Using Lossy JPEG 2000 Compression

For Archival Master Files

Prepared

for the

Library of Congress
Office of Strategic Initiatives

by

Robert Buckley
NewMarket Imaging

Version 1.1

March 12, 2013

Prime Contract/Solicitation No.: LCOSI10T0057
Prime Task Order No. OSI12T0010

Contents

Abstract

This report documents the study that reviewed how different institutions have used the control settings and command line parameters available with various JPEG 2000 encoders to create visually lossless archival and production masters. The study has confirmed the image-dependent nature of the controls and consequently the need for human observers to judge what settings are appropriate for an image type or collection. Almost all memory institutions that use lossy compression use controls that specify compression by size, which generates results that are portable between implementations and institutions. Google Books and the projects associated with it which use slope to specify compression by quality. Harvard Library uses a PSNR-based control process for compression that runs automatically but must be monitored for known failure cases of a global metric for assessing image quality. For the time being, there is no viable alternative to image-dependent observer-based control strategies, although visual discrimination metrics continue to offer the promise of automatic and supervised methods for visually lossless compression. These metrics need to be mindful of the characteristics of the display used in the visual assessment task and their effect on the acceptability thresholds for lossy compression in archival masters.

Introduction

At the heart of an archival system is the preservation or archival file. It “represents the best copy produced by a digitizing organization,” where “best” is defined with respect to the objectives of a particular project or program [1]. Further, it represents “digital content that the organization intends to maintain for the long term without loss of essential features.” For still image content, the archival master has traditionally been an uncompressed TIFF file, as this has been believed to be the best way to preserve essential features as captured in the original scan. However for mass digitization projects, using a compressed format reduces costs compared to an uncompressed format and is thus economically attractive.

The usual arguments against using a compressed image format for an archival file are that it makes the file and its contents more vulnerable to errors, whether the compression is lossless or lossy, and that lossy compression introduces differences from the digitized original that may interfere with achieving a project’s objectives and that could grow with subsequent migrations. However once these concerns have been examined in terms of the ubiquity of error correction [2], noise in the digitization process and the resilience of image-based tasks to compression, lossy compression becomes not only economically attractive but also acceptable for archival masters of digitized originals. In most cases, preserving the original scan without compression is not warranted from a quality standpoint. If substituting a lossy compressed image for an uncompressed image does not affect performance of a given task, then the lossy compression can be described as “functionally lossless.” In visual assessment tasks, the term “visually lossless” is commonly used and is the visual equivalent of “functionally lossless.”

The organizations that have adopted a compressed format for the archival masters in mass digitization projects for the most part are using JPEG 2000 as the compression method[1]. Some use lossless JPEG 2000 compression, which is one step removed from uncompressed, offering reversible compression so that the digitized original is recoverable. Organizations are increasingly using lossy JPEG 2000 compression for the archival master, especially for mass digitization projects. Whether they use lossless or lossy compression for the archival master, most all use lossy JPEG 2000 for the production master file. For those who use lossy JPEG 2000 compression for the archival master, this means the same file can serve as both the archival and production master.

When lossy compression is accepted for the archival image master, the question becomes how much lossy compression to use (and how do you specify it). The two basic methods of specifying compression are "by size" and "by quality." By size means specifying the size of the compressed file or the compression ratio; by quality means using a control that is monotonically related to the quality of the compressed image as measured using a metric such as PSNR[2]. The relationship between the two methods is not a simple one and varies with image content.

Specifying compression by size

JPEG 2000, unlike JPEG, makes it possible to compress a file to a desired size. The size can be specified directly in bytes, or indirectly in terms of a compression ratio or a compressed bit rate. The compression ratio is the size of the uncompressed file to the size of the compressed file, which contains both metadata and image data. (Compression factor is the reciprocal of compression ratio.) The compressed bit rate is the ratio of the compressed image data size to the product of the width and height of the image[3]. In whatever way the file size is specified and passed to the coder, the coder achieves the desired result by retaining only as much image data as is needed to meet the target file size.

If a 24-bit original image is compressed with a compression ratio of 8:1, then the compressed bit rate will be close to but not precisely 3.0. This is because compression ratio is based on file size while compressed bit rate is based on the compressed image data size, which is less than the file size. Similarly, if the compressed bit rate is specified as 3.0, then the compression ratio will not be precisely 8:1. The differences are usually small enough that stating the compression ratio as the number of bits per pixel in the uncompressed image divided by the compressed bit rate is convenient and not significantly different from the value obtained using the definition.

The rest of this section describes the approaches that three institutions have taken to establishing how much compression to use when specifying compression by size.

---

[1] Some use TIFF with LZW lossless compression and a few use Baseline JPEG, which is a lossy compression format more often used for access and delivery.

[2] PSNR or Peak Signal to Noise Ratio is 10 times the logarithm of the square of the maximum pixel value (peak signal) divided by the mean squared error (noise), and is expressed in dB or decibels. The error is the difference between the image before and after compression/decompression; for lossless compression, the error is zero.

[3] The definition of compressed bit rate is from Kakadu v6.0.

National Digital Newspaper Program

The National Digital Newspaper Program (NDNP) is a partnership between the National Endowment for the Humanities and the Library of Congress to provide online access to the digitized content of historic newspapers via **http://chroniclingamerica.loc.gov/** . Newspaper pages are scanned in grayscale at 8 bits per pixel and 300 or 400 dpi. NDNP uses uncompressed TIFF for archival masters of the page scans and lossy JPEG 2000 compression for production masters; the JPEG 2000 production profile is described in [3]. Since the uncompressed TIFF file from the scan is being retained as the archival master during Phase 1 of the Program, there was no a priori requirement for the production master to be lossless and "visually lossless" compression was acceptable. In this case, the intended uses of the production master are viewing and printing. While the program does not use JPEG 2000 for archival masters, the process that it used to specify lossy JPEG 2000 compression was among the first to be publicly documented and can be adapted for use with archival masters[4].

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 halftoned image, 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 for 8-bit gray originals. Figure 1 shows the images of the test areas; for details see Appendix C of [3].

(a) (b) (c)

Figure 1. Test images with (a) halftoned, (b) line-art and (c) text 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

[4] This description of the process is based on Section 3.1 of [3]

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 compression levels. Overall print quality was adequate but less important than visual screen presentation quality.

As a result, NDNP adopted 8:1 compression for the JPEG 2000 production masters. This was judged a good compromise between file size and image quality. Since higher compression ratios may be acceptable for some purposes, quality 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 between 1 and 0.25 bits per pixel).

