Comparative Study of JPEG 2000 Compression and Format Configuration on Image Quality

Lei He, Steven Puglia; Library of Congress; Washington, DC /USA

Abstract

In this paper we present the results of a study on the effects of JPEG 2000 compression and file format configuration on image quality. This study was conducted as part of an institutional reassessment of still image file format choices for master image files, and parallels other institutional efforts to evaluate JPEG 2000 image compression. This work extends the thinking on this topic by focusing on identifying visually lossless compression for different collections types, by relating subjective assessment to objective or metrics-based analyses of image quality, and by evaluating the effect of JPEG 2000 file configuration on image quality. Additional work was done to evaluate the effect of scanner sampling efficiency on lossy image compression, and an assessment of traditional JPEG compression was done for comparison.

1. Introduction

JPEG (Joint Photographic Experts Group) 2000 [1, 2] is the latest image compression standard developed by the International Standard Organization (ISO) and International Telecommunications Union (ITU-T), which complements the current JPEG standard with improved compression performance and new features such as scalability and editability. For example, JPEG 2000 provides both lossless and lossy compression in a single code stream, which can be transmitted progressively by resolution, quality, component, or location; JPEG 2000 allows random code-stream access and processing, i.e., the code streams offer several mechanisms to support spatial random access or region of interest access at varying degrees of granularity. In particular, Part 1 of the standard, ISO/IEC 15444-1, is the Core Coding System containing the features that all decoders must support, in order to be called JPEG 2000 compliant.

Figure 1 shows the main function diagram of JPEG 2000 algorithm. The preprocessing block consists of three steps: tiling, DC level shifting, and color components transformation. Tiling partitions the input image into rectangular, nonoverlapping tiles of equal size (except edges). Color components of each tile can then be transformed from the RGB space to YCbCr (irreversible transformation) or modified YUV (reversible transformation) spaces. If the pixel values are unsigned integers and represented by B bits, an offset of -2^{B-1} is added to convert the intensity values to a range of $[-2^{B-1}, \ 2^{B-1}]$. The second block implements wavelet transform using either the CDF 9/7 wavelets (irreversible) or CDF 5/3 wavelets (reversible) on the intensity values. The Normalization/Quantization block maps the wavelet coefficients to integers for bit-by-bit encoding. The greater the quantization step size, the greater is the compression and the loss of quality. The last block of the encoder implements entropy coding, Embedded Block Coding with Optimal Truncation (EBCOT), which separates the coefficients into bit-planes and encodes each plane in three passes: significance propagation, magnitude refinement, and cleanup passes. Each of these coding passes collects contextual information about the bit-plane data. Such information along with the bit-planes is used by the arithmetic encoder (MQ-coder) to generate the compressed bit-stream. At last, four types of progression (for transmission), i.e., layer, resolution, spatial position, and component, can be achieved by appropriate order of the packets (spatially consistent code blocks) in the code stream. The decoder functions are the inverse of the corresponding encoder blocks.

This paper is organized as follows: Section 2 introduces our experiment plan. Section 3 presents the experimental results, including both subjective assessments and objective measurements of the JPEG 2000 compression effects. We summarize our observations and draw conclusions in Section 4.

2. The Study Plan

The goal of our study was to evaluate the compression effects of JPEG 2000 on image quality, and to develop guidance and recommendations for compression for different collection content to achieve the “visually lossless” quality. For example, with the identification of the “visually lossless” compression by domain experts, we may use JPEG 2000 to replace the commonly used TIFF to save the storage cost in archiving applications, and still preserve the “same” visual quality as in TIFF format. To implement the above goal, we conduct four phases of analysis in our compression experiments:

