1.20 test images. Jack Holm (International Color Consortium) and Axel Rehse (LuraTech Imaging GmbH) are thanked for their helpful comments and suggestions on the "input"- "display" profile issue. Thomas Richter (Accusoft Pegasus) and Scott Houchin (Aerospace Corporation) are thanked for sharing their thoughts on the capture resolution issue, which guided me to towards the current proposed solution. Robert Buckley (Rob Buckley Consulting), Richard Clark (Elysium Ltd) and Barbara Sierman (KB) are all thanked for their feedback on an earlier draft of this paper.

---

## Notes

[n1] The Luratech software also does this for JPX files, which means it is standards-compliant for both formats.

[n2] Although the Adobe plugin produces files that contain features which are only allowed in JPX, it assigns an erroneous value to the "Brand" header field that uniquely identifies a JPEG 2000 file as either JP2 or JPX. As a result, these files are neither valid JP2 nor JPX. Moreover, any file identification tools that are based on byte signatures ("magic numbers") will identify these files as JP2, even though the real format is JPX.

---

## References

- [1] Adobe. Adobe RGB (1998) Color Image Encoding Version 2005-05. San Jose: Adobe Systems Inc., 2005. 29 Dec 2010 <http://www.adobe.com/digitalimag/pdfs/AdobeRGB1998.pdf>.
- [2] Brown, A. Digital Preservation Guidance Note 1: Selecting file formats for long-term preservation. London: The National Archives, 2008. 5 Jan 2011 <http://www.nationalarchives.gov.uk/documents/selecting-file-formats.pdf>.
- [3] Buckley, R. & Sam, R. JPEG 2000 Profile for the National Digital Newspaper Program. Washington: Library of Congress Office of Strategic Initiatives, 2006. 27 Dec 2010 [http://www.loc.gov/ndnp/guidelines/docs/NDNP\\_JP2HistNewsProfile.pdf](http://www.loc.gov/ndnp/guidelines/docs/NDNP_JP2HistNewsProfile.pdf).
- [4] ECI. eciRGB\_v2 - the update of eciRGB 1.0 - Background information. European Color Initiative, 2007. 29 Dec 2010 <http://www.eci.org/doku.php?id=en:colourstandards:workingcolorspaces>.
- [5] Gillese, R., Rog, J. & Verheusen, A. Alternative File Formats for Storing Master Images of Digitisation Projects. Den Haag: Koninklijke Bibliotheek, 2008. 27 Dec 2010 [http://www.kb.nl/hrd/dd/dd\\_links\\_en\\_publicaties/publicaties/Alternative\\_File\\_Formats\\_for\\_Storing\\_Masters\\_2\\_1.pdf](http://www.kb.nl/hrd/dd/dd_links_en_publicaties/publicaties/Alternative_File_Formats_for_Storing_Masters_2_1.pdf).
- [6] Harvey, P. ExifTool. 30 Dec 2010 <http://www.sno.phy.queensu.ca/~phil/exiftool/>.
- [7] Henshaw, C. We need how much storage? London: Wellcome Library, 2010a. 27 Dec 2010 <http://jpeg2000wellcomelibrary.blogspot.com/2010/06/we-need-how-much-storage.html>.
- [8] Henshaw, C. Finding a JPEG 2000 conversion tool. London: Wellcome Library, 2010b. 30 Dec 2010 <http://jpeg2000wellcomelibrary.blogspot.com/2010/07/finding-jpeg-2000-conversion-tool.html>.
- [9] ICC. Specification ICC.1:1998-09 – File Format for Color Profiles. International Color Consortium, 1998. 29 Dec 2010 [http://www.color.org/ICC-1\\_1998-09.pdf](http://www.color.org/ICC-1_1998-09.pdf).

- [10] ISO/IEC. "Information technology — JPEG 2000 image coding system: Core coding system". ISO/IEC 15444-1, Second edition. Geneva: ISO/IEC, 2004a. 28 Dec 2010 <http://www.jpeg.org/public/15444-1annexi.pdf> ("Annex I: JP2 file format syntax" only).
- [11] ISO/IEC. "Information technology — JPEG 2000 image coding system: Extensions". ISO/IEC 15444-2, First edition. Geneva: ISO/IEC, 2004b. 28 Dec 2010 <http://www.jpeg.org/public/15444-2annexm.pdf> ("Annex M: JPX extended file format syntax" only).
- [12] JHOVE - JSTOR/Harvard Object Validation Environment. 30 Dec 2010 <http://hul.harvard.edu/jhove>.
- [13] Kulovits, H., Rauber, A., Kugler, A., Brantl, M., Beinert, T. & Schoger, A. "From TIFF to JPEG 2000? Preservation Planning at the Bavarian State Library Using a Collection of Digitized 16th Century Printings". D-Lib Magazine 15.11/12 (2009). 27 Dec 2010 [doi:10.1045/november2009-kulovits](https://doi.org/10.1045/november2009-kulovits).
- [14] LoC. "Sustainability Factors". Sustainability of Digital Formats - Planning for Library of Congress Collections. Washington: Library of Congress, 2007. 5 Jan 2011 <http://www.digitalpreservation.gov/formats/sustain/sustain.shtml>.
- [15] McLeod, R. & Wheatley, P. Preservation Plan for Microsoft — Update Digital Preservation Team. London: British Library, 2007. 27 Dec 2010 <http://www.bl.uk/aboutus/stratpolprog/ccare/introduction/digital/digpresmicro.pdf>.
- [16] National Library of Norway. Digitization of books in the National Library — methodology and lessons learned. Oslo: National Library of Norway, 2007. 27 Dec 2010 [http://www.nb.no/content/download/2326/18198/version/1/file/digitizing-books\\_sep07.pdf](http://www.nb.no/content/download/2326/18198/version/1/file/digitizing-books_sep07.pdf).
- [17] Vychodil, B. "JPEG2000 - Specifications for The National Library of the Czech Republic". Seminar JPEG 2000 for the Practitioner. London: Wellcome Trust, 16 Nov 2010. 27 Dec 2010 [http://www.dpconline.org/component/docman/doc\\_download/520-jp2knov2010bedrich](http://www.dpconline.org/component/docman/doc_download/520-jp2knov2010bedrich).

---

## About the Author



Johan van der Knijff is a digital preservation researcher at the Koninklijke Bibliotheek, National Library of the Netherlands. His work focuses around preservation-related aspects of digital file formats. He holds an MSc in Physical Geography from Utrecht University (NL), where he specialised in hydrology, geographical information systems and remote sensing. Johan previously worked on the development of hydrological simulation models at the European Commission's Joint Research Center.

---

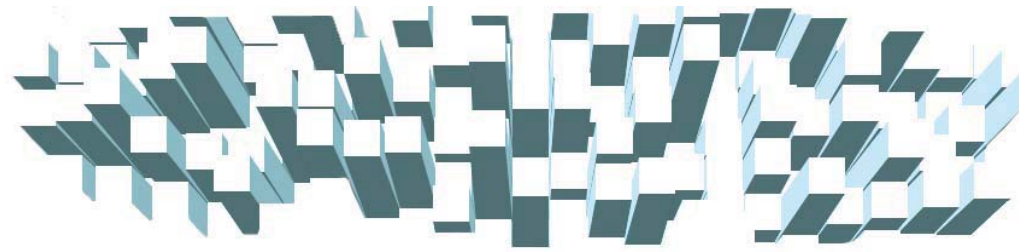
Copyright © 2011 Johan van der Knijff

---

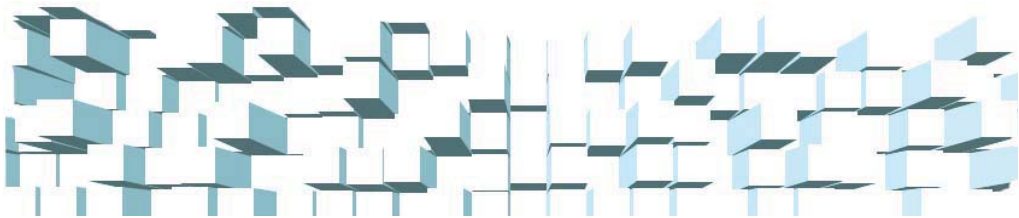
PRINTER-FRIENDLY FORMAT

[Return to Article](#)





JPEG 2000 as a  
Preservation and Access Format  
for the Wellcome Trust Digital Library



Robert Buckley  
Xerox Corporation

Edited by Simon Tanner  
*King's Digital Consultancy Services*  
*King's College London*



[www.kdcs.kcl.ac.uk](http://www.kdcs.kcl.ac.uk)

# Contents

1. SUMMARY OF ISSUES/QUESTIONS .....	3
2. RECOMMENDATIONS.....	4
3 BASIS OF RECOMMENDATIONS/REASONINGS/TESTS DONE.....	5
3.1 COMPRESSION.....	5
3.2 MULTIPLE RESOLUTION LEVELS .....	5
3.3 MULTIPLE QUALITY LAYERS.....	6
3.4 EXAMPLE: TIFF TO JP2 CONVERSION .....	6
3.5 MINIMALLY LOSSY COMPRESSION .....	7
3.6 TESTING REVERSIBLE AND IRREVERSIBLE COMPRESSION.....	7
3.7 FURTHER COMPRESSION FINDINGS.....	8
4 IMPLEMENTATION SOLUTIONS / DISCUSSION.....	9
4.1 COLOR SPECIFICATION.....	9
4.2 CAPTURE RESOLUTION.....	10
4.3 METADATA .....	10
4.3 SUPPORT .....	11
5 CONCLUSION .....	12
APPENDIX 1: JPEG 2000 DATASTREAM PARAMETERS .....	13
APPENDIX 2: REVERSIBLE AND IRREVERSIBLE COMPRESSION .....	14
FIGURE 1: SAMPLE IMAGES.....	15
FIGURE 2: COMPARISON OF IRREVERSIBLE WITH REVERSIBLE COMPRESSION.....	17



## 1. Summary of Issues/questions

The Wellcome Trust is developing a digital library over the next 5 years, anticipating a storage requirement for up to 30 million images. The Wellcome previously has used uncompressed TIFF image files as their archival storage image format. However, the storage requirement for many millions of images suggests that a better compromise is needed between the costs of secure long-term digital storage and the image standards used. It is expected that by using JPEG2000, total storage requirements will be kept at a value that represents an acceptable compromise between economic storage and image quality. Ideally, JPEG2000 could serve as both a preservation format and as an access or production format in a write-once-read-many type environment.

JPEG2000 was chosen as an image preservation format due to its small size and because it offers intelligent compression for preservation and intelligent decompression for access. If a lossy format is used to obtain a relatively high compression, e.g. between 5:1 and 20:1 (in comparison to an uncompressed TIFF file), then the storage requirements desired are achievable. The questions to address are what level of compression is acceptable and delivers the desired balance of image quality and reduced storage footprint.

With regard to the use of JPEG 2000, the questions posed in the brief and addressed in this report are:

- a. What JPEG2000 format(s) is best suited for preservation?
- b. What JPEG2000 format(s) is best suited for access?
- c. Can any single JPEG2000 format adequately serve both preservation and access?
- d. What models exist for the use of descriptive and/or administrative metadata with JPEG2000?
- e. If a JPEG2000 format is recommended for access purposes, what tools can be used to display/manipulate/manage it and any associated or embedded metadata?

This report will describe how a unified approach can enable JPEG2000 to serve for both preservation and access and balance the needs for compressed image size, image quality and decompression performance.

## 2. Recommendations

The majority of materials that will form part of the Wellcome Digital Library are expected to use visually lossless JPEG 2000 compression. Although “visually lossless” compression is lossy, the differences it introduces between the original and the image reconstructed from a compressed version of it are either not noticeable or insignificant and do not interfere with the application and usefulness of the image. Because the original cannot be reconstructed from the compressed image, compression in this case is irreversible. JPEG compression in digital still cameras is a familiar example of irreversible but visually lossless image compression. In mass digitization projects that use JPEG2000, compression ratios around 40:1 have been used for basically textual content. When applied to printed books, it has been found that these compression ratios do not impair the legibility or OCR accuracy of the text.

However, archiving and long-term preservation indicate a more conservative approach to compression and a different trade-off between compressed image size and image quality to meet current and anticipated uses. Still, given the volumes of material being digitized, a lossy format represents an acceptable compromise between quality and economic storage. A compression ratio of 4:1 or 5:1 gives a conservative upper limit on file size and decompressed image quality in the preservation format for the material being digitized. However this material should tolerate higher compression ratios with the results remaining visually lossless.

While most of the materials will use visually lossless compression, it is suggested that a small subset of materials (less than 5% of total) may be candidates for lossless or reversible compression. Reversible means that the original can be reconstructed exactly from the compressed image, i.e. the compression process is reversible.

Nevertheless, this report recommends irreversible JPEG2000 compression for the preservation and access formats of single grayscale or color images. Initially specifying a minimally lossy datastream will result in overall compression ratios around 4:1; the exact value will depend on image content. While this is a particularly conservative compression ratio, the compression can be increased as new materials are captured and even applied retroactively to files with previously captured material. The access format will be a subset of the preservation format with a subset of the resolution levels and quality layers in the preservation format.

In particular, the JPEG 2000 datastream should have the following properties:

- Irreversible compression using the 9-7 wavelet transform and ICT (see Section 3.1) with minimal loss (see Section 3.5)
- Multiple resolutions levels: the number depends on the original image size and the desired size of the smallest image derived from the JPEG 2000 datastream (see Section 3.2)
- Multiple quality layers, where all layers gives minimally lossy compression for preservation (see Sections 3.3 and 3.5)
- Resolution-major progression order (see Section 3.4)
- Tiles for improved codec (coder-decoder) performance, although the final decision regarding the use of tiles and precincts depends on the codec
- Generated using Bypass mode, which creates a compressed datastream that takes less time to compress and decompress (see Section 3.6)
- TLM markers (see Section 3.7)

A formal specification of the JPEG 2000 datastream for this application is given in Appendix 1. The datastream specified there is compatible with Part 1 of the

JPEG 2000 standard; none of the JPEG 2000 datastream extensions defined in Part 2 of the standard are needed.

Further, this report recommends embedding the JPEG 2000 datastream in a JP2 file. The JP2 file should contain:

- A single datastream containing a grayscale or color image whose content can be specified using the sRGB color space (or its grayscale or luminance-chrominance analogue) or a restricted ICC<sup>1</sup> profile, as defined in the JP2 file format specification in Part 1 of the JPEG 2000 standard (see Section 4.1)
- A Capture Resolution value (see Section 4.2)
- Embedded metadata that describes the JPEG 2000 datastream should follow the ANSI/NISO Z39.87-2006 standard and be placed in a XML box following the FileType box in a JP2 file (see Section 4.3)

Using the JP2 file format is sufficient as long as the requirement is for a single datastream whose color content can be specified using sRGB or a restricted ICC profile. While the JPX file format can be used if the color content of the image is specified by a non-sRGB color space or a general ICC profile, the use of a JP2-compatible file format is recommended.

## 3 Basis of recommendations/reasonings/tests done

### 3.1 *Compression*

In general terms, the compression ratio is set for preservation and quality, and JPEG 2000 datastream parameters such as the number of resolution levels and quality layers and tile size are set for access and performance. JPEG 2000 offers smart decompression, where only that portion of the datastream needed to satisfy the requested image view in terms of resolution, quality and location need be accessed and decompressed on demand and just in time.

The JPEG 2000 compression offers both reversible and irreversible compression. Reversible compression in JPEG 2000 uses the 5-3 integer wavelet transform and a reversible component transform (RCT). If no compressed data is discarded, then the original image data is recoverable from the compressed datastream created using these transforms. Irreversible compression uses the 9-7 floating point wavelet transform and an irreversible component transform (ICT), both of which have round-off errors so that the original image data is not recoverable from the compressed datastream, even when no compressed data is discarded. Appendix 2 contains a more detailed discussion of the differences between reversible and irreversible compression in JPEG 2000.

### 3.2 *Multiple resolution levels*

To begin with, it is recommended that JPEG 2000 be used with multiple resolution levels. The first two or three resolution levels facilitate compression; levels beyond that give little more compression but are added so that decompressing just the lowest resolution sub-image in the JPEG 2000 datastream gives a thumbnail of a desired size. For example, with a 5928-by-4872 pixel dimension original and 5 resolution levels, the smallest sub-image would have dimensions that would be 1/32 those of the original, in this case 186 by 153 pixels, which is roughly QQVGA<sup>2</sup> sized. Accordingly, JPEG 2000

---

<sup>1</sup> International Color Consortium, an organization that develops and promotes color management using the ICC profile format ([www.color.org](http://www.color.org))

<sup>2</sup> Quarter-quarter VGA (Video Graphics Array); since VGA is 640 by 480, QQVGA is 160 by 120.

compression with 5 resolution levels is recommended for images of this and similar sizes, which are typical of the sample images provided. In practice, the number of resolution levels would vary with the original image size so that the lowest resolution sub-image has the desired dimensions.