So far no issues have been reported with the use of compression in the production masters, even when the user-selected views derived from them are transcoded to JPEG for delivery to the user browser. As a rule, scan quality is a more important factor in image quality and user satisfaction than compression quality at this level of compression.

It is worth noting that for a fixed compressed bit rate, the amount of distortion that compression introduces into an image depends on the image content. For example, the more complex an image, the more detail or high spatial frequency content it has and the more compressed data (detail) needs to be discarded to achieve the target compressed file size compared to a less complex image with less detail. These differences will be evident in the PSNR values, which will be lower for the more complex images than for the less complex ones when both are compressed to the same bit rate. However, for the types of image content used here—halftoned image, line-art and text—and at these compression ratios, no significant differences were reported between the quality of the image content at the same bit rate. These content types are fairly homogeneous in these test images.

Also worth noting in the NDNP use of JPEG 2000 encoding is the generous use of quality 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.

For maximum quality at maximum resolution, the user would direct the decoder to decompress all layers in reconstructing the image. However, at lower resolutions and for thumbnail images, the user could decompress fewer layers to obtain a decompressed image without objectionable visual artifacts. The use of layers effectively enables higher compression ratios to meet the needs of lower resolution and lower quality applications. Using fewer layers means less compressed data to access and decompress, which speeds up performance. [4]

Altogether, the NDNP 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 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 2 plots the bit rates and compression ratios against layers. 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.

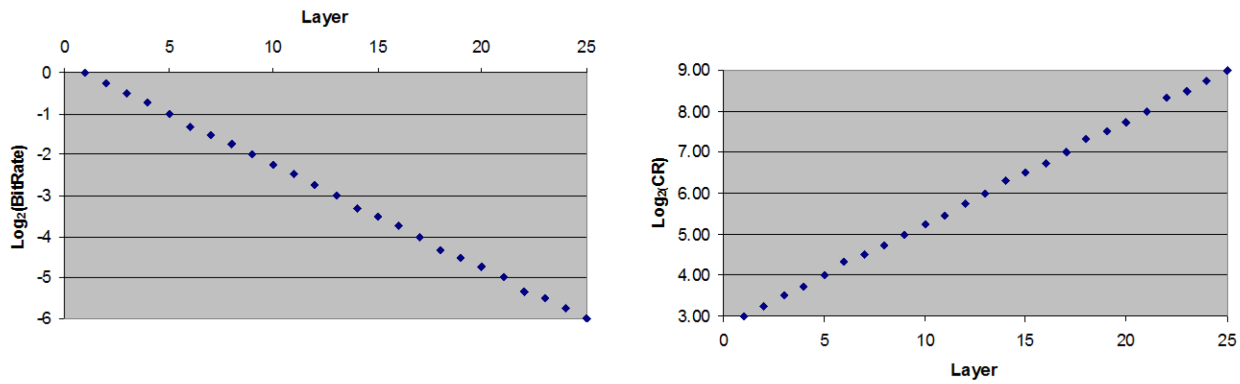


Figure 2: Plots of (a) Log_2 of Bit Rate and (b) Log_2 of Compression Ratio (CR) vs. Layer

Within a fixed bit budget method such as specifying compression by size, 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. With less compressed image data, the greater the difference between the original and the image that can be reconstructed from the compressed image data. However, for the sample images in the NDNP tests, 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 is a relatively small price to pay to obtain the advantages of quality scalability. Future applications may use fewer layers and consequently fewer quality sampling points.

The British Newspaper Archive

The British Newspaper Archive (BNA) is a partnership between the British Library and bright**solid** online publishing to digitize up to 40 million newspaper pages from the British Library's collection over the next 10 years [5]. Pages are scanned in color at 24 bits per pixels and 400 dpi. The page images are then compressed using JPEG2000.

Unlike NDNP, which uses uncompressed TIFF for archival masters and lossy JPEG 2000 for production masters, the BNA uses lossy JPEG 2000 for both archival and production purposes. While recognizing that there is a trade-off between the quality of a compressed file and its size and consequently the cost of file storage, the British Library decided that lossy compression was not only economically desirable but also acceptable. Further, they decided that it was desirable to have the same master file support both archival and production uses, which simplifies file management.

In general terms, a production profile is designed for performance and fast access to the image subsets that support the targeted use cases, while an archival profile is designed for image quality and file integrity. Fortunately, with JPEG 2000 the two are compatible so that one profile can meet both purposes. The profile that the British Library uses for the mass digitization of newspapers in the British Newspaper Archive serves for both production and archival use. For production purposes, the key choices have to do with the use of precincts vs. tiles, progression order, the use of codestream markers and coding modes. For archival purposes, the main choice is the amount of compression and how it is specified.

The British Library profile uses 12 quality layers with the 9-7 irreversible wavelet transform [6]. The layers correspond to compressed bit rates of 10, 8.7, 7, 5.2, 3.4, 2.1, 1.2, 0.6, 0.3, 0.15, 0.075 bits per pixel (corresponding to compression ratios of 2.4, 2.8, 3.4, 4.6, 7.1, 11.4, 20, 40, 80, 160, 320 for a 24-bit per pixel image), plus a maximum bit rate corresponding to the retention of all transformed and quantized image data. Although all the data is retained, there is still some minimal loss because of round-off errors with the floating point transforms and from the quantization used in the lossy compression mode of JPEG 2000: this is the minimally lossy case. As in the case of lossless compression, the bit rate corresponding to the minimally lossy case is not known beforehand and depends on the image content. The first bit rate that is explicitly specified is 10 bits per pixel, corresponding to a compression ratio of 2.4:1. It is conceivable that the minimally lossy compression ratio for some images will be higher, which means the compressed image will have fewer than 12 layers. In any case, the image recovered by decompressing all layers will correspond to the minimal loss with the selected quantization levels.

The minimally lossy case sets a lower bound on the amount of compression loss for an image. Most if not all images can tolerate more loss. Using layers permits increasing the amount of compression after the image is initially compressed by eliminating layers from the compressed image data, an operation that can be performed without decompressing (and recompressing) the image. For example, eliminating the top three layers would be the same as compressing the image with a compression ratio of 3.4:1. The alternative is to set the amount of compression in advance based on the characteristics of the image or the collection from which an image is drawn, or institutional factors such as affordability and budget constraints.