1. Analysis based on image content, mode, size, and resolution
2. Analysis based on imaging sampling efficiency

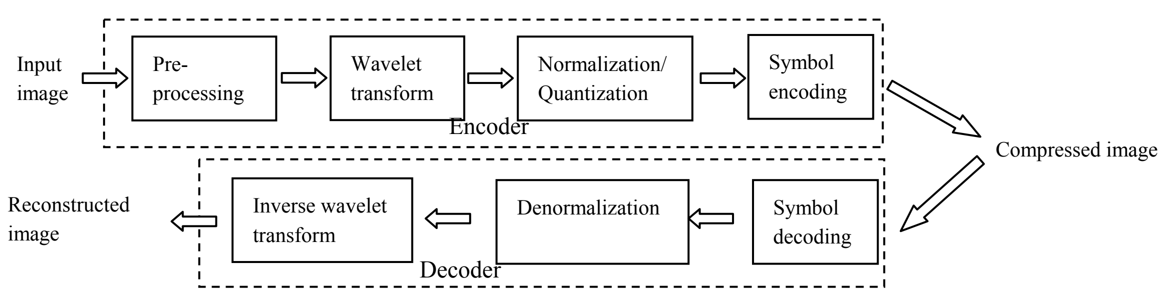

Figure 1. *JPEG 2000 function diagram*

3. Analysis based on JPEG 2000 parameter settings
4. Comparison with the standard JPEG compression

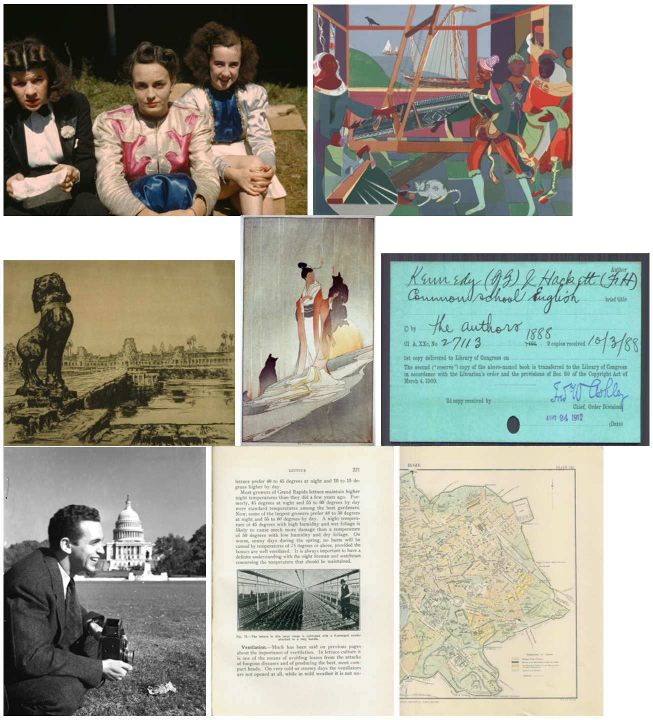

***Figure 2**. Representative image samples of different content, mode and size in our collection. From left to right and from top to bottom: color photo, graphic illustration, monochrome drawing, color drawing, copyright card, gray photo, printed text, and map.*

The first analysis was our main focus, in which we selected a representative sample of different content types, including printed text, photographs, graphic illustrations, maps, drawings, and documents with handwriting, as shown in Figure 2. Samples of different contents, modes (gray, monochrome, and color), size, and resolutions (300 and 400 dpi) are included in our collection, with the objective to investigate the effects of these factors on image compression. Digital images from different scanners were used for the initial phase of the study, and JPEG 2000 derivative images were produced using a sequence of 15 compression ratios applied to each image starting from lossless compression and progressing to 1024:1. Only the compression ratio was changed for this set of compressed images, all other JPEG 2000 configuration parameters (e.g., progression order, number of quality layers, number of decomposition levels, and tile size) were kept the same.

In response to the results of the first analysis, the second phase of the study used test images produced by image processing tools that simulate different scanner sampling efficiencies and followed the same protocol as the first phase. The third phase of the study focused on the effect of JPEG 2000 configuration parameters on image quality, by using the same set of sample images to generate additional derivatives under different configuration parameters for the JPEG 2000 files. Finally, a fourth phase was conducted to evaluate traditional JPEG compression (discrete cosine transform) for comparison purposes.