### 3.3 *Multiple quality layers*

There are two main reasons for using multiple quality layers. One is so that it is possible to decompress fewer layers and therefore less compressed data when accessing lower resolution sub-images. This speeds up decompression without affecting quality since the incremental quality due to the discarded layers is not noticeable at reduced resolutions. The second reason is that multiple quality layers make it possible to deliver subsets of the compressed image corresponding to higher compression ratios, which may be acceptable in some applications. This means there is less data to transmit and process, which improves performance and reduces access times. It also means that it is possible for the access format to be a subset of the preservation format, derived from it by discarding quality layers as the application and quality requirements warrant. The use of quality layers makes it possible to retroactively reduce the storage needs should they be revised downward by discarding quality layers in the preservation format and turning images compressed at 4:1 or 5:1 for example into images compressed at 8:1 or higher, depending on where the quality layer boundaries are defined.

### 3.4 *Example: TIFF to JP2 conversion*

For example, the following command line uses the Kakadu<sup>3</sup> compress function (`kdu_compress`) to convert a TIFF image to a JP2 file that contains an irreversible JPEG 2000 datastream. In particular, it contains a lossy JPEG 2000 datastream with 5 resolution levels and 8 quality layers, corresponding to compression ratios of 4, 8, 16, 32, 64, 128, 256 and 512 to 1 for a 24-bit color image. These correspond to compressed bit rates of 6, 3, 1.5, 0.75, 0.375, 0.1875, 0.09375 and 0.046875 bits per pixel. (A compression ratio of 4 to 1 applied to a color image that originally had 24 bits per pixel means the compressed image will equivalently have a compressed bit rate of 6 bits per pixel.) The Kakadu command line use bit rates rather than compression ratios to specify the amount of compression.

```
kdu_compress -i in.tif -o out.jp2 -rate  
6,3,1.5,0.75,0.375,0.1875,0.09375,0.046875 Creversible=no  
Clevels=5 Stiles={ 1024,1024 } Cblk={ 64,64 } Corder=RPCL
```

The JPEG 2000 datastream created in this example has 1024-by-1024 tiles, 64-by-64 codeblocks and a resolution-major progressive order RPCL, so that the compressed data for the lowest resolution (and therefore smallest) sub-image occurs first in the datastream, followed by the compressed data needed to reconstruct the next lowest resolution sub-image and so on. This data ordering means that the data for a thumbnail image occurs in a contiguous block at the start of the datastream where it can be easily and speedily accessed. This data organization makes it possible to obtain a screen-resolution image quickly from a megabyte or gigabyte sized image compressed using JPEG 2000. Tiles and codeblocks are used to partition the image for processing and make it possible to access portions of the datastream corresponding to sub-regions of the image.

---

<sup>3</sup> <http://www.kakadusoftware.com/>

### 3.5 *Minimally Lossy Compression*

The JPEG 2000 coder in this example would discard transformed and compressed data to obtain a compressed file size corresponding to 4:1 compression. This needs to be compared with the performance of the the minimally lossy coder, where no data is discarded but which is still lossy because of the use of the irreversible transforms. In some cases, depending on the image content, as shown in Section 3.6, the minimally lossy coder can give higher compression ratios than 4:1. Accordingly, it is recommended that a minimally lossy format with multiple quality layers and multiple resolution levels be used for the preservation format. The access format would use reduced quality subsets of the preservation format optionally obtained by discarding layers and using reduced resolution levels.

### 3.6 *Testing reversible and irreversible compression*

The reason to use irreversible compression is that it gives better compression than reversible compression, at the cost of introducing errors (or differences) in the reconstructed image. This section examines this performance tradeoff.

Reversible and irreversible compression were applied to four images provided by the Wellcome Digital Library (Figure 1). A variation on irreversible compression was tested which used coder bypass mode, in which the coder skipped the compression of some of the data. This gave a little less compression, but made the coder (and decoder) run about 20% faster. The Kakadu commands used in these tests are given in Appendix 2.

The compression ratios obtained with these three test are shown in the following table.

Original	Reversible	Irreversible	Irreversible w/bypass
L0051262_Manuscript_Page	2.25	3.45	3.42
L0051320_Line_Drawing	1.82	2.52	2.51
L0051761_Painting	2.46	3.96	3.90
L0051440_Archive_Collection	2.52	4.47	4.41

For these particular images, the compression ratio for irreversible JPEG 2000 was from about 40% to almost 80% better than it was for reversible, and on average over 30% faster (with a further 20% boost with coder bypass mode).

The cost of irreversible compared to reversible is the error it introduces. The error or difference between the reversibly and irreversibly compressed images is about 50 dB, which means the average absolute error value was about 0.5. For one of the sample images, 99.99% of the green component values were the same after decompression as they were before, or at most two counts different. (For the red and blue components, the percentages were 99.79 and 99.35.) This is within the tolerance for scanners: in other words, minimally lossless irreversible JPEG 2000 compression adds about as much noise to an image as a good scanner does.

A region was cropped from one image so that the visual effects of this error on this image could be examined more closely (Figure 2). When they were, the differences were not perceptible on screen or on paper. Unless being able to reconstruct the original scan is a requirement, legal or otherwise, then irreversible compression is clearly advantaged over reversible compression.

### 3.7 Further compression findings

In these tests, the compressed file sizes (and compression ratios) were image dependent and varied with the image content. Images with less detail or variation than these samples would give even higher compression ratios.

An advantage of JPEG 2000 is that it lets the user set the compression ratio, or equivalently the compressed file size, to a specific target value, which the coder achieves by discarding compressed image data. While this feature was not used to set the overall compressed file size in the minimally lossy compression case, it can be used to set the sizes of intermediate images corresponding to the different quality layers. The following Kakadu command line generates a JP2 file with a minimally lossy irreversible JPEG 2000 datstream that complies with the recommendation in this report:

```
kdu_compress -i in.tif -o out.jp2 -rate -, 4, 2.34, 1.36, 0.797,
0.466, 0.272, 0.159, 0.0929, 0.0543, 0.0317, 0.0185
Creversible=no Clevels=5 Stiles={ 1024,1024 } Cblk={ 64,64 }
Corder=RPCL Cmodes=BYPASS
```

The JPEG 2000 datstream in this example has 5 resolution levels and 12 quality layers. Using all 12 layers give a decompressed image with minimal loss. The intermediate layers boundaries are at pre-set compressed bit rates, starting at 4 bits per pixel, corresponding to a compression ratio of 6:1, assuming a 24-bit color original. Thereafter, the layer boundaries are distributed logarithmically up to a compression ratio of 1296:1. The exact values are not critical. What is important is the range of values and there being sufficient values to provide an adequate sampling within the range.

When a datstream has multiple quality layers, it is possible to truncate it at points corresponding to the layer boundaries and obtain derivative datstreams that correspond to higher compression ratios (or lower compressed bit rates). In the previous example, discarding the topmost quality layer produces a datstream corresponding to a compression ratio of 6:1 (compressed bit rate of 4 bits per pixel). Discarding the next layers produces a datstream with a compression ratio of 10.3:1, and so on. As noted previously, some images may have minimally lossy compression ratio greater than 6:1; the layer settings can be adjusted when this happens.

Using layers adds overhead that increases the size of the datstream and therefore decreases the compression ratio. To assess this effect as well as the overhead due to the use of tiles, the four sample images were compressed with one layer and no tiles, with one layer and 1024x1024 tiles, and with 12 layers and no tiles. As the following table shows, adding layers and tiles did decrease the minimally lossy compression ratio, but the effect was only visible in the third place after the decimal and was therefore judged insignificant in comparison to the advantages of using them.

Original	No tiles 1 layer	1024x1024 tiles 1 layer	No tiles 12 layers
L0051262_Manuscript_Page	3.452	3.450	3.443
L0051320_Line_Drawing	2.522	2.521	2.517
L0051761_Painting	3.961	3.957	3.948
L0051440_Archive_Collection	4.477	4.473	4.461

Besides tiles and layers, other datstream components that can improve performance and access within the datstream are markers, such as Tile Length Markers (TLM) which can aid in searching for tile boundaries in a datstream. Their effectiveness depends on whether or not the decoder or



access protocol makes use of them. As a result, recommendations regarding their use depend on the choice of codec.

## 4 Implementation solutions / discussion

This section discusses the file format and metadata recommendations.

One function of a file format is packaging the datastream with metadata that can be used to render, interpret and describe the image in the file. Besides defining the JPEG 2000 datastream and core decoder, Part 1 of the JPEG 2000 standard also defines the JP2 file format which applications may use to encapsulate a JPEG 2000 datastream. A minimal JP2 file consists of four structures or "boxes":

1. JPEG 2000 Signature Box, which identifies the file as a member of the JPEG 2000 file format family
2. File Type Box, which identifies which member of the family it is, the version number and the members of the family it is compatible with
3. JP2 Header Box, which contains image parameters such as resolution and color specification needed for rendering the image
4. Contiguous Codestream Box, which contains the JPEG 2000 datastream

### 4.1 *Color Specification*

How an image was captured or created determines the parameters in the JP2 Header Box, which are subsequently used to render and interpret the image. Among these parameters are the number of components (i.e. whether the image is grayscale or color), an optional resolution value for capture or display, and the color specification. In general the color content of an image can be specified in one of two ways: directly using a named color space, such as sRGB, Adobe RGB 98 or CIELAB, or indirectly using an ICC profile.

The digitization process and the nature of the material being digitized, not the file format, drive the color specification requirements of the application. The issue for the file format is whether or not it supports the color encoding used by the digital materials. What's significant about the JP2 file format is that it supports a limited set of color specifications. For example, the only color space it supports directly is sRGB, including its grayscale and luminance-chrominance analogues. This is a consequence of the JP2 file format having been originally designed with digital cameras in mind.

Besides sRGB, the JP2 file format supports a restricted set of ICC profiles, namely gamma-matrix-style ICC profiles. This style of profile can represent the data encoded by RGB color spaces other than sRGB. The image data is still RGB; it's just that it is specified indirectly by means of an ICC profile. This does not necessarily mean that non-sRGB systems must support ICC workflows; it does mean more sophisticated handling of the color specification in the JP2 file. For example, the system may recognize that the JP2 file contains the ICC profile for Adobe RGB 98 and use an Adobe RGB 98 workflow.

An alternative to JP2 is the Baseline JPX file format, defined in Part 2 of the JPEG 2000 standard. JPX is an extended version of JP2 which, among other things, specifies additional named color spaces, including Adobe RGB 98 and ProPhoto RGB. There are some RGB spaces, such as eciRGBv2, which JPX does not support directly and for which ICC profiles would still be needed for them to be used. The best thing is to use the JP2 file format as long as possible, since it is more widely supported than JPX and its use avoids support for the more advanced features of JPX when only extended color space support is desired.

## 4.2 Capture Resolution

The JP2 Header Box may also contain a capture or display resolutions, indicating the resolution at which the image was captured or the resolution at which it should be displayed. While the JP2 file is required to contain a color specification, it is not required to have either resolution values. Instead, it is up to the application to require it. This report recommends that the JP2 Header Box in the JP2 file contain a capture resolution value, indicating the resolution at which the image contained in the file was scanned. The JP2 file format specification requires that the resolution value be given in pixels per meter.

## 4.3 Metadata

In addition to the four boxes that a JP2 is required to contain, it may optionally contain XML and UUID boxes. Each can contain vendor or application specific information, encoded in an XML box using XML or in a UUID box in a way that is interpreted according to the UUID code (UUID stands for Universally Unique Identifier). These two types of boxes are used to embed metadata in a JP2 file. For example, UUID boxes are used for IPTC<sup>4</sup> or EXIF<sup>5</sup> metadata. An XML box can be used for any XML-encoded data, such as MIX.

While the application and system normally determine the nature and format of the metadata associated with an image, JPEG 2000-specific administrative or technical metadata is within scope for this report. While such metadata may or may not be embedded in a JP2 file, this reports recommends that it be embedded.

JPEG 2000-specific metadata in the JP2 file should follow the ANSI/NISO Z39.87-2006 standard. This standard defines a data dictionary with technical metadata for digital still images. It lists "image/jp2" as an example of a *formatName* value and lists "JPEG2000 Lossy" and "JPEG2000 Lossless" as *compressionScheme* values. Files that implement this recommendation would have "JPEG 2000 Lossy" as their *compressionScheme* value and would also contain a rational *compressionRatio* value.

Compression	compressionScheme	JPEG2000 Lossy
	compressionRatio	<rational value>

While "JPEG2000 Lossy" is the *compressionScheme* value for all files that follow this recommendation and the *compressionRatio* value can be derived from file size and parameters in the JP2 Header Box, it is recommended that they be specified explicitly.

The Z39.87 standard also defines a *SpecialFormatCharacteristics* container to document attributes that are unique to a particular file format and datastream. In the case of JPEG 2000, this container has two sub-containers: one for *CodecCompliance* and the other for *EncodingOptions*. The elements in the *CodecCompliance* container identify by name and version the coder that created the datastream, the profile to which the datastream conforms (Part 1 of the JPEG 2000 standard defines codestream or datastream profiles), and the class of the decoder needed to decompress the image (Part 4 of the JPEG 2000 defines compliance classes). The elements in the *EncodingOptions* container give the size of the tiles, the number of quality layers and the number of resolution levels.

---

<sup>4</sup> International Press Telecommunications Council, creates standards for photo metadata (<http://www.iptc.org/IPTC4XMP/>)

<sup>5</sup> Exchangeable image file format, a standard file format with metadata tags for digital cameras (<http://www.exif.org/>)



The following table shows the hierarchy of *SpecialFormatCharacteristics* containers and elements for JPEG 2000; the column on the right shows the values these elements would have for the data stream generated by the example in Section 3.7.

JPEG2000	CodecCompliance	codec	Kakadu
		codecVersion	6.0
		codestreamProfile	1
		complianceClass	2
	EncodingOptions	tiles	1024x1024
		qualityLayers	12
resolutionLevels		5	

An XML schema for these and the other elements defined in the Z39.87 standard is available at <http://www.loc.gov/standards/mix>.

When metadata is embedded in a JP2 file, it would be convenient if it were near the beginning of the file where it could be found and read quickly. Any XML or UUID boxes containing metadata can immediately follow the JPEG 2000 Signature and FileType boxes, which must be the first two boxes in a JP2 file. This means the metadata can come before the JP2 Header box, which in turn must come before the Contiguous Codestream box. Therefore, the metadata-containing boxes can occur in a JP2 file before any of the image data to which their contents pertain.

### 4.3 Support

JPEG 2000 is supported by several popular image editors, toolkits and viewers. Among them are Adobe Photoshop, Corel Paint Shop Pro, Irfanview, ER Viewer, Apple QuickTime and SDKs from Lead Technologies and Accusoft Pegasus. Aside from the viewers, all offer automated command lines and batch support.

Other sources of JPEG 2000 components and libraries are Kakadu, Luratech<sup>6</sup> and Aware<sup>7</sup>. This is not an exhaustive list. JP2 is not natively supported by Web browsers. In this regard, it is like PDF and TIFF and, like them, Web browser plug-ins are available, such as from Luratech and LizardTech<sup>8</sup>.

A common approach for delivering online images from a JPEG 2000 server is to decode just as much of the JPEG 2000 image as is needed in terms of resolution, quality and position to create the requested view, and then convert the resulting image to JPEG at the server for delivery to a client browser. This avoids the need for a client side plug-in to view JPEG 2000. The National Digital Newspaper Program (NDNP) uses this approach; it offers 1.25 million newspapers pages, all stored as JPEG 2000, on its website at <http://chroniclingamerica.loc.gov/>. While NDNP uses a commercial JPEG 2000 server from Aware, other commercial servers as well as the Djakota<sup>9</sup> open-source JPEG 2000 image server are also available.

The choice between a JPEG 2000 client or a JPEG client with server-side JPEG2000-to-JPEG conversion is a system issue that is largely independent of the JPEG 2000 datastream and file format recommendation in this report.

<sup>6</sup> <http://www.luratech.com/>

<sup>7</sup> <http://www.aware.com/imaging/jpeg2000.htm>

<sup>8</sup> [http://www.lizardtech.com/download/dl\\_options.php?page=plugins](http://www.lizardtech.com/download/dl_options.php?page=plugins)

<sup>9</sup> <http://www.dlib.org/dlib/september08/chute/09chute.html>

## 5 Conclusion

This report has described the use of JPEG 2000 as a preservation and access format for materials in the Wellcome Digital Library. In general terms, the compression ratio is set for preservation and quality, and JPEG 2000 datastream parameters such as the number of resolution levels and quality layers and tile size are set for access and performance.

This report recommends that the preservation format for single grayscale and color images be a JP2 file containing a minimally lossy irreversible JPEG 2000 datastream, typically with five resolution levels and multiple quality layers. To improve performance, especially decompression times on access, the datastream would be generated with tiles and in coder bypass mode.