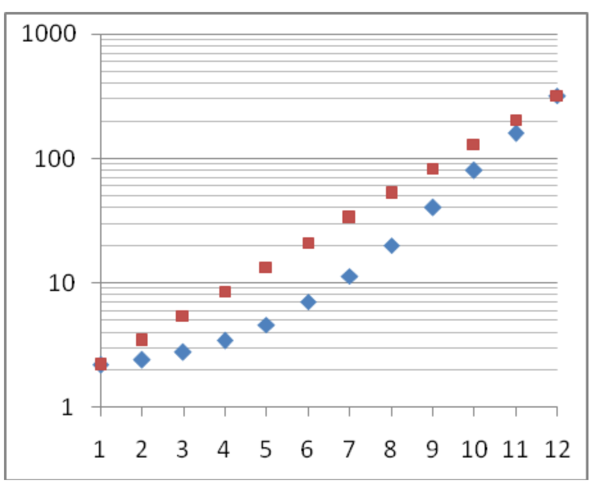


Figure 3: Distribution of Compression Ratios over Quality Layers 1 through 12: red squares show a uniform distribution; blue diamonds, the distribution in the British Library profile

The logarithms of the compression ratios of the layers in the British Library profile are not uniformly distributed across layers, as they are in the NDNP profile. Figure 3, which plots the logarithm of compression ratio on the vertical axis against layer number on the horizontal axis, shows the differences. The red squares correspond to a uniform distribution; the blue diamonds, to the compression ratios used in the British Library profile.

Layer 12 corresponds to an image with a compression ratio of 320:1. Layer 11 in the British Library profiles corresponds to a compression ratio of 160:1. Decompressing Layer 11 means decompressing Layers 11 and 12. In general, decompressing Layer N means decompressing Layers N, N+1, N+2 and so on since the compressed data in a given layer is a subset of the data in the preceding (next lower number) layer. Layer 2 corresponds to a compression ratio of 2.4:1. Layer 1 is shown here with a compression ratio of 2, although as noted earlier the actual value is image dependent.

The non-uniform distribution of compression ratios in the British Library profile was selected to provide a finer sampling of compression ratios in the range 2-10 than a uniform distribution. This is the range in which the archival encoder is expected to operate and archival JPEG 2000 masters will likely have compression ratios of 10 or less. The option still exists to use fewer layers in some production use cases; for example, smaller image derivatives, such as thumbnails, would use fewer layers.

Most of the cost benefits of compression are obtained in the compression ratio range from 2 to 10. This is shown in Figure 4, which plots the savings in storage costs versus compression ratio. The plot is based on a simple model that assumes cost is proportional to storage volume. This is a reasonable model to use for the cost of cloud storage, which is often priced per unit volume per unit time, e.g. TB/month. Figure 4 shows a steep rise in cost savings with low levels of compression, rising to a 90% cost reduction at a 10:1 compression ratio. Diminishing returns set in at that point and the cost savings curve begins to level off; doubling the compression beyond that only achieves a further 5% in cost savings at the risk of significantly increasing the cost of poor quality.

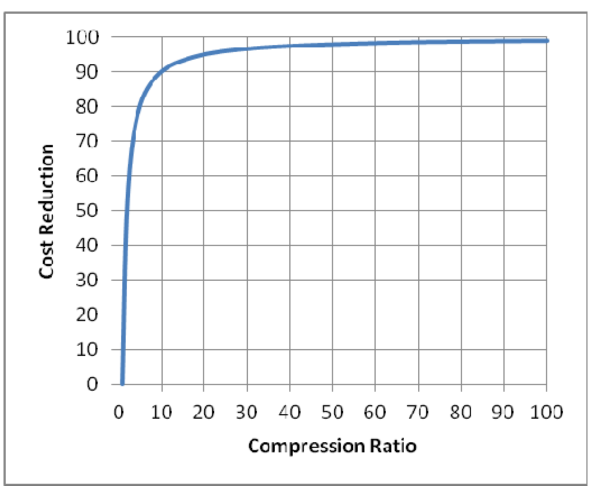


Figure 4: Storage cost reduction versus compression ratio

Pending adoption of a method for setting the compression level on an image-by-image or collection basis, the British Library is using minimally lossy compression, which reserves the ability to “retrofit” or specify a higher compression level in future by discarding quality layers in archived masters.

Wellcome Digital Library

The Wellcome Trust is developing a digital library with the requirement to store up to 30 million images. They previously used uncompressed TIFF image files as archival masters. However, the storage requirement for this volume of images pointed to the need for a better compromise between the costs of secure long-term digital storage and the image standards used.

In its collections, the Wellcome Digital Library holds:

- Printed books – early printed books, modern books (monographs), pamphlets, reports
- Archives – personal papers, institutional papers, unpublished works, mostly 20th century
- Manuscripts – unpublished, handwritten “manuscript books” and related materials, mostly 17th, 18th and 19th century
- Artworks – prints, paintings, posters, drawings and glass slides.

Figure 5 shows sample materials from the Wellcome Library collection.

Figure 5. Sample Wellcome Library materials: Line drawing, manuscript page and page from archive collection.

The Wellcome Trust has adopted lossy JPEG 2000 compression for archival masters for large projects and with it the notion of using the same file for both preservation and production in a write-once-read-many type environment. JPEG 2000 was chosen as the image storage format in order to keep total storage requirements at a value that represented an acceptable compromise between economic storage and image quality. In addition, JPEG2000 offers intelligent compression for preservation and intelligent decompression for production and access. While the Wellcome Library also uses a master JPEG 2000 file with multiple quality layers, the main question is what overall level of compression is acceptable to deliver the desired balance between image quality and reduced storage footprint.

Since the Wellcome Digital Library wanted to use compression to manage their storage requirements, specifying compression "by size" made sense. To determine the acceptable level of compression, they selected and viewed representative images from a "collection," increasing the compression ratio in steps until compression artifacts become visible. They then backed off one step and used that compression ratio for the other images in the collection [7]. An alternate approach that they also used started with a high level of compression that was clearly unacceptable and worked backwards, decreasing the compression until obtaining an image that was "fit for purpose." These approaches are variations on the classic psychophysical method of limits for detecting a stimulus and determining a threshold. The images were viewed at 100% on a color calibrated monitor, and subjects were asked to detect a difference between successive compression steps or between images before and after compression.

If different representative images from a collection gave different compression thresholds, then the lowest compression ratio was used for all the images in the collection. In other words, the "worst performing" image determined the compression value for the entire collection. In results to date, a 10:1 compression ratio was found to work well for books and an 8:1 value for materials from the archive collection.