To evaluate the compression effects on image quality, we compared the compressed images with the original uncompressed images using both qualitative and quantitative measurement methods. For the qualitative or subjective assessment, observers reviewed the compressed images on the same workstation and were asked to identify the lowest compression ratio at which “artifacts” were visible. The average compression ratio identified for the group of observers and the ratio identified by the most critical observer (calculated as the 90^{th} psercentile) can be used to determine a “visually lossless” level for different content types. For the quantitative measurements, we calculated mean square error (MSE), peak signal noise ratio (PSNR), ΔE 2000 (for color images) [3], and structural similarity (SSIM) [4], using the original images as the reference. By mapping the subjective assessment to the objective measurements, we may estimate the compression ratio corresponding to the “visually lossless” level based on the objective error measurements. However, with only a single value calculated, these metrics cannot provide a detailed characterization of the compression effects. To address this issue, we propose a more general metric based on the error (difference) distribution with respect to the image intensity and gradient. These distribution-based measurements provide more comprehensive information about the compression effects on image quality.

3. Experiments

3.1 Visually Lossless Compression Identification

In the first phase of our study, a set of 42 samples of different content types, size, and modes were collected, e.g., 7 gray photos, 10 cartoon drawings (monochrome, gray, and color), 7 color photos, 4 graphic illustrations, 14 cards (copyright and musical cards). For each sample, a sequence of 15 compression ratios are applied to produce the compressed images, i.e., lossless, 8:1, 12:1, 16:1, 20:1, 24:1, 32:1, 40:1, 60:1, 80:1, 108:1, 168:1, 336:1, 600:1, and 1024:1. Specifically, there were two sets of musical cards collected at 300 and 400 dpi, respectively, with each set consisting of 5 samples. All the compressed images were named sequentially and uploaded to a computer for visual inspection. We recruited six domain experts to determine visually the lowest compression level at which they first observed compression artifacts, e.g., edge ringing and background smoothing; all evaluations were done on the same computer and monitor, in a darkened viewing environment. Meanwhile, objective error measurements (MSE, PSNR, and SSIM) were computed for the different compression ratios. Figure 3 shows the error curves of five representative samples at different compression ratios. As expected, the larger the compression ratio, the higher the error values. In addition to the commonly used single value metrics (MSE, PSNR, ΔE 2000, and SSIM), we also propose a new distribution-based measurement for more comprehensive and accurate error characterization. We illustrate the error (e.g., MSE) with respect to the image intensity and gradient, which show the distribution of error at different brightness and image features (edges). Besides these commonly used generic metrics, other more advanced image quality metrics [6, 7] specifically designed for JPEG 2000 may also be adopted for this measurement. Figure 4 shows the corresponding MSE curves for a grayscale photo sample, together with the intensity and gradient magnitude histograms. It can be seen that the intensity error is inverse to the number of pixels at that intensity level, which is also true for all other samples in our experiments. Table 1 lists the subjective assessment of "visually lossless" compression ratios on five representative samples. Lastly, in order to show the compression effects on the pixel values, we calculated and plotted ΔE 2000 curves with respect to the compression ratios on a color photo sample, as shown in Figure 5.