One consequence of using a minimally lossy JPEG 2000 datastream is that the compressed file size will depend on image content, which will create some uncertainty in the overall storage requirements. The ability with JPEG 2000 to create a compressed image with a specified size will reduce this uncertainty, but replace it with some variability in the quality of the compressed images. The sample images could tolerate more than minimally lossy compression; how much more depends on the quality requirements of the Wellcome Library, which will depend on image content. Until these requirements are articulated and validated, and even after they are, using quality layers in the datastream will provide a range of compressed file sizes to satisfy future image quality-file size tradeoffs.

This report recommends that the access format be either the same as the preservation format or a subset of it obtained by discarding quality layers to create a smaller and more compressed file. Requests for a view of a particular portion of an image at a particular size would be satisfied in a just-in-time on-demand fashion by accessing only as many of the tiles (and codeblocks within a tile), resolution levels and quality layers as are needed to obtain the image data for the view.

## Appendix 1: JPEG 2000 Datastream Parameters

This table specified the values for the main parameters of the JPEG 2000 datastream.

JPEG 2000 Datastream Parameters	
Parameter	Value
SI Z marker segment	
Profile	Rsiz=2 (Profile 1)
Image size	Same as scanned original
Tiles	1024 x 1024
Image and tile origin	XOsiz = YOsiz = XTOSiz = YTOsiz = 0
Number of components	Csiz = 1 (grayscale) or 3 (color)
Bit depth	Determined by scan
Subsampling	XRsiz = YRsiz = 1
Marker Locations	
COD, COC, QCD, QCC	Main header only
COD/COC marker segments	
Progression Order	RPCL
Number of decomposition levels	$N_L = 5$
Number of layers	Multiple (see text)
Code-block size	xcb=ycb=6
Code-block style	SPcod, SPcoc = 0000 0001 (Coding Bypass)
Transformation	9-7 irreversible filter
Precinct size	Not explicitly specified

## Appendix 2: Reversible and Irreversible Compression

This appendix describes the operation of a JPEG 2000 coder and the differences between reversible and irreversible compression. The first thing a JPEG 2000 coder does with an RGB color image is to apply a component transform which converts it to something better suited to compression by reducing the redundancy between the red, green and blue components in a color image. The next thing it does is apply multiple wavelet transforms, corresponding to the multiple resolution levels, again with the idea of making the image more suitable for compression by redistributing the energy in the image over different subbands or subimages. The step after that is an optional quantization, where image information is discarded on the premise that it will hardly be missed (if it's not overdone) and will further condition the image for compression. The next step is a coder, which takes advantage of all the prep work that has gone on before and the resulting statistics of the transformed and quantized signal to use fewer bits to represent it; this is where the compression actually occurs. The coder doesn't discard any data and is reversible, although the steps leading up to it may not be. In a JPEG 2000 coder, there is one more step in which the compressed data is organized to define the quality layers boundaries and to support the resolution-major progressive order mentioned earlier.

For JPEG 2000 compression to be reversible, there can't be any quantization or round-off errors in the component and wavelet transforms. To avoid round-off errors, JPEG 2000 has a reversible transform based on integer arithmetic. So for example, JPEG 2000 specifies a reversible wavelet transform, called the 5-3 transform because of the size of the filters it uses. JPEG 2000 also specifies an irreversible wavelet transform, the 9-7 transform, based on floating point operations. The 9-7 transform does a better job than the 5-3 transform of conditioning the image data and so gives better compression, but at the cost of being unable to recover the original data due to round-off errors in its calculations. Because of these round-off errors, the 9-7 transform is not reversible.

The following Kakadu commands were used to generate the compressed images for the tests reported in Section 3.6. The first command generates a reversible JPEG 2000 datastream; the second, an irreversible but minimally lossy datastream; and the third, an irreversible datastream using coder bypass.

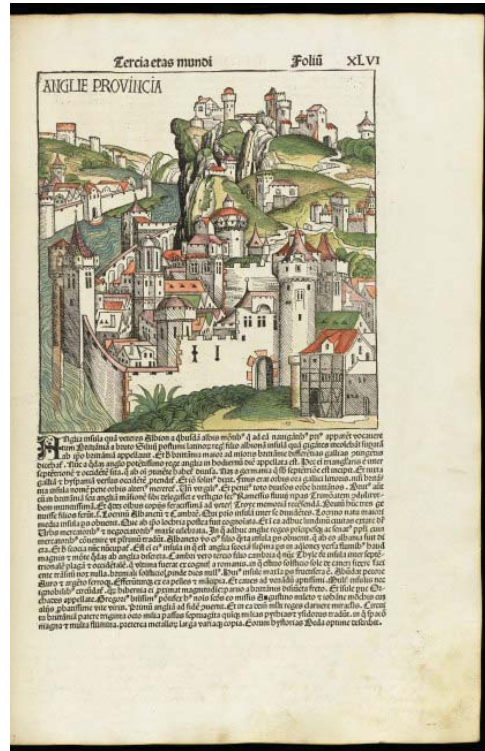
```
kdu_compress -i in.tif -o out.jp2 -rate - Creversible=yes
Clevels=5 Stiles={ 1024,1024 Cblk={ 64,64 } } Corder=RPCL
//reversible

kdu_compress -i in.tif -o out.jp2 -rate - Creversible=no
Clevels=5 Stiles={ 1024,1024 Cblk={ 64,64 } } Corder=RPCL
//irreversible

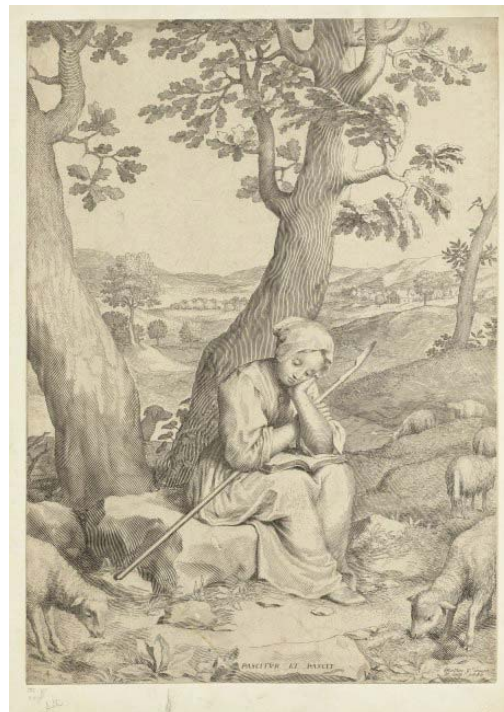
kdu_compress -i in.tif -o out.jp2 -rate - Creversible=no
Clevels=5 Stiles={ 1024,1024 Cblk={ 64,64 } } Corder=RPCL
Cmodes=BYPASS
```

The dash after the rate parameter in these commands indicates that all transformed and quantized data is to be retained and none discarded. Creversible=yes in the first command directs the coder to use the reversible wavelet and components transforms. Creversible=no in the second and third commands directs the use of irreversible transforms. Cmodes=BYPASS in the third command directs the coder to use Bypass mode.

Figure 1: Sample images



L0051262\_Manuscript\_Page

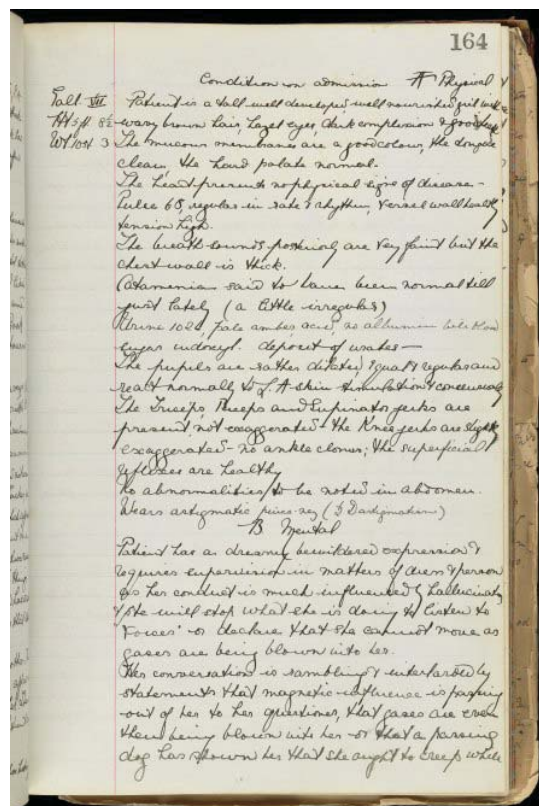




L0051320\_Line\_Drawing

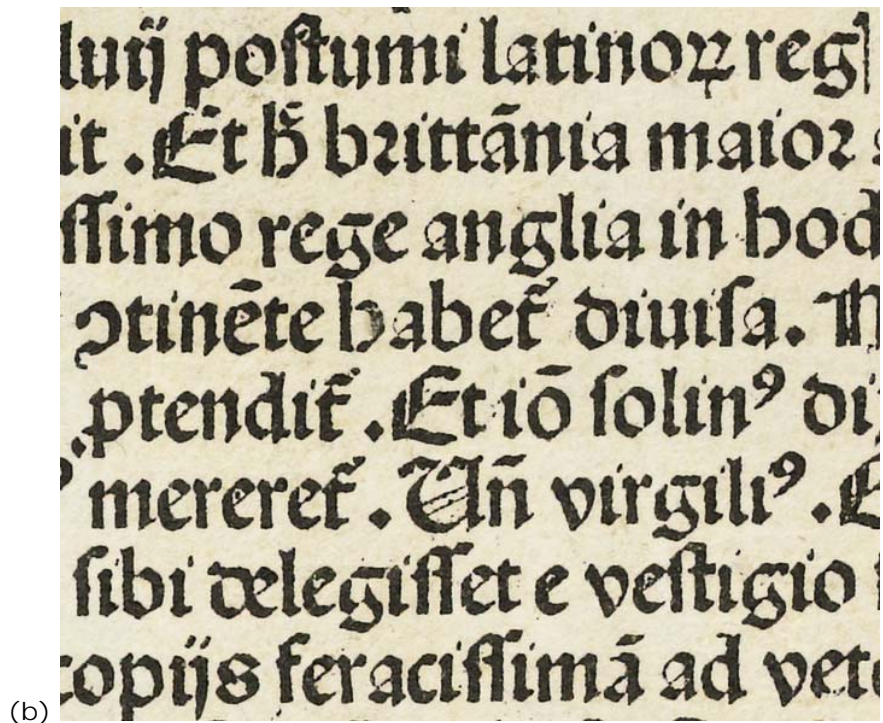
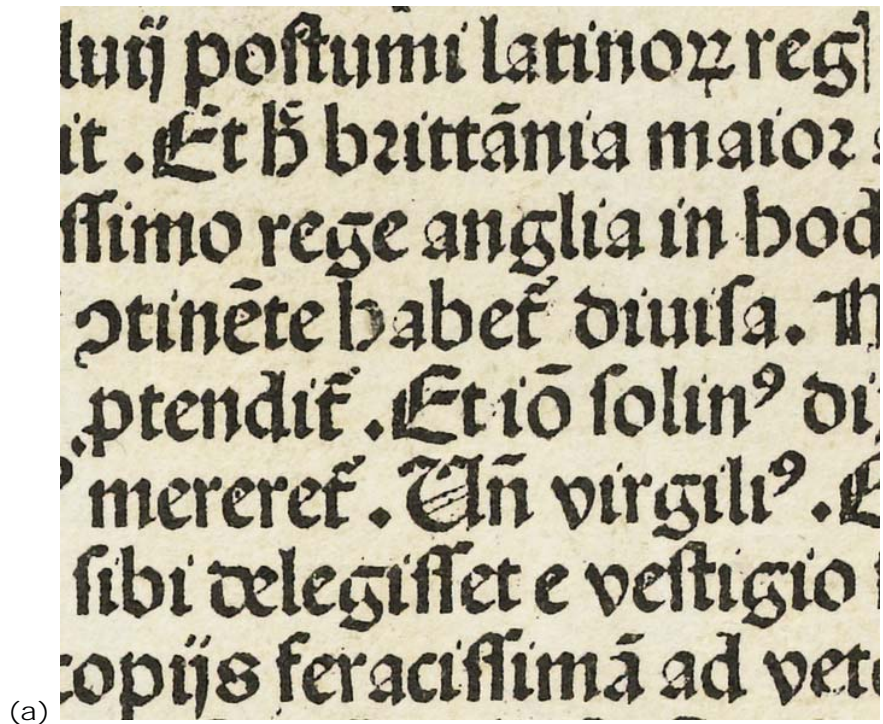


L0051761\_Painting



L0051440\_Archive\_Collection

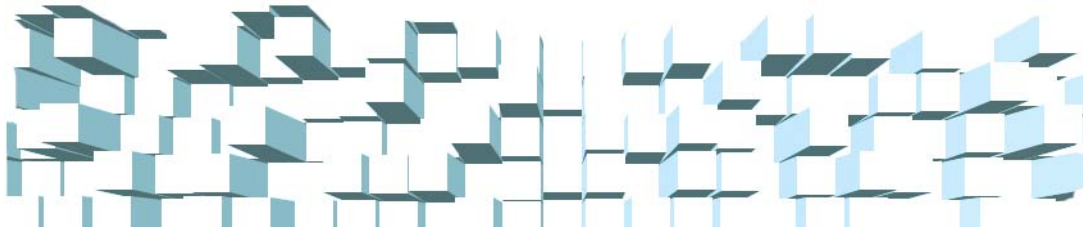
Figure 2: Comparison of irreversible with reversible compression



Comparison of (a) irreversible compression with minimal loss and (b) reversible compression of region from L0051262\_Manuscript\_Page image, reproduced at 150 dpi.



**JPEG 2000 as a  
Preservation and Access Format  
for the Wellcome Trust Digital Library**



**Robert Buckley**  
Xerox Corporation

**Edited by Simon Tanner**  
*King's Digital Consultancy Services*  
*King's College London*



[www.kdcs.kcl.ac.uk](http://www.kdcs.kcl.ac.uk)



## Contents

<b>1. SUMMARY OF ISSUES/QUESTIONS .....</b>	<b>3</b>
<b>2. RECOMMENDATIONS.....</b>	<b>4</b>
<b>3 BASIS OF RECOMMENDATIONS/REASONINGS/TESTS DONE.....</b>	<b>5</b>
3.1 COMPRESSION .....	5
3.2 MULTIPLE RESOLUTION LEVELS .....	5
3.3 MULTIPLE QUALITY LAYERS.....	6
3.4 EXAMPLE: TIFF TO JP2 CONVERSION .....	6
3.5 MINIMALLY LOSSY COMPRESSION .....	7
3.6 TESTING REVERSIBLE AND IRREVERSIBLE COMPRESSION.....	7
3.7 FURTHER COMPRESSION FINDINGS.....	8
<b>4 IMPLEMENTATION SOLUTIONS / DISCUSSION.....</b>	<b>9</b>
4.1 COLOR SPECIFICATION .....	9
4.2 CAPTURE RESOLUTION.....	10
4.3 METADATA .....	10
4.3 SUPPORT .....	11
<b>5 CONCLUSION .....</b>	<b>12</b>
<b>APPENDIX 1: JPEG 2000 DATASTREAM PARAMETERS .....</b>	<b>13</b>
<b>APPENDIX 2: REVERSIBLE AND IRREVERSIBLE COMPRESSION .....</b>	<b>14</b>
<b>FIGURE 1: SAMPLE IMAGES.....</b>	<b>15</b>
<b>FIGURE 2: COMPARISON OF IRREVERSIBLE WITH REVERSIBLE COMPRESSION.....</b>	<b>17</b>

## 1. Summary of Issues/questions

The Wellcome Trust is developing a digital library over the next 5 years, anticipating a storage requirement for up to 30 million images. The Wellcome previously has used uncompressed TIFF image files as their archival storage image format. However, the storage requirement for many millions of images suggests that a better compromise is needed between the costs of secure long-term digital storage and the image standards used. It is expected that by using JPEG2000, total storage requirements will be kept at a value that represents an acceptable compromise between economic storage and image quality. Ideally, JPEG2000 could serve as both a preservation format and as an access or production format in a write-once-read-many type environment.

JPEG2000 was chosen as an image preservation format due to its small size and because it offers intelligent compression for preservation and intelligent decompression for access. If a lossy format is used to obtain a relatively high compression, e.g. between 5:1 and 20:1 (in comparison to an uncompressed TIFF file), then the storage requirements desired are achievable. The questions to address are what level of compression is acceptable and delivers the desired balance of image quality and reduced storage footprint.