An advantage of a size-based approach to setting compression levels is that it enables a user to know or estimate in advance how much space is needed for storage, which is useful for planning purposes. However, a point that has emerged from the use of a size-based approach is that the same compression level or ratio that may be acceptable for some images will not be for others. In other words, the same compression level can give different quality levels for different images. In order for the same compression level to achieve acceptable quality levels across a set of images, it should be applied conservatively or to images with the same or similar content, which is how it has been applied in the cases described so far. Typically the most demanding images and the most discerning observer set the compression level. A more direct way of achieving acceptable quality levels would be to specify compression by quality.

Specifying compression by quality

The idea behind compression by quality is to set the amount of compression with a method that, unlike specifying compression by size, produces a result that depends on or adapts to the image content.

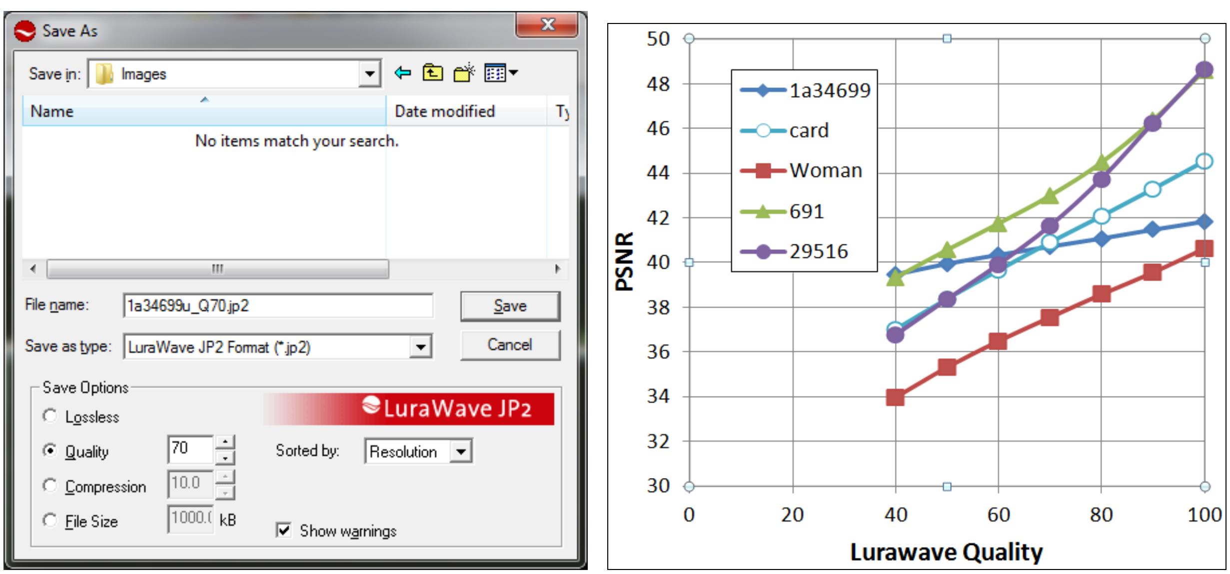
Compression by quality is familiar to JPEG users. When users elect to save an image as a JPEG file, they are typically presented with a slider along which they can adjust a pointer from "low quality, most compression" at one end to "high quality, least compression" at the other. The higher the quality factor, the finer the quantization of the coefficients of the transformed image and the smaller the difference between the compressed and original uncompressed images. Because of content differences, the same quality factor will give different compression ratios for different images; specifying compression by size is only possible with JPEG by using an iterative process. One potential drawback of quality-based approaches is that the scales for adjusting quality have not been standardized and there are multiple ways of controlling and measuring quality.

The rest of this section describes different methods that vendors offer and organizations have used for specifying compression by quality.

Quality Factor – Luratech, Adobe

The Luratech JP2 SDK and command line interface offers Quality as a parameter when saving an image as a JP2 file. Figure 6(a) shows the Save As dialog box from the LuraWave SmartCompress tool, where Quality is one of the Save Options; compression ratio and compressed file size are also Save Options. The range of Quality values is 1 to 100, with 1 the minimum and 100 the maximum. The lower the Quality value, the coarser the quantization of the subbands of the wavelet-transformed image beyond the smallest. In the tests performed on the JPEG 2000 compression of book pages, the California Digital Library and Internet Archive reported that only a professional digital photographer could detect compression artifacts at a Quality value of 70 [8].

Figure 6(b) plots PSNR against Quality for the samples images shown in Figure 7[5]. For the color images in this sample, the plot of PSNR vs. Quality is a straight line over the range of Quality values from 40 to 100; the intercept and slope of the line are image dependent. For the grayscale images in this sample, the plot was slightly bowed. A priori there is no reason to believe that different images with the same Quality value will have the same visual quality or that there exists a single Quality value that represents the threshold at which compression artifacts become visible across a set of images.



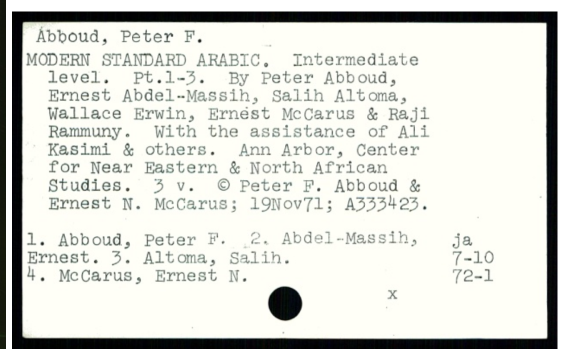
(a) (b)

[5] The sample images were compressed using LuraWave SmartCompress Version 2.1.05.03 from Luratech.

Figure 6. Luratech Quality setting (a) LuraWave SmartCompress Save As Dialog Box with Quality Save Option set to 70; (b) Plot of PSNR vs. Quality for sample images in Figure 7.

Adobe Photoshop also offers a Quality setting when saving an image as a JPF file, including a JP2-compatible JPF file. However, the Photoshop Quality setting works differently from the Luratech one. For all images, the plot of compressed file size vs. Quality is a straight-line, with compressed file size decreasing as Quality decreases from an image-dependent maximum at a Quality of 100[6]. For a Quality value of Q, the compressed file size is Q% of the maximum file size. The quantization in the Photoshop encoder does not change with the Quality setting, so the different-sized files at different Quality settings are obtained by discarding compressed data. In effect, Photoshop Quality controls the compression factor on an image-by-image basis.