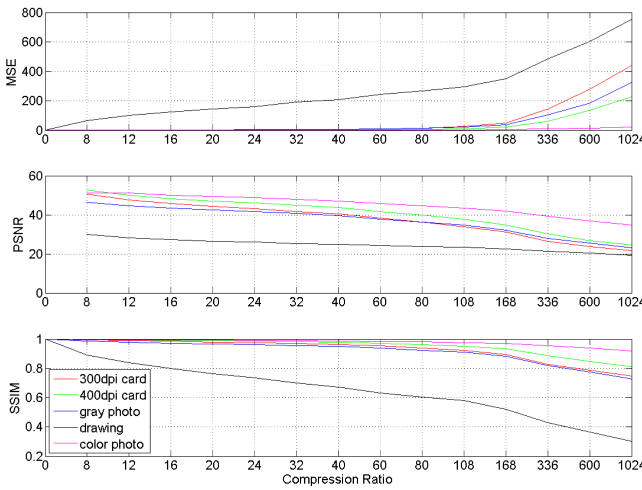

***Figure 3**. Objective error measurements on five samples.*

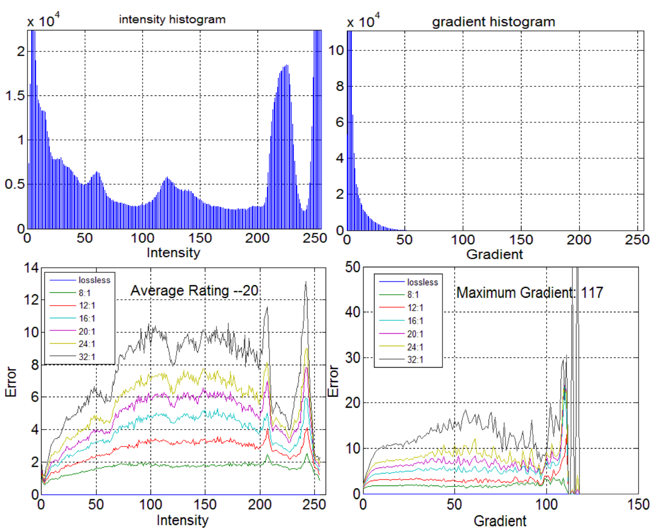

***Figure 4**. Error (MSE) distribution with respective to intensity and gradient.*

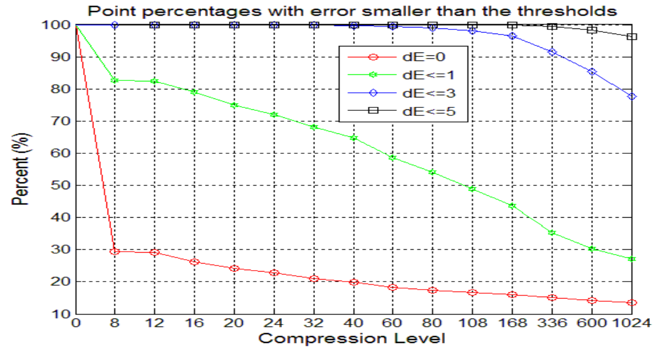

***Figure 5**. Percentage of points with error smaller than the thresholds.*

With the above results, we summarize our observations as follows:

- Image size is not necessarily related to the objective error measurements. However, some people identified the artifacts earlier on smaller size images, i.e., at lower compression levels on smaller size images. This may be due to the software used to view the images, which by default zooms large images to fill the screen, i.e., the noticeability of compression artifacts is reduced due to resizing to fit large images on the monitor.
- Color images usually have smaller error variation across different compression ratios than other contents. Compared with grayscale photos, generally observers identified artifacts at higher compression ratios on color images, i.e., it is more difficult to identify the compression artifacts in color photo.
- Image content affects human visual perception on artifact identification. With the objective error measurements, on average color photos usually have less error than musical cards (text), whose error is less than those of gray photos and copyright cards (text). The color/monochrome cartoon prints/drawings usually had the highest error. The subjective assessments also follow such pattern, see Figure 3 for an example.
- From the curves of the objective error with respect to the image intensity and gradient, we observed no specific correlation between the error and the intensity levels. On the other hand, large error appears at high gradient magnitudes, i.e., image edges/textures usually have larger error than smooth regions, see the rapid error increasing shown in Figure 4, especially at high compression ratios.
- Even minimally lossy compression (e.g., 8:1) introduces error (ΔE curves in Figure 5) to most pixels (e.g., 60%). For Just Noticeable Difference, ΔE = 1, our minimum lossy compression resulted in errors for 10% - 80% pixels for different image content types. The effects on different

Table 1: JPEG 2000 subjective assessment of "lossless compression" (# of people)

Samples	Average ratios	8:1	12:1	16:1	20:1	24:1	32:1	40:1	60:1	80:1	108:1	168:1	336:1	600:1	1024:1
300dpi card	32		1		1			4							
400dpi card	41		1	1				1	3						
Gray photo	20	1		2		2	1								
Drawing	11	2	3	1											
Color photo	71					2				2	2				

categories are different, e.g., color images have many fewer pixels affected than those in fine prints and glass negatives, which are less than those in cartoon drawings.