With regard to the use of JPEG 2000, the questions posed in the brief and addressed in this report are:

- a. What JPEG2000 format(s) is best suited for preservation?
- b. What JPEG2000 format(s) is best suited for access?
- c. Can any single JPEG2000 format adequately serve both preservation and access?
- d. What models exist for the use of descriptive and/or administrative metadata with JPEG2000?
- e. If a JPEG2000 format is recommended for access purposes, what tools can be used to display/manipulate/manage it and any associated or embedded metadata?

This report will describe how a unified approach can enable JPEG2000 to serve for both preservation and access and balance the needs for compressed image size, image quality and decompression performance.

## 2. Recommendations

The majority of materials that will form part of the Wellcome Digital Library are expected to use visually lossless JPEG 2000 compression. Although “visually lossless” compression is lossy, the differences it introduces between the original and the image reconstructed from a compressed version of it are either not noticeable or insignificant and do not interfere with the application and usefulness of the image. Because the original cannot be reconstructed from the compressed image, compression in this case is irreversible. JPEG compression in digital still cameras is a familiar example of irreversible but visually lossless image compression. In mass digitization projects that use JPEG2000, compression ratios around 40:1 have been used for basically textual content. When applied to printed books, it has been found that these compression ratios do not impair the legibility or OCR accuracy of the text.

However, archiving and long-term preservation indicate a more conservative approach to compression and a different trade-off between compressed image size and image quality to meet current and anticipated uses. Still, given the volumes of material being digitized, a lossy format represents an acceptable compromise between quality and economic storage. A compression ratio of 4:1 or 5:1 gives a conservative upper limit on file size and decompressed image quality in the preservation format for the material being digitized. However this material should tolerate higher compression ratios with the results remaining visually lossless.

While most of the materials will use visually lossless compression, it is suggested that a small subset of materials (less than 5% of total) may be candidates for lossless or reversible compression. Reversible means that the original can be reconstructed exactly from the compressed image, i.e. the compression process is reversible.

Nevertheless, this report recommends irreversible JPEG2000 compression for the preservation and access formats of single grayscale or color images. Initially specifying a minimally lossy datastream will result in overall compression ratios around 4:1; the exact value will depend on image content. While this is a particularly conservative compression ratio, the compression can be increased as new materials are captured and even applied retroactively to files with previously captured material. The access format will be a subset of the preservation format with a subset of the resolution levels and quality layers in the preservation format.

In particular, the JPEG 2000 datastream should have the following properties:

- Irreversible compression using the 9-7 wavelet transform and ICT (see Section 3.1) with minimal loss (see Section 3.5)
- Multiple resolutions levels: the number depends on the original image size and the desired size of the smallest image derived from the JPEG 2000 datastream (see Section 3.2)
- Multiple quality layers, where all layers gives minimally lossy compression for preservation (see Sections 3.3 and 3.5)
- Resolution-major progression order (see Section 3.4)
- Tiles for improved codec (coder-decoder) performance, although the final decision regarding the use of tiles and precincts depends on the codec
- Generated using Bypass mode, which creates a compressed datastream that takes less time to compress and decompress (see Section 3.6)
- TLM markers (see Section 3.7)

A formal specification of the JPEG 2000 datastream for this application is given in Appendix 1. The datastream specified there is compatible with Part 1 of the

JPEG 2000 standard; none of the JPEG 2000 datastream extensions defined in Part 2 of the standard are needed.

Further, this report recommends embedding the JPEG 2000 datastream in a JP2 file. The JP2 file should contain:

- A single datastream containing a grayscale or color image whose content can be specified using the sRGB color space (or its grayscale or luminance-chrominance analogue) or a restricted ICC<sup>1</sup> profile, as defined in the JP2 file format specification in Part 1 of the JPEG 2000 standard (see Section 4.1)
- A Capture Resolution value (see Section 4.2)
- Embedded metadata that describes the JPEG 2000 datastream should follow the ANSI/NISO Z39.87-2006 standard and be placed in a XML box following the FileType box in a JP2 file (see Section 4.3)

Using the JP2 file format is sufficient as long as the requirement is for a single datastream whose color content can be specified using sRGB or a restricted ICC profile. While the JPX file format can be used if the color content of the image is specified by a non-sRGB color space or a general ICC profile, the use of a JP2-compatible file format is recommended.

## **3 Basis of recommendations/reasonings/tests done**

### **3.1 Compression**

In general terms, the compression ratio is set for preservation and quality, and JPEG 2000 datastream parameters such as the number of resolution levels and quality layers and tile size are set for access and performance. JPEG 2000 offers smart decompression, where only that portion of the datastream needed to satisfy the requested image view in terms of resolution, quality and location need be accessed and decompressed on demand and just in time.

The JPEG 2000 compression offers both reversible and irreversible compression. Reversible compression in JPEG 2000 uses the 5-3 integer wavelet transform and a reversible component transform (RCT). If no compressed data is discarded, then the original image data is recoverable from the compressed datastream created using these transforms. Irreversible compression uses the 9-7 floating point wavelet transform and an irreversible component transform (ICT), both of which have round-off errors so that the original image data is not recoverable from the compressed datastream, even when no compressed data is discarded. Appendix 2 contains a more detailed discussion of the differences between reversible and irreversible compression in JPEG 2000.

### **3.2 Multiple resolution levels**

To begin with, it is recommended that JPEG 2000 be used with multiple resolution levels. The first two or three resolution levels facilitate compression; levels beyond that give little more compression but are added so that decompressing just the lowest resolution sub-image in the JPEG 2000 datastream gives a thumbnail of a desired size. For example, with a 5928-by-4872 pixel dimension original and 5 resolution levels, the smallest sub-image would have dimensions that would be 1/32 those of the original, in this case 186 by 153 pixels, which is roughly QQVGA<sup>2</sup> sized. Accordingly, JPEG 2000

---

<sup>1</sup> International Color Consortium, an organization that develops and promotes color management using the ICC profile format ([www.color.org](http://www.color.org))

<sup>2</sup> Quarter-quarter VGA (Video Graphics Array); since VGA is 640 by 480, QQVGA is 160 by 120.

compression with 5 resolution levels is recommended for images of this and similar sizes, which are typical of the sample images provided. In practice, the number of resolution levels would vary with the original image size so that the lowest resolution sub-image has the desired dimensions.

### 3.3 *Multiple quality layers*

There are two main reasons for using multiple quality layers. One is so that it is possible to decompress fewer layers and therefore less compressed data when accessing lower resolution sub-images. This speeds up decompression without affecting quality since the incremental quality due to the discarded layers is not noticeable at reduced resolutions. The second reason is that multiple quality layers make it possible to deliver subsets of the compressed image corresponding to higher compression ratios, which may be acceptable in some applications. This means there is less data to transmit and process, which improves performance and reduces access times. It also means that it is possible for the access format to be a subset of the preservation format, derived from it by discarding quality layers as the application and quality requirements warrant. The use of quality layers makes it possible to retroactively reduce the storage needs should they be revised downward by discarding quality layers in the preservation format and turning images compressed at 4:1 or 5:1 for example into images compressed at 8:1 or higher, depending on where the quality layer boundaries are defined.

### 3.4 *Example: TIFF to JP2 conversion*

For example, the following command line uses the Kakadu<sup>3</sup> compress function (`kdu_compress`) to convert a TIFF image to a JP2 file that contains an irreversible JPEG 2000 datastream. In particular, it contains a lossy JPEG 2000 datastream with 5 resolution levels and 8 quality layers, corresponding to compression ratios of 4, 8, 16, 32, 64, 128, 256 and 512 to 1 for a 24-bit color image. These correspond to compressed bit rates of 6, 3, 1.5, 0.75, 0.375, 0.1875, 0.09375 and 0.046875 bits per pixel. (A compression ratio of 4 to 1 applied to a color image that originally had 24 bits per pixel means the compressed image will equivalently have a compressed bit rate of 6 bits per pixel.) The Kakadu command line use bit rates rather than compression ratios to specify the amount of compression.

```
kdu_compress -i in.tif -o out.jp2 -rate  
6,3,1.5,0.75,0.375,0.1875,0.09375,0.046875 Creversible=no  
Clevels=5 Stiles={ 1024,1024} Cblk={ 64,64} Corder=RPCL
```

The JPEG 2000 datastream created in this example has 1024-by-1024 tiles, 64-by-64 codeblocks and a resolution-major progressive order RPCL, so that the compressed data for the lowest resolution (and therefore smallest) sub-image occurs first in the datastream, followed by the compressed data needed to reconstruct the next lowest resolution sub-image and so on. This data ordering means that the data for a thumbnail image occurs in a contiguous block at the start of the datastream where it can be easily and speedily accessed. This data organization makes it possible to obtain a screen-resolution image quickly from a megabyte or gigabyte sized image compressed using JPEG 2000. Tiles and codeblocks are used to partition the image for processing and make it possible to access portions of the datastream corresponding to sub-regions of the image.

---

<sup>3</sup> <http://www.kakadusoftware.com/>

### 3.5 *Minimally Lossy Compression*

The JPEG 2000 coder in this example would discard transformed and compressed data to obtain a compressed file size corresponding to 4:1 compression. This needs to be compared with the performance of the the minimally lossy coder, where no data is discarded but which is still lossy because of the use of the irreversible transforms. In some cases, depending on the image content, as shown in Section 3.6, the minimally lossy coder can give higher compression ratios than 4:1. Accordingly, it is recommended that a minimally lossy format with multiple quality layers and multiple resolution levels be used for the preservation format. The access format would use reduced quality subsets of the preservation format optionally obtained by discarding layers and using reduced resolution levels.

### 3.6 *Testing reversible and irreversible compression*

The reason to use irreversible compression is that it gives better compression than reversible compression, at the cost of introducing errors (or differences) in the reconstructed image. This section examines this performance tradeoff.

Reversible and irreversible compression were applied to four images provided by the Wellcome Digital Library (Figure 1). A variation on irreversible compression was tested which used coder bypass mode, in which the coder skipped the compression of some of the data. This gave a little less compression, but made the coder (and decoder) run about 20% faster. The Kakadu commands used in these tests are given in Appendix 2.

The compression ratios obtained with these three test are shown in the following table.

Original	Reversible	Irreversible	Irreversible w/bypass
L0051262_Manuscript_Page	2.25	3.45	3.42
L0051320_Line_Drawing	1.82	2.52	2.51
L0051761_Painting	2.46	3.96	3.90
L0051440_Archive_Collection	2.52	4.47	4.41

For these particular images, the compression ratio for irreversible JPEG 2000 was from about 40% to almost 80% better than it was for reversible, and on average over 30% faster (with a further 20% boost with coder bypass mode).

The cost of irreversible compared to reversible is the error it introduces. The error or difference between the reversibly and irreversibly compressed images is about 50 dB, which means the average absolute error value was about 0.5. For one of the sample images, 99.99% of the green component values were the same after decompression as they were before, or at most two counts different. (For the red and blue components, the percentages were 99.79 and 99.35.) This is within the tolerance for scanners: in other words, minimally lossless irreversible JPEG 2000 compression adds about as much noise to an image as a good scanner does.

A region was cropped from one image so that the visual effects of this error on this image could be examined more closely (Figure 2). When they were, the differences were not perceptible on screen or on paper. Unless being able to reconstruct the original scan is a requirement, legal or otherwise, then irreversible compression is clearly advantaged over reversible compression.



### 3.7 Further compression findings

In these tests, the compressed file sizes (and compression ratios) were image dependent and varied with the image content. Images with less detail or variation than these samples would give even higher compression ratios.

An advantage of JPEG 2000 is that it lets the user set the compression ratio, or equivalently the compressed file size, to a specific target value, which the coder achieves by discarding compressed image data. While this feature was not used to set the overall compressed file size in the minimally lossy compression case, it can be used to set the sizes of intermediate images corresponding to the different quality layers. The following Kakadu command line generates a JP2 file with a minimally lossy irreversible JPEG 2000 datastream that complies with the recommendation in this report:

```
kdu_compress -i in.tif -o out.jp2 -rate -, 4, 2.34, 1.36, 0.797,
0.466, 0.272, 0.159, 0.0929, 0.0543, 0.0317, 0.0185
Creversible=no Clevels=5 Stiles={ 1024,1024} Cblk={64,64}
Corder=RPCL Cmodes=BYPASS
```

The JPEG 2000 datastream in this example has 5 resolution levels and 12 quality layers. Using all 12 layers give a decompressed image with minimal loss. The intermediate layers boundaries are at pre-set compressed bit rates, starting at 4 bits per pixel, corresponding to a compression ratio of 6:1, assuming a 24-bit color original. Thereafter, the layer boundaries are distributed logarithmically up to a compression ratio of 1296:1. The exact values are not critical. What is important is the range of values and there being sufficient values to provide an adequate sampling within the range.

When a datastream has multiple quality layers, it is possible to truncate it at points corresponding to the layer boundaries and obtain derivative datastreams that correspond to higher compression ratios (or lower compressed bit rates). In the previous example, discarding the topmost quality layer produces a datastream corresponding to a compression ratio of 6:1 (compressed bit rate of 4 bits per pixel). Discarding the next layers produces a datastream with a compression ratio of 10.3:1, and so on. As noted previously, some images may have minimally lossy compression ratio greater than 6:1; the layer settings can be adjusted when this happens.

Using layers adds overhead that increases the size of the datastream and therefore decreases the compression ratio. To assess this effect as well as the overhead due to the use of tiles, the four sample images were compressed with one layer and no tiles, with one layer and 1024x1024 tiles, and with 12 layers and no tiles. As the following table shows, adding layers and tiles did decrease the minimally lossy compression ratio, but the effect was only visible in the third place after the decimal and was therefore judged insignificant in comparison to the advantages of using them.

Original	No tiles 1 layer	1024x1024 tiles 1 layer	No tiles 12 layers
L0051262_Manuscript_Page	3.452	3.450	3.443
L0051320_Line_Drawing	2.522	2.521	2.517
L0051761_Painting	3.961	3.957	3.948
L0051440_Archive_Collection	4.477	4.473	4.461

Besides tiles and layers, other datastream components that can improve performance and access within the datastream are markers, such as Tile Length Markers (TLM) which can aid in searching for tile boundaries in a datastream. Their effectiveness depends on whether or not the decoder or

access protocol makes use of them. As a result, recommendations regarding their use depend on the choice of codec.

## 4 Implementation solutions / discussion

This section discusses the file format and metadata recommendations.

One function of a file format is packaging the datastream with metadata that can be used to render, interpret and describe the image in the file. Besides defining the JPEG 2000 datastream and core decoder, Part 1 of the JPEG 2000 standard also defines the JP2 file format which applications may use to encapsulate a JPEG 2000 datastream. A minimal JP2 file consists of four structures or "boxes":

1. JPEG 2000 Signature Box, which identifies the file as a member of the JPEG 2000 file format family
2. File Type Box, which identifies which member of the family it is, the version number and the members of the family it is compatible with
3. JP2 Header Box, which contains image parameters such as resolution and color specification needed for rendering the image
4. Contiguous Codestream Box, which contains the JPEG 2000 datastream

### 4.1 Color Specification

How an image was captured or created determines the parameters in the JP2 Header Box, which are subsequently used to render and interpret the image. Among these parameters are the number of components (i.e. whether the image is grayscale or color), an optional resolution value for capture or display, and the color specification. In general the color content of an image can be specified in one of two ways: directly using a named color space, such as sRGB, Adobe RGB 98 or CIELAB, or indirectly using an ICC profile.

The digitization process and the nature of the material being digitized, not the file format, drive the color specification requirements of the application. The issue for the file format is whether or not it supports the color encoding used by the digital materials. What's significant about the JP2 file format is that it supports a limited set of color specifications. For example, the only color space it supports directly is sRGB, including its grayscale and luminance-chrominance analogues. This is a consequence of the JP2 file format having been originally designed with digital cameras in mind.

Besides sRGB, the JP2 file format supports a restricted set of ICC profiles, namely gamma-matrix-style ICC profiles. This style of profile can represent the data encoded by RGB color spaces other than sRGB. The image data is still RGB; it's just that it is specified indirectly by means of an ICC profile. This does not necessarily mean that non-sRGB systems must support ICC workflows; it does mean more sophisticated handling of the color specification in the JP2 file. For example, the system may recognize that the JP2 file contains the ICC profile for Adobe RGB 98 and use an Adobe RGB 98 workflow.

An alternative to JP2 is the Baseline JPX file format, defined in Part 2 of the JPEG 2000 standard. JPX is an extended version of JP2 which, among other things, specifies additional named color spaces, including Adobe RGB 98 and ProPhoto RGB. There are some RGB spaces, such as eciRGBv2, which JPX does not support directly and for which ICC profiles would still be needed for them to be used. The best thing is to use the JP2 file format as long as possible, since it is more widely supported than JPX and its use avoids support for the more advanced features of JPX when only extended color space support is desired.