1a34699



Abboud, Peter F.
MODERN STANDARD ARABIC. Intermediate level. Pt.1-3. By Peter Abboud, Ernest Abdel-Massih, Salih Altoma, Wallace Erwin, Ernest McCarus & Raji Rammuny. With the assistance of Ali Kasimi & others. Ann Arbor, Center for Near Eastern & North African Studies. 3 v. © Peter F. Abboud & Ernest N. McCarus; 19Nov71; A333423.

1. Abboud, Peter F. 2. Abdel-Massih, Ernest. 3. Altoma, Salih. 4. McCarus, Ernest N.

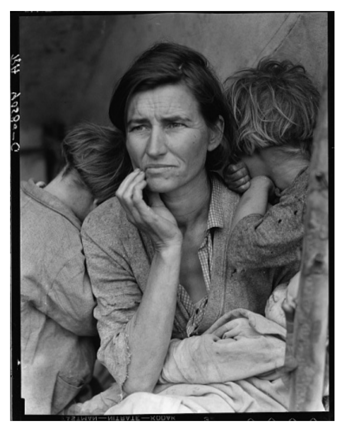
ja
7-10
72-1

x

card

Woman

691



29516

Figure 7. Test Images: 1a34699 (6667x5177, 24-bit color); card (1542x945, 24-bit color); Woman (2048x2560, 24-bit color); 691 (1306x1294, 8-bit gray); 29516 (6648x8358, 8-gray)

Slope –Google Books, Internet Archive

When specifying compression by size, JPEG 2000 encoders use post-compression rate-distortion optimization to select a subset of the compressed data that gives a compressed file with the specified

[6] Selected sample images were compressed using Adobe Photoshop CS3.

size and the lowest distortion [9]. Not all compressed data equally effects the distortion or error, so that the compressed data to include first in the file is the data that gives the largest reduction in distortion for the smallest increase in file size. This optimization produces a slope value that estimates the change in distortion with the change in file size.

An alternative to directly specifying the compressed bit rate is directly specifying this slope value. This is an option in the Kakadu and Aware command line interfaces[7]. Slopes are inversely related to bit rates, so that the lower the slope, the higher the bit rate and the smaller the distortion. In Kakadu, the slope values can range from 0, corresponding to image with all the compressed data with none discarded to 65535, corresponding to an image where all the compressed data has been discarded. The useful operating range for the slope parameter in the Kakadu encoder is between about 49000 and 53000; the higher the slope number, the lower the quality.

At the JPEG 2000 Summit in May 2011, Google described their use of slope values in the JPEG 2000 encoding for the Google Books project [10]. Using the Kakadu encoder, they reported acceptable quality in day-to-day operations at a slope value of 50980; images exported after image processing used a quality value of 51492. The Michigan Digitization Project, which is related to Google Books, requires that all continuous tone images be delivered in JPEG 2000 format and that the default slope rate distortion value for encoding grayscale images will be 51492 [11]. In the study that Chapman et al. reported on in 2007 [8], a slope value of 51492 gave perfect or acceptable quality ratings 93% of the time with text pages and 87% of the time with non-text pages. By comparison, a slope value of 52516 gave acceptable quality ratings only 5% of the time with text and non-text pages; the other 95% of the time the pages were judged marginal or unacceptable.

A sampling of color and gray-scale page images from the Internet Archive shows that they were created with a slope value of 50995 using Kakadu V5.1. The corresponding compression ratios ranged from 8:1 for a gray scale page with text and a halftoned illustration to 351:1 for a blank color page with some show-through, with 52:1 for a color page with black-and-white text. Just as images with the same compression ratio will have different amounts of distortion depending on the image content, images created using the same slope value will have different compression ratios depending on their content.

Figure 8 plots PSNR against slope value for the sample images of Figure 7. As the plot clearly shows, the relationship is image-dependent. Compression artifacts are visible in all images created with a slope value of 52516. In some images, notably Image 691, artifacts were visible at a slope value of 50995 when they were viewed at 100% magnification on a laptop screen in an uncontrolled environment that can make some claim to the way in which users would typically view images[8]. Kakadu V6.0 was used to create the compressed images with slope values between 49200 and 52516.

[7] Since the use of visual weights is the default in the Kakadu command line interface, their use must be turned off when using the slope to minimize distortion in the sense of PSNR or mean squared error.
[8] Uncontrolled here means free form and not measured so the environment may not be able to be replicated. See the Discussion section for a discussion of the potential for display characteristics to affect compression results.

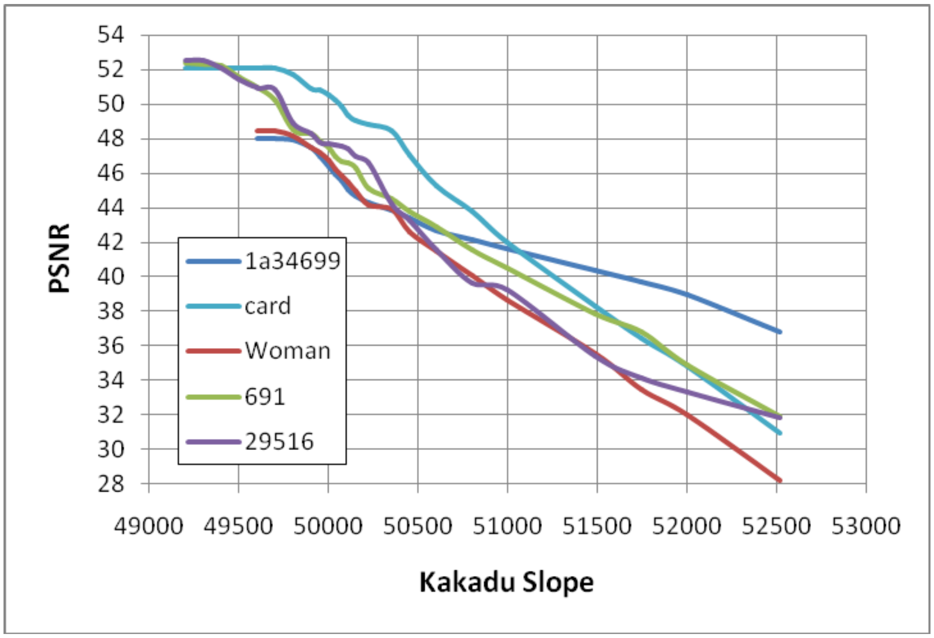


Figure 8. Plots of PSNR vs. Slope for sample images in Figure 7

PSNR – Harvard Library

The plots in Figures 6(b) and 8 show that the relationship between PSNR and the Quality and Slope control settings is image dependent. For any one image, this relationship can be characterized and then applied to images that share the same relationship, provided they can be identified, to obtain a desired target PSNR value. An alternative is to use an image model that allows an analytic derivation of the relationship between control settings and PSNR (or some other image metric). The desired PSNR would then be obtained when compressing images that the model describes. This leaves aside for the moment the question as to how well PSNR estimates visual quality across different images subject to the same treatment, in this case JPEG 2000 compression.