- In practice, $\Delta E = 5$ is used to distinguish one color from the next in standard images [4]. Except for the cartoon drawings, all image categories had a high tolerance to moderate compressions (e.g., up to 16:1 in our experiment), with more than 90% pixels having a $\Delta E < 5$.
- Lossless compression (about 2:1) has no effect on objective measurements.
- Higher resolution (e.g., 400ppi) images always have smaller error at different compression ratios, thus people usually identify the artifacts at lower compression ratios on low resolution (e.g., 300ppi) images than those of high resolution images.

With our observations, we recommend the ratios for "visually lossless" compression for different image content types. Here we show two sets of ratios based on the average and 90th percentile ratings of the domain experts, as shown in Table 2.

Table 2: Recommended ratios for "visually lossless" compression

	Mean	90th percentile
Grayscale photo	21	8
Color photo	46	12
Drawings	20	8
Graphics	42	12
Cards	41	16

3.2 Imaging Sampling Efficiency Effects

This experiment analyzes the effects imaging sampling efficiency on JPEG 2000 compression. We scanned 14 samples of different content types at 300dpi and 400dpi, respectively. For each resolution setting, we produced images at three different sampling efficiencies for both grayscale and color modes using image processing: 62%, 85%, and 105% for 300dpi set; and 65%, 85%, 101% for 400dpi set. Again, for each sample collected with a particular sampling efficiency, we compress it with the same 15 ratios as in the first analysis. Thus in total we have 14 × 2 (mode) × 2 (resolution) × 3 (efficiency) = 168 sets of images.

The same analysis process as in our first study was applied to the above image sets, i.e., we collect both human subjective assessment for "visually lossless" compression ratio identification and objective error measurements at different ratios. Here are our observations based on the results:

- On average, higher sampling efficiency leads to larger objective error, thus people identify compression artifacts at lower compression level.
- Other observations regarding image size, content, mode, and resolution are the same as our previous study, e.g., image size is not related to the visual assessment and objective error measurements; with smaller error, it is more difficult to identify the compression artifacts for higher resolution images; color images usually have smaller error than gray photos with the same compression ratios, thus people generally identify the artifacts at lower compression levels on grayscale images; human visual perception also depends on image content, as in the first analysis; objective error measurement is independent to the image intensity, and increases as the image gradient does; small compression introduces error to most pixels, but the compression error is small even under a high compression ratio, e.g., 20:1.

We propose "visually lossless" compression ratios with respect to different sampling efficiencies. For high (101% and 105%), middle (85%), and low (62% and 65%) sampling efficiencies, we list the average and 90th percentile expert ratings in Table 3, 4 and 5 that are organized by different categories (content type, information type, and scanning resolution).

Table 3: Recommended ratios for "visually lossless" compression with respect to the sampling efficiency and content type

Content Type	Encoding	Samp Eff	Mean	90th Percen	StDev
Book	Color	1.01 to 1.05	46	18	53
		0.85	44	20	31
		0.62 to 0.65	55	24	29
	Grayscale	1.01 to 1.05	25	10	29
		0.85	25	10	22
		0.62 to 0.65	32	12	29
Graphic Illustration	Color	1.01 to 1.05	41	12	41
		0.85	47	12	59
		0.62 to 0.65	70	16	108
	Grayscale	1.01 to 1.05	34	8	57
		0.85	32	8	35
		0.62 to 0.65	37	12	54
Index Card	Color	1.01 to 1.05	41	21	17
		0.85	41	21	17
		0.62 to 0.65	45	20	18
	Grayscale	1.01 to 1.05	21	16	6
		0.85	23	16	6
		0.62 to 0.65	28	16	13
Map	Color	1.01 to 1.05	28	13	18
		0.85	36	12	43
		0.62 to 0.65	46	13	43
	Grayscale	1.01 to 1.05	18	8	20
		0.85	19	8	12
		0.62 to 0.65	25	8	22
Photo	Color	1.01 to 1.05	42	16	46
		0.85	41	16	44
		0.62 to 0.65	48	16	44
	Grayscale	1.01 to 1.05	33	12	46
		0.85	32	12	29
		0.62 to 0.65	43	16	57