## 4.2 Capture Resolution

The JP2 Header Box may also contain a capture or display resolutions, indicating the resolution at which the image was captured or the resolution at which it should be displayed. While the JP2 file is required to contain a color specification, it is not required to have either resolution values. Instead, it is up to the application to require it. This report recommends that the JP2 Header Box in the JP2 file contain a capture resolution value, indicating the resolution at which the image contained in the file was scanned. The JP2 file format specification requires that the resolution value be given in pixels per meter.

## 4.3 Metadata

In addition to the four boxes that a JP2 is required to contain, it may optionally contain XML and UUID boxes. Each can contain vendor or application specific information, encoded in an XML box using XML or in a UUID box in a way that is interpreted according to the UUID code (UUID stands for Universally Unique Identifier). These two types of boxes are used to embed metadata in a JP2 file. For example, UUID boxes are used for IPTC<sup>4</sup> or EXIF<sup>5</sup> metadata. An XML box can be used for any XML-encoded data, such as MIX.

While the application and system normally determine the nature and format of the metadata associated with an image, JPEG 2000-specific administrative or technical metadata is within scope for this report. While such metadata may or may not be embedded in a JP2 file, this reports recommends that it be embedded.

JPEG 2000-specific metadata in the JP2 file should follow the ANSI/NISO Z39.87-2006 standard. This standard defines a data dictionary with technical metadata for digital still images. It lists "image/jp2" as an example of a *formatName* value and lists "JPEG2000 Lossy" and "JPEG2000 Lossless" as *compressionScheme* values. Files that implement this recommendation would have "JPEG 2000 Lossy" as their *compressionScheme* value and would also contain a rational *compressionRatio* value.

Compression	compressionScheme	JPEG2000 Lossy
	compressionRatio	<rational value>

While "JPEG2000 Lossy" is the *compressionScheme* value for all files that follow this recommendation and the *compressionRatio* value can be derived from file size and parameters in the JP2 Header Box, it is recommended that they be specified explicitly.

The Z39.87 standard also defines a *SpecialFormatCharacteristics* container to document attributes that are unique to a particular file format and datastream. In the case of JPEG 2000, this container has two sub-containers: one for *CodecCompliance* and the other for *EncodingOptions*. The elements in the *CodecCompliance* container identify by name and version the coder that created the datastream, the profile to which the datastream conforms (Part 1 of the JPEG 2000 standard defines codestream or datastream profiles), and the class of the decoder needed to decompress the image (Part 4 of the JPEG 2000 defines compliance classes). The elements in the *EncodingOptions* container give the size of the tiles, the number of quality layers and the number of resolution levels.

<sup>4</sup> International Press Telecommunications Council, creates standards for photo metadata (<http://www.iptc.org/IPTC4XMP/>)

<sup>5</sup> Exchangeable image file format, a standard file format with metadata tags for digital cameras (<http://www.exif.org/>)

The following table shows the hierarchy of *SpecialFormatCharacteristics* containers and elements for JPEG 2000; the column on the right shows the values these elements would have for the data stream generated by the example in Section 3.7.

JPEG2000	CodecCompliance	codec	Kakadu
		codecVersion	6.0
		codestreamProfile	1
		complianceClass	2
	EncodingOptions	tiles	1024x1024
		qualityLayers	12
		resolutionLevels	5

An XML schema for these and the other elements defined in the Z39.87 standard is available at <http://www.loc.gov/standards/mix>.

When metadata is embedded in a JP2 file, it would be convenient if it were near the beginning of the file where it could be found and read quickly. Any XML or UUID boxes containing metadata can immediately follow the JPEG 2000 Signature and FileType boxes, which must be the first two boxes in a JP2 file. This means the metadata can come before the JP2 Header box, which in turn must come before the Contiguous Codestream box. Therefore, the metadata-containing boxes can occur in a JP2 file before any of the image data to which their contents pertain.

### 4.3 Support

JPEG 2000 is supported by several popular image editors, toolkits and viewers. Among them are Adobe Photoshop, Corel Paint Shop Pro, Irfanview, ER Viewer, Apple QuickTime and SDKs from Lead Technologies and Accusoft Pegasus. Aside from the viewers, all offer automated command lines and batch support.

Other sources of JPEG 2000 components and libraries are Kakadu, Luratech<sup>6</sup> and Aware<sup>7</sup>. This is not an exhaustive list. JP2 is not natively supported by Web browsers. In this regard, it is like PDF and TIFF and, like them, Web browser plug-ins are available, such as from Luratech and LizardTech<sup>8</sup>.

A common approach for delivering online images from a JPEG 2000 server is to decode just as much of the JPEG 2000 image as is needed in terms of resolution, quality and position to create the requested view, and then convert the resulting image to JPEG at the server for delivery to a client browser. This avoids the need for a client side plug-in to view JPEG 2000. The National Digital Newspaper Program (NDNP) uses this approach; it offers 1.25 million newspapers pages, all stored as JPEG 2000, on its website at <http://chroniclingamerica.loc.gov/>. While NDNP uses a commercial JPEG 2000 server from Aware, other commercial servers as well as the Djakota<sup>9</sup> open-source JPEG 2000 image server are also available.

The choice between a JPEG 2000 client or a JPEG client with server-side JPEG2000-to-JPEG conversion is a system issue that is largely independent of the JPEG 2000 datastream and file format recommendation in this report.

<sup>6</sup> <http://www.luratech.com/>

<sup>7</sup> <http://www.aware.com/imaging/jpeg2000.htm>

<sup>8</sup> [http://www.lizardtech.com/download/dl\\_options.php?page=plugins](http://www.lizardtech.com/download/dl_options.php?page=plugins)

<sup>9</sup> <http://www.dlib.org/dlib/september08/chute/09chute.html>

## 5 Conclusion

This report has described the use of JPEG 2000 as a preservation and access format for materials in the Wellcome Digital Library. In general terms, the compression ratio is set for preservation and quality, and JPEG 2000 datastream parameters such as the number of resolution levels and quality layers and tile size are set for access and performance.

This report recommends that the preservation format for single grayscale and color images be a JP2 file containing a minimally lossy irreversible JPEG 2000 datastream, typically with five resolution levels and multiple quality layers. To improve performance, especially decompression times on access, the datastream would be generated with tiles and in coder bypass mode.

One consequence of using a minimally lossy JPEG 2000 datastream is that the compressed file size will depend on image content, which will create some uncertainty in the overall storage requirements. The ability with JPEG 2000 to create a compressed image with a specified size will reduce this uncertainty, but replace it with some variability in the quality of the compressed images. The sample images could tolerate more than minimally lossy compression; how much more depends on the quality requirements of the Wellcome Library, which will depend on image content. Until these requirements are articulated and validated, and even after they are, using quality layers in the datastream will provide a range of compressed file sizes to satisfy future image quality-file size tradeoffs.

This report recommends that the access format be either the same as the preservation format or a subset of it obtained by discarding quality layers to create a smaller and more compressed file. Requests for a view of a particular portion of an image at a particular size would be satisfied in a just-in-time on-demand fashion by accessing only as many of the tiles (and codeblocks within a tile), resolution levels and quality layers as are needed to obtain the image data for the view.

## Appendix 1: JPEG 2000 Datastream Parameters

This table specified the values for the main parameters of the JPEG 2000 datastream.

<b>JPEG 2000 Datastream Parameters</b>	
<b>Parameter</b>	<b>Value</b>
<b>SIZ marker segment</b>	
Profile	Rsiz=2 (Profile 1)
Image size	Same as scanned original
Tiles	1024 x 1024
Image and tile origin	XOsiz = YOsiz = XTOsiz = YTOsiz = 0
Number of components	Csiz = 1 (grayscale) or 3 (color)
Bit depth	Determined by scan
Subsampling	XRsiz = YRsiz = 1
<b>Marker Locations</b>	
COD, COC, QCD, QCC	Main header only
<b>COD/COC marker segments</b>	
Progression Order	RPCL
Number of decomposition levels	$N_L = 5$
Number of layers	Multiple (see text)
Code-block size	xcb=ycb=6
Code-block style	SPcod, SPcoc = 0000 0001 (Coding Bypass)
Transformation	9-7 irreversible filter
Precinct size	Not explicitly specified

## Appendix 2: Reversible and Irreversible Compression

This appendix describes the operation of a JPEG 2000 coder and the differences between reversible and irreversible compression. The first thing a JPEG 2000 coder does with an RGB color image is to apply a component transform which converts it to something better suited to compression by reducing the redundancy between the red, green and blue components in a color image. The next thing it does is apply multiple wavelet transforms, corresponding to the multiple resolution levels, again with the idea of making the image more suitable for compression by redistributing the energy in the image over different subbands or subimages. The step after that is an optional quantization, where image information is discarded on the premise that it will hardly be missed (if it's not overdone) and will further condition the image for compression. The next step is a coder, which takes advantage of all the prep work that has gone on before and the resulting statistics of the transformed and quantized signal to use fewer bits to represent it; this is where the compression actually occurs. The coder doesn't discard any data and is reversible, although the steps leading up to it may not be. In a JPEG 2000 coder, there is one more step in which the compressed data is organized to define the quality layers boundaries and to support the resolution-major progressive order mentioned earlier.

For JPEG 2000 compression to be reversible, there can't be any quantization or round-off errors in the component and wavelet transforms. To avoid round-off errors, JPEG 2000 has a reversible transform based on integer arithmetic. So for example, JPEG 2000 specifies a reversible wavelet transform, called the 5-3 transform because of the size of the filters it uses. JPEG 2000 also specifies an irreversible wavelet transform, the 9-7 transform, based on floating point operations. The 9-7 transform does a better job than the 5-3 transform of conditioning the image data and so gives better compression, but at the cost of being unable to recover the original data due to round-off errors in its calculations. Because of these round-off errors, the 9-7 transform is not reversible.

The following Kakadu commands were used to generate the compressed images for the tests reported in Section 3.6. The first command generates a reversible JPEG 2000 datastream; the second, an irreversible but minimally lossy datastream; and the third, an irreversible datastream using coder bypass.

```
kdu_compress -i in.tif -o out.jp2 -rate - Creversible=yes  
Clevels=5 Stiles={ 1024,1024 Cblk={ 64,64 } } Corder=RPCL  
//reversible
```

```
kdu_compress -i in.tif -o out.jp2 -rate - Creversible=no  
Clevels=5 Stiles={ 1024,1024 Cblk={ 64,64 } } Corder=RPCL  
//irreversible
```

```
kdu_compress -i in.tif -o out.jp2 -rate - Creversible=no  
Clevels=5 Stiles={ 1024,1024 Cblk={ 64,64 } } Corder=RPCL  
Cmodes=BYPASS
```

The dash after the rate parameter in these commands indicates that all transformed and quantized data is to be retained and none discarded. Creversible=yes in the first command directs the coder to use the reversible wavelet and components transforms. Creversible=no in the second and third commands directs the use of irreversible transforms. Cmodes=BYPASS in the third command directs the coder to use Bypass mode.



# Figure 1: Sample images



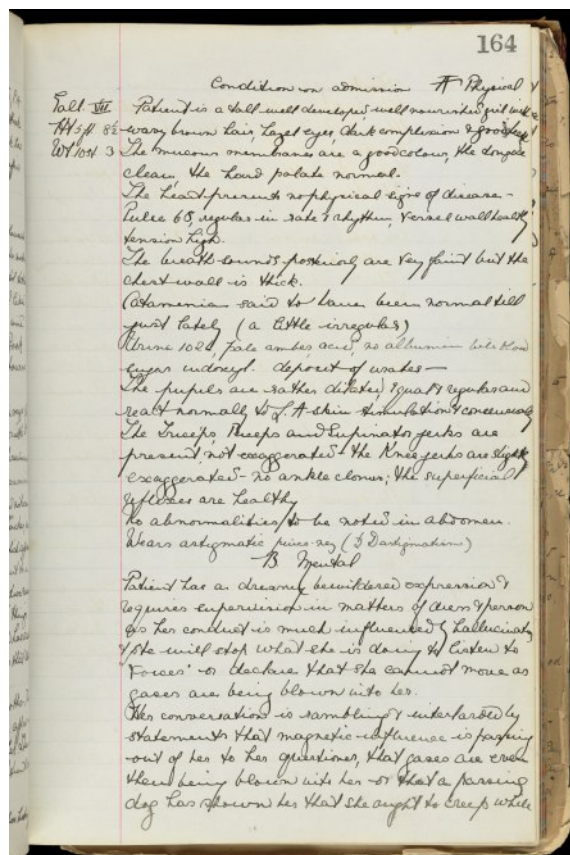
L0051262\_Manuscript\_Page



L0051320\_Line\_Drawing



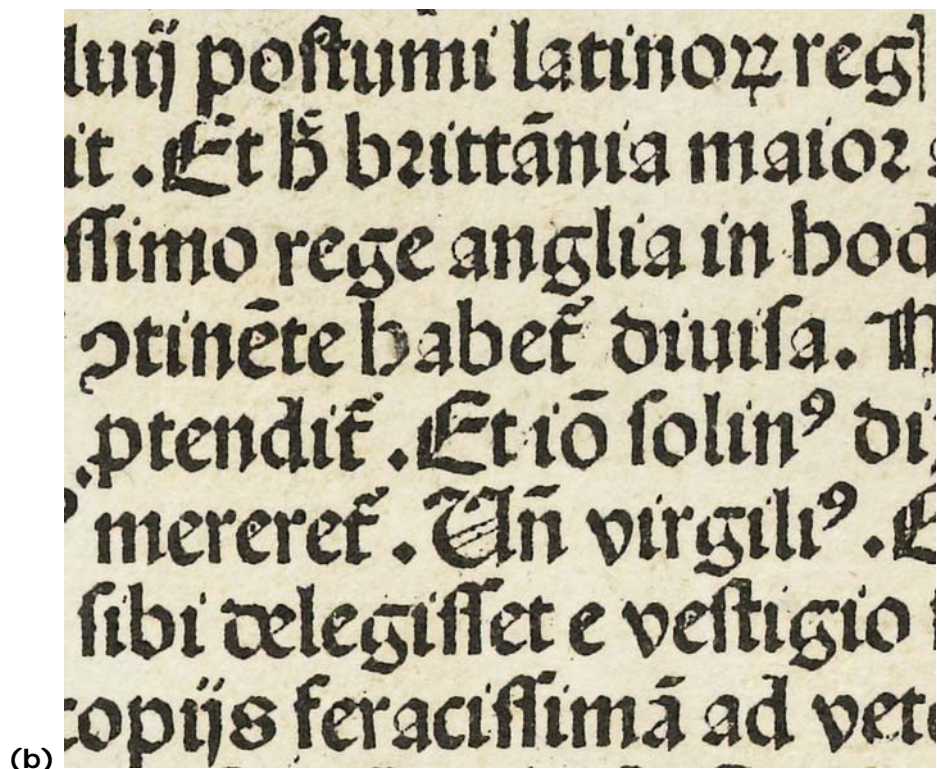
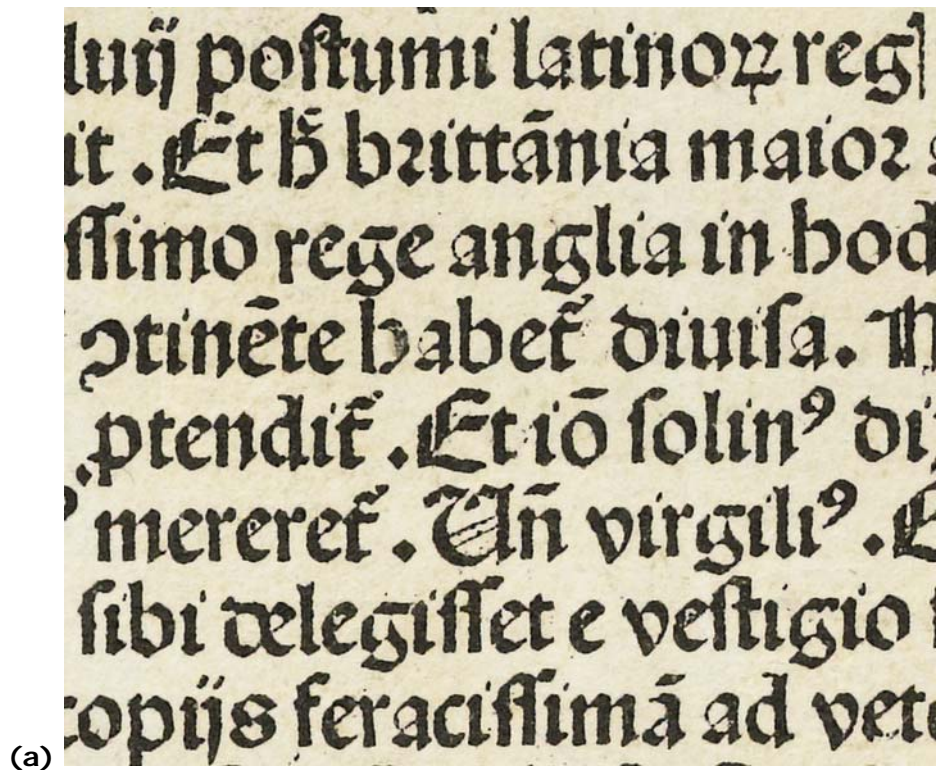
L0051761\_Painting



L0051440\_Archive\_Collection



**Figure 2: Comparison of irreversible with reversible compression**



Comparison of (a) irreversible compression with minimal loss and (b) reversible compression of region from L0051262\_Manuscript\_Page image, reproduced at 150 dpi.