The Aware JPEG 2000 encoder command line interface allows users to specify a target PSNR value when encoding an image. At the JPEG 2000 Summit, Harvard Library reported obtaining high-fidelity copies of page images using the Aware JPEG 2000 encoder with a PSNR setting of 46 db [12]. However, they also reported that this approach could produce overly compressed images with noticeable compression artifacts. Figure 9 shows an example from [12].

Figure 9. Comparison of (a) uncompressed and (b) compressed images; see text for description of (b).

These failures occurred in images with large uniform areas and abrupt transitions and that depart significantly from the underlying model. To guard against this kind of failure and over-compression, Harvard uses a script that compresses the file to a target PSNR value and then checks to see if the compression ratio exceeds a pre-determined maximum value. If it does, then the target PSNR value is incremented by one and the file compressed again. This continues until the compressed ratio falls below the maximum or until a maximum PSNR value reached, in which case a warning is issued.

This behavior also highlights the limitations of a global metric such as PSNR—a single number to describe the effect of compression over the entire image, although compression does not affect all parts of an image equally. Larger distortions and very visible compression artifacts in some parts of the image can be offset by smaller distortions in the rest of the image to give an average distortion for the entire image that would be acceptable if it occurred uniformly over the image.

Discussion

This study has reviewed how different institutions have used the control settings and command line parameters available with various JPEG 2000 encoders to create visually lossless archival and production masters. The criterion for judging whether a given level of lossy compression is acceptable is that it introduces no visible or functional differences in the decompressed image compared to the original image before compression and that in all the ways that matter the decompressed is an acceptable surrogate for the original.

The following table summarizes the controls available in selected encoders for specifying compression by size and by quality.

	Compression by size				Compression by Quality		
Encoder	Compressed bit rate	Compression ratio	Compressed file size	Compression factor	Quality factor	Slope	PSNR
Kakadu	X					X	
Luratech		X	X		X		
Aware	X	X				X	X
Jasper[1]				X			
OpenJPEG[2]		X					X
Photoshop			X		X		
Paint Shop Pro			X	X			

1. ImageMagick uses Jasper for JPEG 2000 encoding and adds a quality parameter that is mapped to the compression factor in Jasper by a non-linear function.
2. GIMP requires the OpenJPEG library for JPEG 2000 encoding and decoding.

It is worth noting that the different controls listed under “Compression by size” in the table are all expressions of the same underlying metric and are simply related one to another so that Compression by size is a standardized and easily replicated approach to image compression. Different implementations produce comparable and compatible results. By comparison, the controls listed under “Compression by quality” in the table are all different, and the settings and results for Quality factor in particular are not portable between implementations.

Whichever approach is used to control JPEG 2000 compression, they all get mapped onto two mechanisms that the encoder uses when creating compressed data: specifying the subband quantization steps and deciding which parts of the compressed data stream are retained and written to the output file.

As has been observed, the mapping from the settings or values of any of these controls to the visibility threshold for compression artifacts is image dependent. As a result, mappings such as the ones that NDNP or the Wellcome Library uses for compression ratio and that Google Books uses for slope are applied conservatively and to a collection or type of image. So far there has not been a mapping that applies to all images nor is there likely to be one when the criterion is detecting compression artifacts since these controls and the associated metrics are global. They use a single number that applies to the entire image when visual discrimination is usually based on local features.

PSNR is a different control from the others in the table in the sense that one cannot apply it directly to the encoding process and expect to get back a compressed image with the specified PSNR value. The exception is when the image fits the model that was assumed in the mapping from PSNR value to the compression controls. Otherwise to obtain a compressed image with a specific PSNR value, an iterative process is needed that goes through multiple compress-compare cycles until the desired PSNR value is obtained. By comparison, if you specify the compression ratio or slope value, then you will get a compressed image with the specified compression ratio or slope value.

Even when a specific PSNR value is achieved, it is well known that PSNR is not a reliable predictor of image quality or the image quality differences across images and image treatments, such as compression [13]. A metric that shows a better relationship to image quality is the SSIM (Structural Similarity) index [14]. Besides being a global metric, it can be applied locally, which may be useful in identifying regions in an image whose structural content has been most disturbed by compression. A local metric would be superior in mimicking the process that observers use to evaluate image compression and determine the visibility threshold for compression artifacts. However, the SSIM index is more challenging analytically than PSNR and achieving a target value will require an iterative process. The global SSIM Index has been used to assess the effects of JPEG 2000 compression on the scale from visually lossless to visually lossy compression [14, 15]. Our interest here is in finding the boundary or threshold between the two for a class or collection of images.

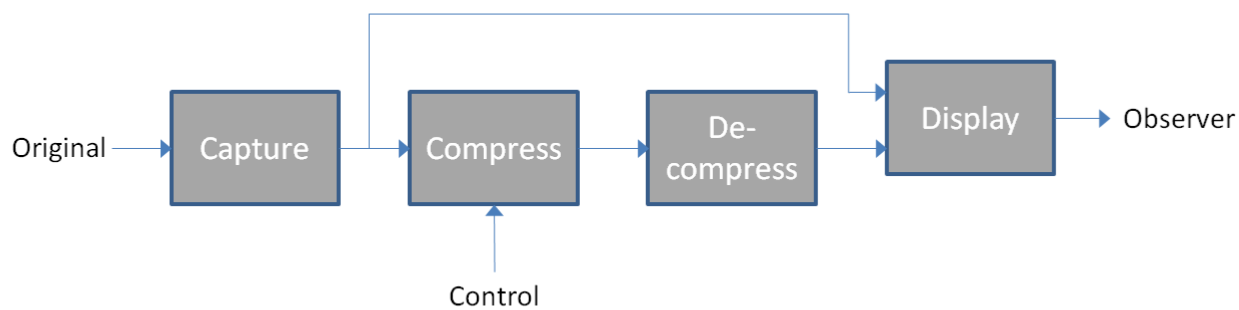


Figure 10. Block diagram of system for assessing compression artifacts

Figure 10 shows the steps involved in the visual assessments reviewed in this study. An original is captured, compressed using specified control parameters and then decompressed. The resulting decompressed image and the original captured image are displayed to an observer who compares them, looking for differences and artifacts. Because the observer can only base an assessment on the images presented by the display, it has been implicit the display faithfully presents to the observer the compression artifacts that matter to the assessment. This assumption needs to be examined.