Table 4: Recommended ratios for "visually lossless" compression with respect to the sampling efficiency and information type

Info Type	Encoding	Samp Eff	Mean	90th Percen	StDev
Continuous Tone	Color	1.01 to 1.05	32	16	18
		0.85	32	14	17
		0.62 to 0.65	39	16	20
	Grayscale	1.01 to 1.05	23	12	13
		0.85	24	12	11
		0.62 to 0.65	32	14	14
Continuous Tone Color	Color	1.01 to 1.05	57	16	68
		0.85	55	16	64
		0.62 to 0.65	62	16	63
	Grayscale	1.01 to 1.05	49	13	69
		0.85	45	17	41
		0.62 to 0.65	59	18	86

Line	Color	1.01 to 1.05	38	12	34
		0.85	34	13	29
		0.62 to 0.65	47	17	45
	Grayscale	1.01 to 1.05	25	8	30
		0.85	28	8	32
		0.62 to 0.65	26	12	19
Line, Shading, and Color	Color	1.01 to 1.05	38	12	38
		0.85	48	12	62
		0.62 to 0.65	69	14	108
	Grayscale	1.01 to 1.05	32	8	56
		0.85	29	8	31
		0.62 to 0.65	36	10	55
Text	Color	1.01 to 1.05	47	20	47
		0.85	45	23	27
		0.62 to 0.65	53	23	25
	Grayscale	1.01 to 1.05	25	12	24
		0.85	26	12	19
		0.62 to 0.65	32	16	25
Text, with Halftone	Color	1.01 to 1.05	30	17	12
		0.85	33	16	19
		0.62 to 0.65	44	24	26
	Grayscale	1.01 to 1.05	19	8	14
		0.85	19	8	11
		0.62 to 0.65	23	12	15

Table 5: Recommended ratios for "visually lossless" compression with respect to the sampling efficiency and scanning resolution

Scanning Resolution	Encoding	Samp Eff	Mean	90th Percentile	StDev
300ppi	Color	1.05	38	16	28
		0.85	46	16	53
		0.62	48	16	40
	Grayscale	1.05	23	9	22
		0.85	29	8	25
		0.62	33	12	37
400ppi	Color	1.01	45	16	53
		0.85	39	16	27
		0.65	59	17	74
	Grayscale	1.01	35	8	53
		0.85	28	12	28
		0.65	39	13	51

3.3 JPEG 2000 Configuration Parameter Effects

This experiment tested the effects of JPEG 2000 parameter settings on the compression performance. We used the Matlab™ *imwrite* function to set the parameters. We chose 12 samples from each category of grayscale photo, color photo (color and monochrome), hand drawing (color and monochrome), and copyright cards. Configuration parameters include CompressionRatio, ProgressionOrder ('LRCP', 'RLCP', 'RPCL', 'PCRL' or 'CPRL'), QualityLayers (max 20, default 1), ReductionLevels (i.e., decomposition levels, max 8), and TileSize (mininum [128 128], default image size).