3 of 13 DOCUMENTS

States News Service

December 7, 2011 Wednesday

## **WITH 'GOOGLE EARTH' FOR MARS, EXPLORE THE RED PLANET FROM HOME**

**BYLINE:** States News Service

**LENGTH:** 975 words

**DATELINE:** Tucson, AZ

The following information was released by the University of Arizona:

By Shelley Littin, NASA Space Grant intern, University Communications

A new software tool developed by the HiRISE team in the UA's Lunar and Planetary Lab allows members of the public to download high-resolution images of the Martian landscape almost instantaneously and explore the surface of the Red Planet from their own desktops.

Imagine zooming in over the surface of Mars, sweeping over sand dunes and circling around the rims of craters - all from your home desktop.

With HiView, the image-viewing tool recently released by the High Resolution Imaging Science Experiment, or HiRISE, team at the University of Arizona's Lunar and Planetary Lab, you can do just that.

"HiView is intended to be a tool that both scientists and the general public can use to explore the images at the highest possible quality level, and do so quickly and easily," said Rodney Heyd, the ground data systems manager for the HiRISE mission.

Mounted aboard NASA's Mars Reconnaissance Orbiter, the HiRISE camera sweeps the surface of the red planet collecting image data.

"HiRISE is a scanning imager," explained Bradford Castalia, the systems analyst and principle developer who designed the HiView software. "It scans the surface of Mars from the spacecraft and gathers information as it's going. And it has to do this very, very fast."

"Imagine you want to take a picture of a single grain of sand using a camera pointing out through a hole in the bottom of your car as you drive over that grain of sand at 200 miles an hour," said Castalia. "You're going to take a picture of one grain of sand, and you want it to be really sharp. That's what HiRISE is doing. It's pushing the electronics to the extreme."

The images produced by HiRISE are in the gigabyte size range: Up to tens of thousands of pixels across and more than 100,000 pixels high, the images are big enough to be murals on your living room wall. The current volume of image data from the HiRISE camera that has been made available to the public exceeds 67 Terabytes, said Heyd.

"This is the kind of research project that is fundamentally a data production operation. The research associated with it follows on from the production of the data," said Castalia. "We needed a way to get that

data out to people: to the science community and to the public at large."

"The need for a tool like HiView was known from the very beginning of the HiRISE project," said Heyd. "The HiRISE camera produces images so large that it can take an hour or more to download a single image. Such large images also have large storage requirements."

The HiRISE team adopted an image file format known as JPEG2000 to transfer and store the large volumes of imagery from the HiRISE camera. "This file format allows software supporting the JPEG2000 networking capabilities to download and view just a portion of any image in less than 30 seconds, resolving the need to spend a long time downloading an entire image and ensuring there is enough local storage space to hold the image," said Heyd.

In HiView, users can select a portion of the image and download only that portion of the image, so the user doesn't have to wait hours for the entire image file to download. "Once a region has been selected for viewing, only the area that fits on the computer screen will be downloaded, plus a small additional region around the edges to make panning around the image more seamless," said Heyd.

"HiView is the simplest way to access HiRISE imagery. Once the application has been downloaded to your computer, you can drag and drop a link to any image on the HiRISE website to the HiView application and view that image at the same quality that our scientists are using to analyze the features found in the imagery."

HiView is equipped with a set of data exploration tools so that users can move beyond just viewing the images and explore the data scientifically if they are so inclined.

Castalia hopes that the tools will entice people of all ages and backgrounds to pursue their interests in science. "It's really easy to use and very accessible," said Castalia. "And this allows high school or elementary school students to use HiView to see beautiful HiRISE imagery. I'm hoping that out of the corner of their eyes, they will see these tools and get excited about the science."

More than just an image-viewing tool, HiView also has image enhancement capabilities similar to the functions of Photoshop, and it is versatile in terms of the programs it is designed to work with, unlike many software applications.

A statistics tool gives information about the distribution of image values in a particular area of the image. A tool called a data mapper shows viewers a graph of source data and display data that allows users to apply contrast stretch to compare the source data and display data. HiView takes particular advantage of computers with multiple processing units and large or multiple desktops to provide a better use experience.

HiView isn't limited to HiRISE imagery. "HiView is able to read and write image data using conventional file formats so users can employ image formats that are suitable to their purpose," said Castalia. Members of the HiRISE science team use the program to get a first sense of what an observation contains, said Castalia.

"The pictures are beautiful, but there's real science in there," said Castalia. "That's part of what HiView allows people to do is to explore the science that's there in the imagery. People can really experience what it means to be involved in a Mars mission. All you have to do is get in there and explore."

"Given the volume of imagery that HiRISE produces no one has had time to fully analyze our entire data set," added Heyd. "So anyone with an interest in examining the imagery has the potential to discover something new."

**LOAD-DATE:** December 8, 2011

**LANGUAGE:** ENGLISH

**PUBLICATION-TYPE:** Newswire

WITH 'GOOGLE EARTH' FOR MARS, EXPLORE THE RED PLANET FROM HOME States News Service  
December 7, 2011 Wednesday

Copyright 2011 States News Service



**JPEG 2000 as a  
Preservation and Access Format  
for the Wellcome Trust Digital Library**



**Robert Buckley**  
Xerox Corporation

**Edited by Simon Tanner**  
*King's Digital Consultancy Services*  
*King's College London*



[www.kdcs.kcl.ac.uk](http://www.kdcs.kcl.ac.uk)

## Contents

<b>1. SUMMARY OF ISSUES/QUESTIONS .....</b>	<b>3</b>
<b>2. RECOMMENDATIONS.....</b>	<b>4</b>
<b>3 BASIS OF RECOMMENDATIONS/REASONINGS/TESTS DONE.....</b>	<b>5</b>
3.1 COMPRESSION .....	5
3.2 MULTIPLE RESOLUTION LEVELS.....	5
3.3 MULTIPLE QUALITY LAYERS.....	6
3.4 EXAMPLE: TIFF TO JP2 CONVERSION .....	6
3.5 MINIMALLY LOSSY COMPRESSION .....	7
3.6 TESTING REVERSIBLE AND IRREVERSIBLE COMPRESSION.....	7
3.7 FURTHER COMPRESSION FINDINGS.....	8
<b>4 IMPLEMENTATION SOLUTIONS / DISCUSSION.....</b>	<b>9</b>
4.1 COLOR SPECIFICATION .....	9
4.2 CAPTURE RESOLUTION.....	10
4.3 METADATA .....	10
4.3 SUPPORT .....	11
<b>5 CONCLUSION .....</b>	<b>12</b>
<b>APPENDIX 1: JPEG 2000 DATASTREAM PARAMETERS .....</b>	<b>13</b>
<b>APPENDIX 2: REVERSIBLE AND IRREVERSIBLE COMPRESSION .....</b>	<b>14</b>
<b>FIGURE 1: SAMPLE IMAGES.....</b>	<b>15</b>
<b>FIGURE 2: COMPARISON OF IRREVERSIBLE WITH REVERSIBLE COMPRESSION.....</b>	<b>17</b>

## 1. Summary of Issues/questions

The Wellcome Trust is developing a digital library over the next 5 years, anticipating a storage requirement for up to 30 million images. The Wellcome previously has used uncompressed TIFF image files as their archival storage image format. However, the storage requirement for many millions of images suggests that a better compromise is needed between the costs of secure long-term digital storage and the image standards used. It is expected that by using JPEG2000, total storage requirements will be kept at a value that represents an acceptable compromise between economic storage and image quality. Ideally, JPEG2000 could serve as both a preservation format and as an access or production format in a write-once-read-many type environment.

JPEG2000 was chosen as an image preservation format due to its small size and because it offers intelligent compression for preservation and intelligent decompression for access. If a lossy format is used to obtain a relatively high compression, e.g. between 5:1 and 20:1 (in comparison to an uncompressed TIFF file), then the storage requirements desired are achievable. The questions to address are what level of compression is acceptable and delivers the desired balance of image quality and reduced storage footprint.

With regard to the use of JPEG 2000, the questions posed in the brief and addressed in this report are:

- a. What JPEG2000 format(s) is best suited for preservation?
- b. What JPEG2000 format(s) is best suited for access?
- c. Can any single JPEG2000 format adequately serve both preservation and access?
- d. What models exist for the use of descriptive and/or administrative metadata with JPEG2000?
- e. If a JPEG2000 format is recommended for access purposes, what tools can be used to display/manipulate/manage it and any associated or embedded metadata?

This report will describe how a unified approach can enable JPEG2000 to serve for both preservation and access and balance the needs for compressed image size, image quality and decompression performance.

## 2. Recommendations

The majority of materials that will form part of the Wellcome Digital Library are expected to use visually lossless JPEG 2000 compression. Although “visually lossless” compression is lossy, the differences it introduces between the original and the image reconstructed from a compressed version of it are either not noticeable or insignificant and do not interfere with the application and usefulness of the image. Because the original cannot be reconstructed from the compressed image, compression in this case is irreversible. JPEG compression in digital still cameras is a familiar example of irreversible but visually lossless image compression. In mass digitization projects that use JPEG2000, compression ratios around 40:1 have been used for basically textual content. When applied to printed books, it has been found that these compression ratios do not impair the legibility or OCR accuracy of the text.

However, archiving and long-term preservation indicate a more conservative approach to compression and a different trade-off between compressed image size and image quality to meet current and anticipated uses. Still, given the volumes of material being digitized, a lossy format represents an acceptable compromise between quality and economic storage. A compression ratio of 4:1 or 5:1 gives a conservative upper limit on file size and decompressed image quality in the preservation format for the material being digitized. However this material should tolerate higher compression ratios with the results remaining visually lossless.

While most of the materials will use visually lossless compression, it is suggested that a small subset of materials (less than 5% of total) may be candidates for lossless or reversible compression. Reversible means that the original can be reconstructed exactly from the compressed image, i.e. the compression process is reversible.

Nevertheless, this report recommends irreversible JPEG2000 compression for the preservation and access formats of single grayscale or color images. Initially specifying a minimally lossy datastream will result in overall compression ratios around 4:1; the exact value will depend on image content. While this is a particularly conservative compression ratio, the compression can be increased as new materials are captured and even applied retroactively to files with previously captured material. The access format will be a subset of the preservation format with a subset of the resolution levels and quality layers in the preservation format.

In particular, the JPEG 2000 datastream should have the following properties:

- Irreversible compression using the 9-7 wavelet transform and ICT (see Section 3.1) with minimal loss (see Section 3.5)
- Multiple resolutions levels: the number depends on the original image size and the desired size of the smallest image derived from the JPEG 2000 datastream (see Section 3.2)
- Multiple quality layers, where all layers gives minimally lossy compression for preservation (see Sections 3.3 and 3.5)
- Resolution-major progression order (see Section 3.4)
- Tiles for improved codec (coder-decoder) performance, although the final decision regarding the use of tiles and precincts depends on the codec
- Generated using Bypass mode, which creates a compressed datastream that takes less time to compress and decompress (see Section 3.6)
- TLM markers (see Section 3.7)

A formal specification of the JPEG 2000 datastream for this application is given in Appendix 1. The datastream specified there is compatible with Part 1 of the

JPEG 2000 standard; none of the JPEG 2000 datastream extensions defined in Part 2 of the standard are needed.

Further, this report recommends embedding the JPEG 2000 datastream in a JP2 file. The JP2 file should contain:

- A single datastream containing a grayscale or color image whose content can be specified using the sRGB color space (or its grayscale or luminance-chrominance analogue) or a restricted ICC<sup>1</sup> profile, as defined in the JP2 file format specification in Part 1 of the JPEG 2000 standard (see Section 4.1)
- A Capture Resolution value (see Section 4.2)
- Embedded metadata that describes the JPEG 2000 datastream should follow the ANSI/NISO Z39.87-2006 standard and be placed in a XML box following the FileType box in a JP2 file (see Section 4.3)

Using the JP2 file format is sufficient as long as the requirement is for a single datastream whose color content can be specified using sRGB or a restricted ICC profile. While the JPX file format can be used if the color content of the image is specified by a non-sRGB color space or a general ICC profile, the use of a JP2-compatible file format is recommended.

## 3 Basis of recommendations/reasonings/tests done

### 3.1 *Compression*

In general terms, the compression ratio is set for preservation and quality, and JPEG 2000 datastream parameters such as the number of resolution levels and quality layers and tile size are set for access and performance. JPEG 2000 offers smart decompression, where only that portion of the datastream needed to satisfy the requested image view in terms of resolution, quality and location need be accessed and decompressed on demand and just in time.

The JPEG 2000 compression offers both reversible and irreversible compression. Reversible compression in JPEG 2000 uses the 5-3 integer wavelet transform and a reversible component transform (RCT). If no compressed data is discarded, then the original image data is recoverable from the compressed datastream created using these transforms. Irreversible compression uses the 9-7 floating point wavelet transform and an irreversible component transform (ICT), both of which have round-off errors so that the original image data is not recoverable from the compressed datastream, even when no compressed data is discarded. Appendix 2 contains a more detailed discussion of the differences between reversible and irreversible compression in JPEG 2000.

### 3.2 *Multiple resolution levels*

To begin with, it is recommended that JPEG 2000 be used with multiple resolution levels. The first two or three resolution levels facilitate compression; levels beyond that give little more compression but are added so that decompressing just the lowest resolution sub-image in the JPEG 2000 datastream gives a thumbnail of a desired size. For example, with a 5928-by-4872 pixel dimension original and 5 resolution levels, the smallest sub-image would have dimensions that would be 1/32 those of the original, in this case 186 by 153 pixels, which is roughly QQVGA<sup>2</sup> sized. Accordingly, JPEG 2000

---

<sup>1</sup> International Color Consortium, an organization that develops and promotes color management using the ICC profile format ([www.color.org](http://www.color.org))

<sup>2</sup> Quarter-quarter VGA (Video Graphics Array); since VGA is 640 by 480, QQVGA is 160 by 120.



compression with 5 resolution levels is recommended for images of this and similar sizes, which are typical of the sample images provided. In practice, the number of resolution levels would vary with the original image size so that the lowest resolution sub-image has the desired dimensions.

### 3.3 *Multiple quality layers*

There are two main reasons for using multiple quality layers. One is so that it is possible to decompress fewer layers and therefore less compressed data when accessing lower resolution sub-images. This speeds up decompression without affecting quality since the incremental quality due to the discarded layers is not noticeable at reduced resolutions. The second reason is that multiple quality layers make it possible to deliver subsets of the compressed image corresponding to higher compression ratios, which may be acceptable in some applications. This means there is less data to transmit and process, which improves performance and reduces access times. It also means that it is possible for the access format to be a subset of the preservation format, derived from it by discarding quality layers as the application and quality requirements warrant. The use of quality layers makes it possible to retroactively reduce the storage needs should they be revised downward by discarding quality layers in the preservation format and turning images compressed at 4:1 or 5:1 for example into images compressed at 8:1 or higher, depending on where the quality layer boundaries are defined.

### 3.4 *Example: TIFF to JP2 conversion*

For example, the following command line uses the Kakadu<sup>3</sup> compress function (`kdu_compress`) to convert a TIFF image to a JP2 file that contains an irreversible JPEG 2000 datastream. In particular, it contains a lossy JPEG 2000 datastream with 5 resolution levels and 8 quality layers, corresponding to compression ratios of 4, 8, 16, 32, 64, 128, 256 and 512 to 1 for a 24-bit color image. These correspond to compressed bit rates of 6, 3, 1.5, 0.75, 0.375, 0.1875, 0.09375 and 0.046875 bits per pixel. (A compression ratio of 4 to 1 applied to a color image that originally had 24 bits per pixel means the compressed image will equivalently have a compressed bit rate of 6 bits per pixel.) The Kakadu command line use bit rates rather than compression ratios to specify the amount of compression.

```
kdu_compress -i in.tif -o out.jp2 -rate  
6,3,1.5,0.75,0.375,0.1875,0.09375,0.046875 Creversible=no  
Clevels=5 Stiles={ 1024,1024} Cblk={ 64,64} Corder=RPCL
```

The JPEG 2000 datastream created in this example has 1024-by-1024 tiles, 64-by-64 codeblocks and a resolution-major progressive order RPCL, so that the compressed data for the lowest resolution (and therefore smallest) sub-image occurs first in the datastream, followed by the compressed data needed to reconstruct the next lowest resolution sub-image and so on. This data ordering means that the data for a thumbnail image occurs in a contiguous block at the start of the datastream where it can be easily and speedily accessed. This data organization makes it possible to obtain a screen-resolution image quickly from a megabyte or gigabyte sized image compressed using JPEG 2000. Tiles and codeblocks are used to partition the image for processing and make it possible to access portions of the datastream corresponding to sub-regions of the image.