The spatial frequency response (SFR) of the capture device can affect the compressibility of the image. For example, decreasing the SFR or sampling efficiency will increase the lossless or minimally lossy compression ratio for an image, all other things being equal. Also, when two images are compressed to the same lossy compression ratio, the image with the lower SFR will have a lower mean square error. The SFR of the display device can also affect the compressibility of an image, or more precisely the visual assessment of the compressibility of an image. For example, two images, one compressed and the other uncompressed, that are distinguishable on one display may be indistinguishable on a display with a lower SFR. In other words, a poor display may make an image look more compressible than it really is or should be.

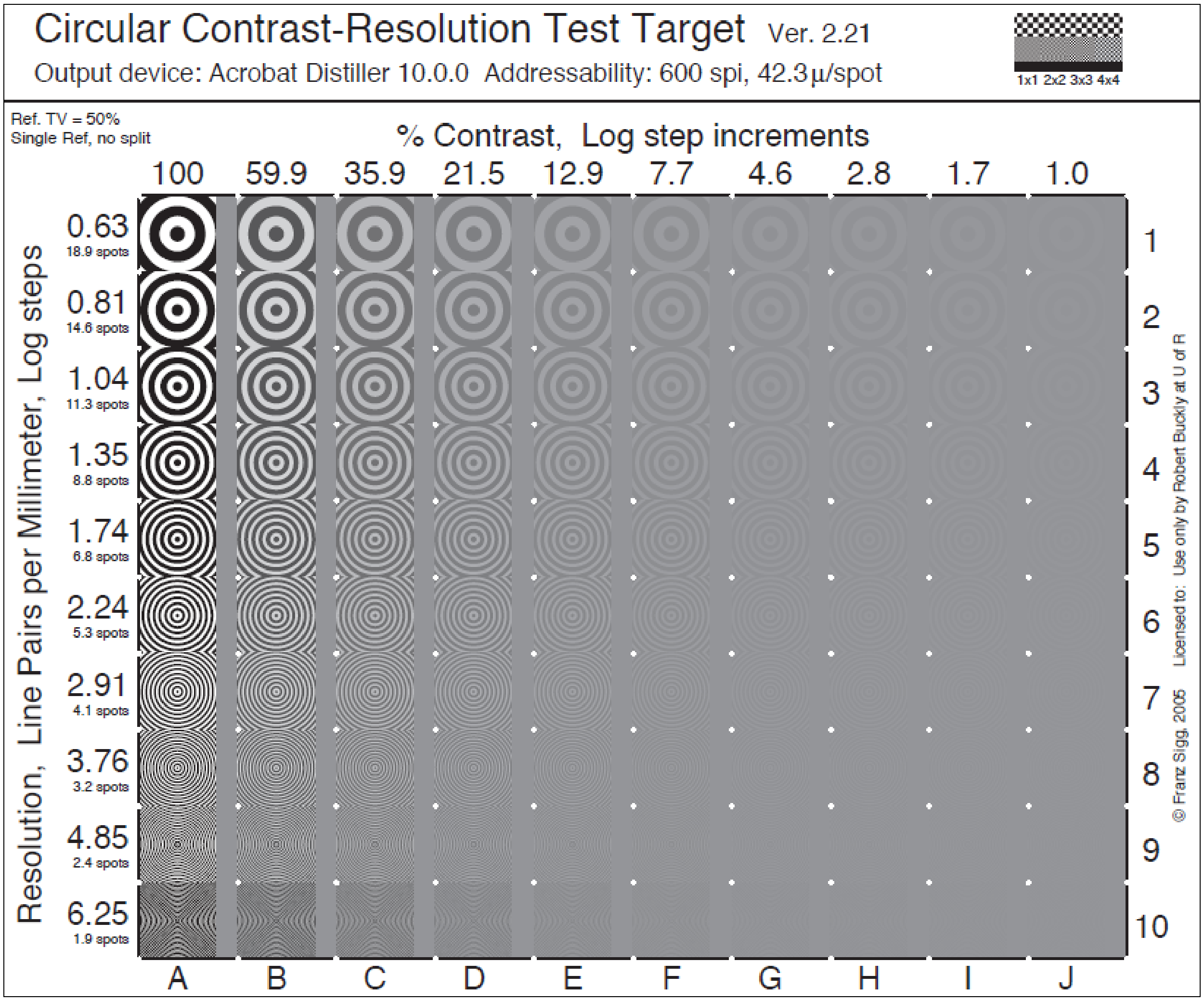


Figure 11. Contrast-Resolution Test Target

Figure 11 shows the Contrast-Resolution Test Target that Franz Sigg of RIT designed and created for assessing the contrast-resolution characteristics of a printer [16]. The target contains 10 rows and 10 columns of concentric circles of varying size and contrast. Using concentric circles means that edges are available at all orientations. The circles in each column have the same contrast ratio, starting with 100% in the left hand column and decreasing in equal-sized log steps to 1% in the right-hand column. The average tone value of the concentric circles is 50%, which is also the tone value of the gray strips separating the columns. The circles in each row have the same resolution, starting with 0.63 line pairs per mm (about 32 dpi) in the top row and increasing in equal log steps to 6.25 line pairs per mm in the bottom row.

When the target resolution exceeds the device resolution, the circles can no longer be resolved and the Moiré patterns like those in the lower left corner of the target become visible. Where the SFR of the device falls off, low-contrast, high-frequency detail cannot be reproduced and the circles blend into uniform gray patches like those in the lower right corner of the target. This is also the region where transform coders like JPEG and JPEG 2000 operate, coarsely quantizing the high spatial frequency content of the image and eliminating low contrast, high frequency edges.