---

<sup>3</sup> <http://www.kakadusoftware.com/>

### 3.5 *Minimally Lossy Compression*

The JPEG 2000 coder in this example would discard transformed and compressed data to obtain a compressed file size corresponding to 4:1 compression. This needs to be compared with the performance of the the minimally lossy coder, where no data is discarded but which is still lossy because of the use of the irreversible transforms. In some cases, depending on the image content, as shown in Section 3.6, the minimally lossy coder can give higher compression ratios than 4:1. Accordingly, it is recommended that a minimally lossy format with multiple quality layers and multiple resolution levels be used for the preservation format. The access format would use reduced quality subsets of the preservation format optionally obtained by discarding layers and using reduced resolution levels.

### 3.6 *Testing reversible and irreversible compression*

The reason to use irreversible compression is that it gives better compression than reversible compression, at the cost of introducing errors (or differences) in the reconstructed image. This section examines this performance tradeoff.

Reversible and irreversible compression were applied to four images provided by the Wellcome Digital Library (Figure 1). A variation on irreversible compression was tested which used coder bypass mode, in which the coder skipped the compression of some of the data. This gave a little less compression, but made the coder (and decoder) run about 20% faster. The Kakadu commands used in these tests are given in Appendix 2.

The compression ratios obtained with these three test are shown in the following table.

Original	Reversible	Irreversible	Irreversible w/bypass
L0051262_Manuscript_Page	2.25	3.45	3.42
L0051320_Line_Drawing	1.82	2.52	2.51
L0051761_Painting	2.46	3.96	3.90
L0051440_Archive_Collection	2.52	4.47	4.41

For these particular images, the compression ratio for irreversible JPEG 2000 was from about 40% to almost 80% better than it was for reversible, and on average over 30% faster (with a further 20% boost with coder bypass mode).

The cost of irreversible compared to reversible is the error it introduces. The error or difference between the reversibly and irreversibly compressed images is about 50 dB, which means the average absolute error value was about 0.5. For one of the sample images, 99.99% of the green component values were the same after decompression as they were before, or at most two counts different. (For the red and blue components, the percentages were 99.79 and 99.35.) This is within the tolerance for scanners: in other words, minimally lossless irreversible JPEG 2000 compression adds about as much noise to an image as a good scanner does.

A region was cropped from one image so that the visual effects of this error on this image could be examined more closely (Figure 2). When they were, the differences were not perceptible on screen or on paper. Unless being able to reconstruct the original scan is a requirement, legal or otherwise, then irreversible compression is clearly advantaged over reversible compression.

### 3.7 Further compression findings

In these tests, the compressed file sizes (and compression ratios) were image dependent and varied with the image content. Images with less detail or variation than these samples would give even higher compression ratios.

An advantage of JPEG 2000 is that it lets the user set the compression ratio, or equivalently the compressed file size, to a specific target value, which the coder achieves by discarding compressed image data. While this feature was not used to set the overall compressed file size in the minimally lossy compression case, it can be used to set the sizes of intermediate images corresponding to the different quality layers. The following Kakadu command line generates a JP2 file with a minimally lossy irreversible JPEG 2000 datastream that complies with the recommendation in this report:

```
kdu_compress -i in.tif -o out.jp2 -rate -, 4, 2.34, 1.36, 0.797,
0.466, 0.272, 0.159, 0.0929, 0.0543, 0.0317, 0.0185
Creversible=no Clevels=5 Stiles={ 1024,1024} Cblk={64,64}
Corder=RPCL Cmodes=BYPASS
```

The JPEG 2000 datastream in this example has 5 resolution levels and 12 quality layers. Using all 12 layers give a decompressed image with minimal loss. The intermediate layers boundaries are at pre-set compressed bit rates, starting at 4 bits per pixel, corresponding to a compression ratio of 6:1, assuming a 24-bit color original. Thereafter, the layer boundaries are distributed logarithmically up to a compression ratio of 1296:1. The exact values are not critical. What is important is the range of values and there being sufficient values to provide an adequate sampling within the range.

When a datastream has multiple quality layers, it is possible to truncate it at points corresponding to the layer boundaries and obtain derivative datastreams that correspond to higher compression ratios (or lower compressed bit rates). In the previous example, discarding the topmost quality layer produces a datastream corresponding to a compression ratio of 6:1 (compressed bit rate of 4 bits per pixel). Discarding the next layers produces a datastream with a compression ratio of 10.3:1, and so on. As noted previously, some images may have minimally lossy compression ratio greater than 6:1; the layer settings can be adjusted when this happens.

Using layers adds overhead that increases the size of the datastream and therefore decreases the compression ratio. To assess this effect as well as the overhead due to the use of tiles, the four sample images were compressed with one layer and no tiles, with one layer and 1024x1024 tiles, and with 12 layers and no tiles. As the following table shows, adding layers and tiles did decrease the minimally lossy compression ratio, but the effect was only visible in the third place after the decimal and was therefore judged insignificant in comparison to the advantages of using them.

Original	No tiles 1 layer	1024x1024 tiles 1 layer	No tiles 12 layers
L0051262_Manuscript_Page	3.452	3.450	3.443
L0051320_Line_Drawing	2.522	2.521	2.517
L0051761_Painting	3.961	3.957	3.948
L0051440_Archive_Collection	4.477	4.473	4.461

Besides tiles and layers, other datastream components that can improve performance and access within the datastream are markers, such as Tile Length Markers (TLM) which can aid in searching for tile boundaries in a datastream. Their effectiveness depends on whether or not the decoder or

access protocol makes use of them. As a result, recommendations regarding their use depend on the choice of codec.

## 4 Implementation solutions / discussion

This section discusses the file format and metadata recommendations.

One function of a file format is packaging the datastream with metadata that can be used to render, interpret and describe the image in the file. Besides defining the JPEG 2000 datastream and core decoder, Part 1 of the JPEG 2000 standard also defines the JP2 file format which applications may use to encapsulate a JPEG 2000 datastream. A minimal JP2 file consists of four structures or "boxes":

1. JPEG 2000 Signature Box, which identifies the file as a member of the JPEG 2000 file format family
2. File Type Box, which identifies which member of the family it is, the version number and the members of the family it is compatible with
3. JP2 Header Box, which contains image parameters such as resolution and color specification needed for rendering the image
4. Contiguous Codestream Box, which contains the JPEG 2000 datastream

### 4.1 Color Specification

How an image was captured or created determines the parameters in the JP2 Header Box, which are subsequently used to render and interpret the image. Among these parameters are the number of components (i.e. whether the image is grayscale or color), an optional resolution value for capture or display, and the color specification. In general the color content of an image can be specified in one of two ways: directly using a named color space, such as sRGB, Adobe RGB 98 or CIELAB, or indirectly using an ICC profile.

The digitization process and the nature of the material being digitized, not the file format, drive the color specification requirements of the application. The issue for the file format is whether or not it supports the color encoding used by the digital materials. What's significant about the JP2 file format is that it supports a limited set of color specifications. For example, the only color space it supports directly is sRGB, including its grayscale and luminance-chrominance analogues. This is a consequence of the JP2 file format having been originally designed with digital cameras in mind.

Besides sRGB, the JP2 file format supports a restricted set of ICC profiles, namely gamma-matrix-style ICC profiles. This style of profile can represent the data encoded by RGB color spaces other than sRGB. The image data is still RGB; it's just that it is specified indirectly by means of an ICC profile. This does not necessarily mean that non-sRGB systems must support ICC workflows; it does mean more sophisticated handling of the color specification in the JP2 file. For example, the system may recognize that the JP2 file contains the ICC profile for Adobe RGB 98 and use an Adobe RGB 98 workflow.

An alternative to JP2 is the Baseline JPX file format, defined in Part 2 of the JPEG 2000 standard. JPX is an extended version of JP2 which, among other things, specifies additional named color spaces, including Adobe RGB 98 and ProPhoto RGB. There are some RGB spaces, such as eciRGBv2, which JPX does not support directly and for which ICC profiles would still be needed for them to be used. The best thing is to use the JP2 file format as long as possible, since it is more widely supported than JPX and its use avoids support for the more advanced features of JPX when only extended color space support is desired.

## 4.2 Capture Resolution

The JP2 Header Box may also contain a capture or display resolutions, indicating the resolution at which the image was captured or the resolution at which it should be displayed. While the JP2 file is required to contain a color specification, it is not required to have either resolution values. Instead, it is up to the application to require it. This report recommends that the JP2 Header Box in the JP2 file contain a capture resolution value, indicating the resolution at which the image contained in the file was scanned. The JP2 file format specification requires that the resolution value be given in pixels per meter.

## 4.3 Metadata

In addition to the four boxes that a JP2 is required to contain, it may optionally contain XML and UUID boxes. Each can contain vendor or application specific information, encoded in an XML box using XML or in a UUID box in a way that is interpreted according to the UUID code (UUID stands for Universally Unique Identifier). These two types of boxes are used to embed metadata in a JP2 file. For example, UUID boxes are used for IPTC<sup>4</sup> or EXIF<sup>5</sup> metadata. An XML box can be used for any XML-encoded data, such as MIX.

While the application and system normally determine the nature and format of the metadata associated with an image, JPEG 2000-specific administrative or technical metadata is within scope for this report. While such metadata may or may not be embedded in a JP2 file, this reports recommends that it be embedded.

JPEG 2000-specific metadata in the JP2 file should follow the ANSI/NISO Z39.87-2006 standard. This standard defines a data dictionary with technical metadata for digital still images. It lists "image/jp2" as an example of a *formatName* value and lists "JPEG2000 Lossy" and "JPEG2000 Lossless" as *compressionScheme* values. Files that implement this recommendation would have "JPEG 2000 Lossy" as their *compressionScheme* value and would also contain a rational *compressionRatio* value.

Compression	compressionScheme	JPEG2000 Lossy
	compressionRatio	<rational value>

While "JPEG2000 Lossy" is the *compressionScheme* value for all files that follow this recommendation and the *compressionRatio* value can be derived from file size and parameters in the JP2 Header Box, it is recommended that they be specified explicitly.

The Z39.87 standard also defines a *SpecialFormatCharacteristics* container to document attributes that are unique to a particular file format and datastream. In the case of JPEG 2000, this container has two sub-containers: one for *CodecCompliance* and the other for *EncodingOptions*. The elements in the *CodecCompliance* container identify by name and version the coder that created the datastream, the profile to which the datastream conforms (Part 1 of the JPEG 2000 standard defines codestream or datastream profiles), and the class of the decoder needed to decompress the image (Part 4 of the JPEG 2000 defines compliance classes). The elements in the *EncodingOptions* container give the size of the tiles, the number of quality layers and the number of resolution levels.

<sup>4</sup> International Press Telecommunications Council, creates standards for photo metadata (<http://www.iptc.org/IPTC4XMP/>)

<sup>5</sup> Exchangeable image file format, a standard file format with metadata tags for digital cameras (<http://www.exif.org/>)



The following table shows the hierarchy of *SpecialFormatCharacteristics* containers and elements for JPEG 2000; the column on the right shows the values these elements would have for the data stream generated by the example in Section 3.7.

JPEG2000	CodecCompliance	codec	Kakadu
		codecVersion	6.0
		codestreamProfile	1
		complianceClass	2
	EncodingOptions	tiles	1024x1024
		qualityLayers	12
		resolutionLevels	5

An XML schema for these and the other elements defined in the Z39.87 standard is available at <http://www.loc.gov/standards/mix>.

When metadata is embedded in a JP2 file, it would be convenient if it were near the beginning of the file where it could be found and read quickly. Any XML or UUID boxes containing metadata can immediately follow the JPEG 2000 Signature and FileType boxes, which must be the first two boxes in a JP2 file. This means the metadata can come before the JP2 Header box, which in turn must come before the Contiguous Codestream box. Therefore, the metadata-containing boxes can occur in a JP2 file before any of the image data to which their contents pertain.

### 4.3 Support

JPEG 2000 is supported by several popular image image editors, toolkits and viewers. Among them are Adobe Photoshop, Corel Paint Shop Pro, Irfanview, ER Viewer, Apple QuickTime and SDKs from Lead Technologies and Accusoft Pegasus. Aside from the viewers, all offer automated command lines and batch support.

Other sources of JPEG 2000 components and libraries are Kakadu, Luratech<sup>6</sup> and Aware<sup>7</sup>. This is not an exhaustive list. JP2 is not natively supported by Web browsers. In this regard, it is like PDF and TIFF and, like them, Web browser plug-ins are available, such as from Luratech and LizardTech<sup>8</sup>.

A common approach for delivering online images from a JPEG 2000 server is to decode just as much of the JPEG 2000 image as is needed in terms of resolution, quality and position to create the requested view, and then convert the resulting image to JPEG at the server for delivery to a client browser. This avoids the need for a client side plug-in to view JPEG 2000. The National Digital Newspaper Program (NDNP) uses this approach; it offers 1.25 million newspapers pages, all stored as JPEG 2000, on its website at <http://chroniclingamerica.loc.gov/>. While NDNP uses a commercial JPEG 2000 server from Aware, other commercial servers as well as the Djakota<sup>9</sup> open-source JPEG 2000 image server are also available.

The choice between a JPEG 2000 client or a JPEG client with server-side JPEG2000-to-JPEG conversion is a system issue that is largely independent of the JPEG 2000 datastream and file format recommendation in this report.

<sup>6</sup> <http://www.luratech.com/>

<sup>7</sup> <http://www.aware.com/imaging/jpeg2000.htm>

<sup>8</sup> [http://www.lizardtech.com/download/dl\\_options.php?page=plugins](http://www.lizardtech.com/download/dl_options.php?page=plugins)

<sup>9</sup> <http://www.dlib.org/dlib/september08/chute/09chute.html>

## 5 Conclusion

This report has described the use of JPEG 2000 as a preservation and access format for materials in the Wellcome Digital Library. In general terms, the compression ratio is set for preservation and quality, and JPEG 2000 datastream parameters such as the number of resolution levels and quality layers and tile size are set for access and performance.

This report recommends that the preservation format for single grayscale and color images be a JP2 file containing a minimally lossy irreversible JPEG 2000 datastream, typically with five resolution levels and multiple quality layers. To improve performance, especially decompression times on access, the datastream would be generated with tiles and in coder bypass mode.

One consequence of using a minimally lossy JPEG 2000 datastream is that the compressed file size will depend on image content, which will create some uncertainty in the overall storage requirements. The ability with JPEG 2000 to create a compressed image with a specified size will reduce this uncertainty, but replace it with some variability in the quality of the compressed images. The sample images could tolerate more than minimally lossy compression; how much more depends on the quality requirements of the Wellcome Library, which will depend on image content. Until these requirements are articulated and validated, and even after they are, using quality layers in the datastream will provide a range of compressed file sizes to satisfy future image quality-file size tradeoffs.

This report recommends that the access format be either the same as the preservation format or a subset of it obtained by discarding quality layers to create a smaller and more compressed file. Requests for a view of a particular portion of an image at a particular size would be satisfied in a just-in-time on-demand fashion by accessing only as many of the tiles (and codeblocks within a tile), resolution levels and quality layers as are needed to obtain the image data for the view.

## Appendix 1: JPEG 2000 Datastream Parameters

This table specified the values for the main parameters of the JPEG 2000 datastream.

<b>JPEG 2000 Datastream Parameters</b>	
<b>Parameter</b>	<b>Value</b>
<b>SIZ marker segment</b>	
Profile	Rsiz=2 (Profile 1)
Image size	Same as scanned original
Tiles	1024 x 1024
Image and tile origin	XOsiz = YOsiz = XTOsiz = YTOsiz = 0
Number of components	Csiz = 1 (grayscale) or 3 (color)
Bit depth	Determined by scan
Subsampling	XRsiz = YRsiz = 1
<b>Marker Locations</b>	
COD, COC, QCD, QCC	Main header only
<b>COD/COC marker segments</b>	
Progression Order	RPCL
Number of decomposition levels	$N_L = 5$
Number of layers	Multiple (see text)
Code-block size	xcb=ycb=6
Code-block style	SPcod, SPcoc = 0000 0001 (Coding Bypass)
Transformation	9-7 irreversible filter
Precinct size	Not explicitly specified



## Appendix 2: Reversible and Irreversible Compression

This appendix describes the operation of a JPEG 2000 coder and the differences between reversible and irreversible compression. The first thing a JPEG 2000 coder does with an RGB color image is to apply a component transform which converts it to something better suited to compression by reducing the redundancy between the red, green and blue components in a color image. The next thing it does is apply multiple wavelet transforms, corresponding to the multiple resolution levels, again with the idea of making the image more suitable for compression by redistributing the energy in the image over different subbands or subimages. The step after that is an optional quantization, where image information is discarded on the premise that it will hardly be missed (if it's not overdone) and will further condition the image for compression. The next step is a coder, which takes advantage of all the prep work that has gone on before and the resulting statistics of the transformed and quantized signal to use fewer bits to represent it; this is where the compression actually occurs. The coder doesn't discard any data and is reversible, although the steps leading up to it may not be. In a JPEG 2000 coder, there is one more step in which the compressed data is organized to define the quality layers boundaries and to support the resolution-major progressive order mentioned earlier.