What makes this significant is that in setting thresholds for the visibility of compression artifacts, it is important to know if artifacts are not visible because the human visual system cannot detect them or because the display can't resolve them. Other than reporting what display was used in testing, the display and its characteristics are not mentioned in reports of the testing of compression settings. This means that the results, besides being image dependent, may also be display dependent. It is worth testing differences between displays (and viewing conditions) to see how great a source of variability they may be in the results reported by different institutions and in different tests.

Finally, the capture device can also be a source of variability in the image path of Figure 10. A focus on compression as a lossy or noisy process should be accompanied by a recognition that capture is also a noisy process.

Conclusions

As a rule, the amount of compression that is acceptable depends on the task, the image and the compression method. This report has reviewed and presented results based on the visual assessment of digitized images from cultural heritage institutions that were compressed using JPEG 2000 to create visually lossless archival and production masters. In particular, it has reviewed and analyzed the different control settings and command line parameters that were used by various JPEG 2000 encoders to create master files.

The controls can be classified as specifying compression by size or by quality. "By size" controls are easier to share and replicate between implementations and institutions and they are the most widely used method for specifying compression. However, they cannot be generalized between images or across image classes because the relationship between compression ratio and quality is content dependent. The "by quality" slope control is related to the "by size" controls because of the mechanism that JPEG 2000 encoders use to obtain a compressed file with a specified size. The Google Books project uses the slope control in applications that focus on books and textual materials and where the size-quality tradeoff usually tilts toward size (and speed). Harvard Library uses a "by quality" PSNR control in a compression process that, once the PSNR value has been set, can run automatically, although it must check for the failure cases to which global metrics such as PSNR are vulnerable. The issue is using controls based on global metrics when visual task performance is based on local features.

The study has confirmed the image-dependent nature of the controls currently available and consequently the need for human observers to judge the control settings that are appropriate for an image type or collection. The criterion for an acceptable level of lossy compression has been that there is no functional difference between the images before and after compression, where "functional" translates to "visual "when the task is a comparison made by a human observer, as it is here. For the time being, there is simply no substitute for a human observer when the criterion for an acceptable level of compression is no visible or significant difference between original and compressed images.

While the control settings to meet the needs of the task are image dependent, they can and have been applied to images of the same general type or within the same collection. Until an image type or class is more closely defined and circumscribed for the purposes of compression, the controls have been applied conservatively. Using quality layers in the compressed data does not affect the overall quality while providing options for reduced quality selections that can satisfy other applications or criteria.

An alternative to an image-dependent observer-based control strategy would reduce the reliance on observers. This would enable the process for determining the threshold of acceptable compression to be automated and to scale more gracefully and less expensively as collections and the demand for compressed masters grow. The place to look for an alternative to observers is a visual discrimination metric that can mimic how human observers detect compression artifacts and determine near-threshold differences between images. A general metric would be costly to develop. Fortunately the metric for this application can be constrained by JPEG 2000 compression artifacts, which holds the promise of lower development costs and superior performance through tuning.

While the image dependence of the compression thresholds is unavoidable, it can be better understood and characterized through the use of a synthetic target. Like the metric, the target can be developed with JPEG 2000 compression in mind. In addition, a target that incorporated elements of the target in Figure 11 would also make it possible to assess the display dependence of compression thresholds.

Adding an education component to these developments offers the benefit of increasing confidence in the use of lossy JPEG 2000 compression for archival masters. This would be an important addition to the cost benefits of compression, which are already well understood.

References

[1] Federal Agencies Digitization Guideline, http://www.digitizationguidelines.gov/glossary.php

[2] Robert Buckley, "Error Correction and Recovery of JPEG 2000 Compressed Images in Digital Preservation," Proc. 9th IS&T Archiving Conference, Copenhagen, pp. 270-274, June 2012

[3] Robert Buckley and Roger Sam, "JPEG Profile for the National Digital Newspaper Program," available at http://www.loc.gov/ndnp/guidelines/docs/NDNP_JP2HistNewsProfile.pdf

[4] Robert Buckley, William Stumbo and Jim Reid, "The Use of JPEG2000 in the Information Packages of the OAIS Reference Model," Proc. IS&T 2007 Archiving Conference, pp. 24-28, Arlington, VA, May 2007

[5] http://www.britishnewspaperarchive.co.uk/help/about

[6] Sean Martin and Robert Buckley, "British Library JPEG 2000 Profile for Bulk Digitisation," presentation at the Library of Congress JPEG 2000 Summit Workshop, available at http://www.digitizationguidelines.gov/still-image/documents/Martin.pdf

[7] Christy Henshaw, "Report from the Wellcome Library and the JP2K-UK Group," presentation at the Library of Congress JPEG 2000 Summit Workshop, available at http://www.digitizationguidelines.gov/still-image/documents/Henshaw.pdf

[8] Stephen Chapman, Laurent Duplouy, John Kunze, Stuart Blair, Stephen Abrams, Catherine Lupovici, Ann Jensen, and Dan Johnston, "Page Image Compression for Mass Digitization," Proc. IS&T 2007 Archiving Conference, pp. 37-42, Arlington, VA, May 2007

[9] David S. Taubman, "High performance scalable image compression with EBCOT," IEEE Transactions on Image Processing, vol. 9, no. 7, pp. 1158-1170, 2000

[10] Jeff Breidenbach, "JPEG 2000 and Google Books," presentation at the Library of Congress JPEG 2000 Summit Workshop, available at http://www.digitizationguidelines.gov/still-image/documents/Breidenbach.pdf

[11] The University of Michigan, University Library Digital Library Production Services, Digital Conversion Unit, http://www.hathitrust.org/documents/UMDigitizationSpecs20100827.pdf

[12] Bill Comstock, "Using PSNR Thresholds to Modulate the Degree of Lossy Compression in JPEG 2000 Files," presentation at the Library of Congress JPEG 2000 Summit Workshop, available at http://www.digitizationguidelines.gov/resources/jpeg2000.html

[13] Z. Wang and A. C. Bovik, "Mean squared error: love it or leave it? - A new look at signal fidelity measures," IEEE Signal Processing Magazine, vol. 26, no. 1, pp. 98-117, Jan. 2009

[14] Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," IEEE Transactions on Image Processing, vol. 13, no. 4, pp. 600-612, Apr. 2004

[15] Richard Clark, "JPEG 2000 standards in digital preservation," J. Phys.: Conf. Ser. 231 012015, 2010

[16] Franz Sigg, "Testing for Resolution and Contrast," pp. 1-5 in Test Targets 6.0, RIT School of Print Media, Nov. 2006