For JPEG 2000 compression to be reversible, there can't be any quantization or round-off errors in the component and wavelet transforms. To avoid round-off errors, JPEG 2000 has a reversible transform based on integer arithmetic. So for example, JPEG 2000 specifies a reversible wavelet transform, called the 5-3 transform because of the size of the filters it uses. JPEG 2000 also specifies an irreversible wavelet transform, the 9-7 transform, based on floating point operations. The 9-7 transform does a better job than the 5-3 transform of conditioning the image data and so gives better compression, but at the cost of being unable to recover the original data due to round-off errors in its calculations. Because of these round-off errors, the 9-7 transform is not reversible.

The following Kakadu commands were used to generate the compressed images for the tests reported in Section 3.6. The first command generates a reversible JPEG 2000 datastream; the second, an irreversible but minimally lossy datastream; and the third, an irreversible datastream using coder bypass.

```
kdu_compress -i in.tif -o out.jp2 -rate - Creversible=yes  
Clevels=5 Stiles={ 1024,1024 Cblk={ 64,64 } } Corder=RPCL  
//reversible
```

```
kdu_compress -i in.tif -o out.jp2 -rate - Creversible=no  
Clevels=5 Stiles={ 1024,1024 Cblk={ 64,64 } } Corder=RPCL  
//irreversible
```

```
kdu_compress -i in.tif -o out.jp2 -rate - Creversible=no  
Clevels=5 Stiles={ 1024,1024 Cblk={ 64,64 } } Corder=RPCL  
Cmodes=BYPASS
```

The dash after the rate parameter in these commands indicates that all transformed and quantized data is to be retained and none discarded. Creversible=yes in the first command directs the coder to use the reversible wavelet and components transforms. Creversible=no in the second and third commands directs the use of irreversible transforms. Cmodes=BYPASS in the third command directs the coder to use Bypass mode.

# Figure 1: Sample images



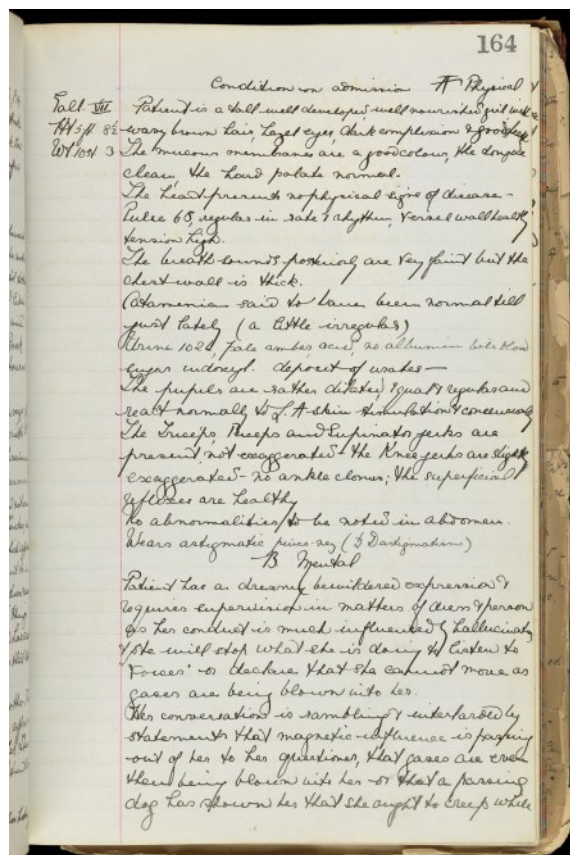
L0051262\_Manuscript\_Page



L0051320\_Line\_Drawing



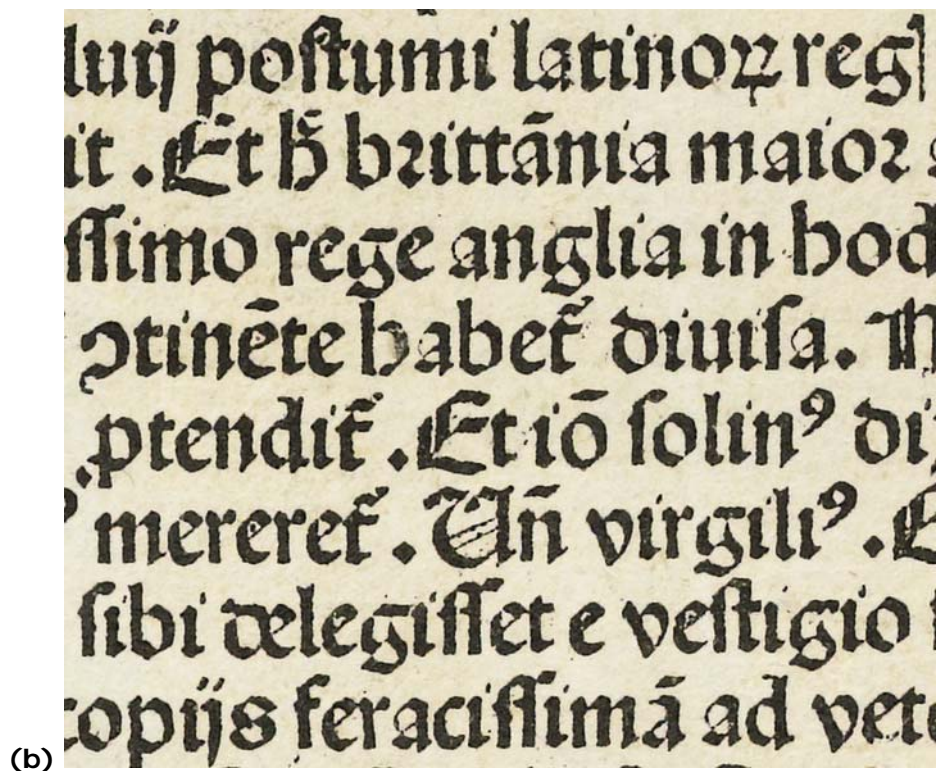
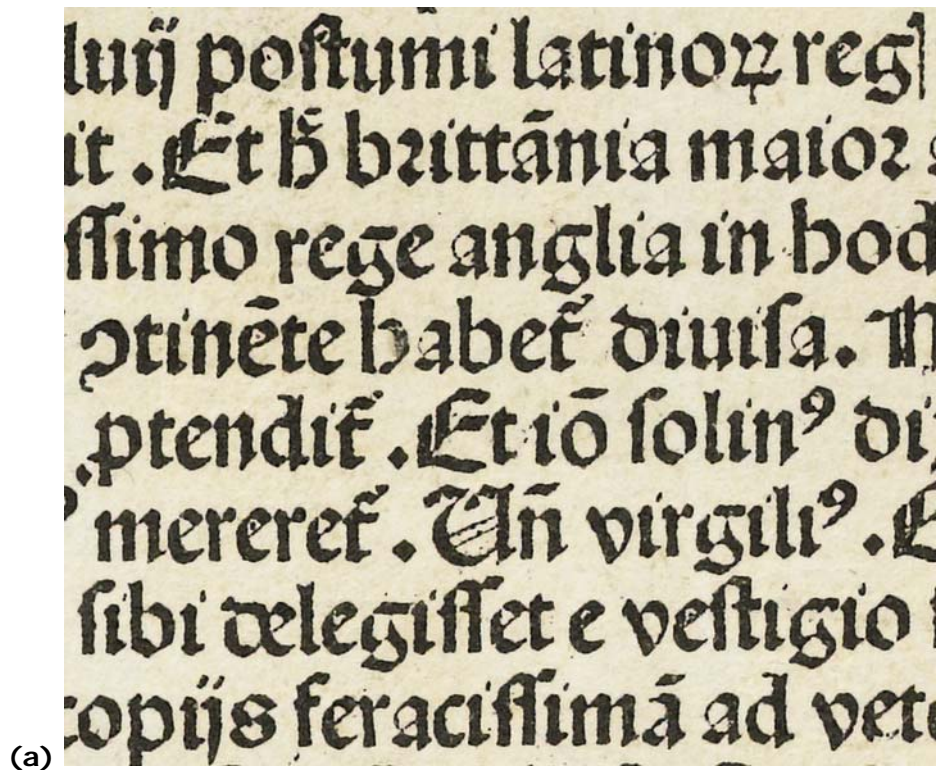
L0051761\_Painting



L0051440\_Archive\_Collection



**Figure 2: Comparison of irreversible with reversible compression**



Comparison of (a) irreversible compression with minimal loss and (b) reversible compression of region from L0051262\_Manuscript\_Page image, reproduced at 150 dpi.

## **JPEG 2000 Page Image Compression for Large-Scale Digitization at Yale**

Quality Control Working Group

Large-Scale Digitization Project

Yale University Library

January 17, 2008

The Quality Control Working Group has examined numerous examples of JPEG 2000 page images both lossless and at various levels of lossy compression. We recommend that Yale adopt lossy compression at Kirtas' quality level 90 as the standard for Yale's copy of the images produced in the Microsoft/Kirtas large-scale digitization project. We are confident that this level of compression provides the best compromise between image quality and manageable file size for an access-oriented project of this type.

In her white paper, *Preservation in the Age of Large-Scale Digitization*, Oya Rieger (Interim Assistant University Librarian for Digital Libraries and Information Technologies at Cornell) argues persuasively that the nature of large-scale digitization necessitates a change in the way that libraries have traditionally viewed the choice of file formats (and many other issues) for digital imaging projects.<sup>1</sup> First of all, large-scale digitization of library material produces such vast quantities of digital data (roughly 100,000 page images per working day in our case) that some form of image compression becomes essential in order to enable affordable long-term storage of the content. Even with the level of compression recommended by the Yale working group, the first 100,000 volumes will occupy roughly 25 terabytes of disc storage (50 TB when replicated) – a quantity approaching double the original size of the entire Rescue Repository, just to give a local perspective. Secondly, the primary goal of these Google, Microsoft, and Open Content Alliance projects is enduring access and full-text indexing rather than digital reformatting. The digital file will not replace the physical book. Consequently, the use of a compressed file format that supports searching and projected scholarly use of the content should be fully adequate to the need.

JPEG 2000 (commonly abbreviated as JP2) is an ISO standard that is well supported and increasingly adopted by both the commercial and academic communities. It utilizes a highly effective compression scheme that achieves dramatic reduction in file size without serious loss of quality (outperforming the earlier JPEG standard). A JP2 file can deliver images at multiple resolutions, ranging from a thumbnail to a highly zoomed detail – ideal for the rapid delivery of image files via the web. Metadata of various types can be encapsulated within a JP2 file, ensuring that metadata about an image will never be separated from the image. While current web browsers do not provide native support for JP2, server-based solutions or browser plug-ins can resolve this (probably temporary) issue.

Even some preservation-oriented repositories are considering the adoption of JPEG 2000 in its lossless form. The National Library of the Netherlands, for instance, is challenging the commonly accepted objections to the use of compression in digital preservation and argues that there may even be advantages over uncompressed TIFFs, most notably the fact that future data

---

<sup>1</sup> Oya Rieger. September, 2007. *Preservation in the Age of Large-Scale Digitization: A White Paper*. Washington, D.C.: Council on Library and Information Resources.  
Available at <http://www.clir.org/activities/details/mdpres.html>.

migration will require the processing of a much smaller quantity of data. This library suspects that “some form of compression can be performed on objects intended for long-term preservation as long as you keep to some common sense rules” and plans to conduct extensive tests to determine the suitability of JPEG 2000 for a digital archive.<sup>2</sup>

### Harvard & CDL Studies

Several libraries, including Harvard and the California Digital Library, have conducted studies designed to determine what level of compression is acceptable for large-scale digitization projects.<sup>3</sup> The principle behind their approach is to “get just enough data” to satisfy current and foreseeable needs of users:

“Three overarching functional requirements of mass text digitization favor the get just enough rationale over the get more. Digitization procedures in these projects must:

- enable very fast scanning of bound volumes (to reduce costs associated with human handling time);
- yield page image masters adequate for OCR and production of images with readable (legible) content when rendered as soft- and hard-copy outputs; and
- result in small file sizes, with the two-fold benefit of speeding up online transfer (ingest and access) and minimizing per unit (page/volume) storage costs.”

The studies presented reviewers with sets of images compressed at different levels using various software codecs and asked them to rate the quality of the images as Perfect, Acceptable, Marginal, and Unacceptable when compared to an uncompressed master image or a lossless JP2 image. Results consistently revealed that reviewers judged images acceptable or better at surprisingly high levels of lossy compression.

An attached set of screen shots shows the effects of JP2 lossy compression in one example from the Harvard test suite (full image sets at [http://preserve.harvard.edu/massdig/hul\\_study/](http://preserve.harvard.edu/massdig/hul_study/)). Only in the last instance (the LuraTech Q30 level) do the compression artifacts become clearly noticeable and fairly objectionable (file size reduced from 3420 KB to 290 KB). Even then, the text is eminently readable. At the Q70 and Q50 quality levels, compression artifacts are hard to detect, yet dramatic reduction in files sizes is still achieved.

### Analysis at Yale

Studies at Yale of samples provided by Kirtas closely parallel and strongly confirm the results yielded in the Harvard study. An attached set of screen shots shows examples of both visual image and plain text content compressed at different quality levels (since Kirtas is not using the

---

<sup>2</sup> Judith Rog. 2007. “Compression and Digital Preservation: Do They Go Together?” Pp. 80-83 in Archiving 2007, vol. 4. Proceedings of a conference of the Society of Imaging Science and Technology, May 21-24, 2007, Arlington, Va. Available at <http://www.imaging.org/store/epub.cfm?abstrid=34425>.

<sup>3</sup> Stephen Chapman, Laurent Duploux, John Kunze, Stuart Blair, Stephen Abrams, Catherine Lupovici, Ann Jensen, Dan Johnston. 2007. “Page Image Compression for Mass Digitization.” Pp. 37-42 in Archiving 2007, vol. 4. Proceedings of a conference of the Society of Imaging Science and Technology, May 21-24, 2007, Arlington, Va. Available at <http://www.imaging.org/store/epub.cfm?abstrid=34415> or [http://preserve.harvard.edu/massdig/hul\\_study/IST\\_PageImageCompression\\_preprint.pdf](http://preserve.harvard.edu/massdig/hul_study/IST_PageImageCompression_preprint.pdf) (full preprint).

## JPEG 2000 Compression

LuraTech software, the numerical quality levels are not directly comparable to those cited in the Harvard study). Since our goals include a high level of readability for text and visual images of a quality high enough for future inclusion in the Image Commons, we quickly narrowed our examination to the higher levels of quality offered (and recommended) by Kirtas.

With plain black and white text, undesirable consequences become apparent only at relatively high levels of compression, manifested in a fringing or halo effect surrounding the letters on a page and most obvious at high levels of magnification. These effects are substantially reduced in our case because we elected early in the Yale project to remove background shading on book pages and produce a purely black-on-white image for pages containing only black text. In addition, this approach results in improved OCR, reduction in file size, better readability on screen, and cleaner printed copies. In the examples the difference between quality levels 90 and 80 are very hard to discern.

With photographs, maps, illustrations, manuscript facsimiles, and other non-text content, negative effects are visible at more modest levels of compression – the effects vary widely, however, depending on the particular image. In the (worst case) example provided, you can see a difference between the lossless image and the compressed image even at the highest quality level (90) if you zoom in on a small detail. At lower zoom levels the difference is much harder to detect.

The attached table lists quality assessments and file sizes for a variety of illustrations and text. There are only small differences in quality between the level 80 and level 90 images; there was no readily visible difference in 9 of the 17 samples even at high magnification. Level 90 images averaged 18% of the size of the lossless version for mixed content and 40% of lossless size for plain text. Level 80 images averaged 12.8% of the size of the lossless version for mixed content and 33.6% of lossless size for plain text. For mixed content the level 80 files were 29% smaller than level 90; for plain text the level 80 files were 16.6% smaller than level 90. Thus the overall savings in repository disk space for level 80 compared to level 90 will be somewhere in between these two figures, most likely in the neighborhood of 20%.

In the end the decision rests on priorities. Is it worth investing in 20% more repository disk space in order to attain a level of quality one step closer to the lossless original? Given the objectives of the Yale project and our desire to preserve files that will be suitable for a variety of future purposes, the Quality Control Working Group recommends Kirtas' quality level 90 lossy JPEG 2000 format as the best compromise between superior image quality and reasonable storage requirements.