To test the effect of a particular parameter, we fixed the compression ratio at different levels and changed the settings of that parameter. We chose three compression ratios of 8:1, 40:1, and 600:1. Table 6 shows the MSE measurements for one color photo sample. The progression order determines the priority of coefficients (layer, resolution, component or position) in transmission and reconstruction, thus it does not change the objective error measurements. The order selection is application-specific. The parameter of quality layers indicates the truncated bitrate for each pixel. In most samples, different layers produce the same or very close results with similar objective measurements, especially at low compression ratios. At high compression ratios, a small number of layers (e.g., 5) is better than a larger number (e.g., 10 or 20) based on the objective error measure. Given a fixed compression ratio, the resulting image size is fixed. Therefore, more quality layers require more bits to be assigned to the layer headers, which result in more truncation on bits assigned to real data. On the other hand, more quality layers provide flexibility in delivering image regions of different quality to the applications. Reduction level specifies the decomposition level to conduct the discrete wavelet transform. At low compression ratios, different reduction levels also produce very close results. However, for most image samples with high compression ratios, a small reduction level (e.g., 2) has worse error measure than that of a high reduction level (e.g., 4 and 8). This is because different frequency components of the signal become well separated at high decomposition levels, which enables less detail (high frequency contents) loss in reconstruction. Lastly, the tile size parameter indicates the number of blocks to divide the original image. For all our samples, small tile size (128 × 128) results in larger error. When the size is large enough, e.g., 4096 × 4096 in our test, the results are the same as the original image size (i.e., no tiles). The more tiles (blocks) used, bit-errors will be restricted to individual blocks, which prevents error propagation in transmission. Moreover, small tiles are useful for reducing local memory buffering requirements to implement the wavelet transform, which is important for inexpensive hardware applications such as digital cameras. However, a large number of tiles usually introduces block effect with larger compression error. Smaller tiles also reduce the number of decomposition levels in the transform and this forces smaller code blocks to be used in the subbands that are smaller than the desired code block size (64 × 64).

3.4 JPEG 2000 vs. JPEG

This experiment showed the superiority of JPEG 2000 over standard JPEG. We choose six samples from the first analysis and used Adobe Photoshop™ to produce JPG files (by changing the quality and blur parameters) with similar size to the compressed JP2 files. We repeated the subjective analysis and computed the

Table 6: JPEG 2000 configuration parameter settings of a color photo example (MSE)

Compression ratio	Progression order	Quality layers			Reduction levels			Tile size			
	LRCP, RLCP, RPCL, PCRL, CPRL	5	10	20	2	4	8	128	512	1024	4096
8:1	1.26	1.27	1.27	1.27	1.24	1.22	1.27	1.51	1.27	1.27	1.24
40:1	2.87	2.88	2.88	2.88	3.36	2.84	2.89	3.41	2.91	2.89	2.87
600:1	14.36	14.50	14.56	14.61	199.91	17.37	14.39	81.75	16.77	15.10	14.42

objective measurements for these JPG files. As expected, people identified compression artifacts at lower compression ratios than those of JPEG 2000. Meanwhile, the objective error measurements of JPG files are larger than the corresponding JP2 files, i.e., those with similar image size.

4. Conclusion

We analyzed JPEG 2000 compression effects on image quality, with both subjective assessments and objective error measurements. We implemented four phases of study to analyze the effects of different factors in the compression, including image content, size, resolution, mode, sampling efficiency, and the JPEG 2000 configuration parameters (i.e., progression order, quality layers, reduction levels, and tile size). With these preliminary studies, we summarize our observations and recommend compression ratios for "visually lossless" compression based on collection content and sampling efficiency. For future work, we will recruit more imaging experts to assist with work on identifying "visually lossless" compression ratios in an automated manner; it is critical to construct a statistically valid model relating subjective assessment to objective error measurements. Thus we can automate image quality assessment, i.e., calculating an image error measurement and creating a predictive model for whether a level of lossy compression is acceptable or not.

References

[1] David Taubman, Michael Marcellin, JPEG 2000: Image Compression Fundamentals, Standards and Practice (Springer, 2002).

[2] Peter Schelkens, Athanassios Skodras, Touradj Ebrahimi, The JPEG 2000 Suite (Wiley, 2009).

[3] M.R. Luo, G. Cui, B. Rigg, "The Development of the CIE 2000 Colour-Difference Formula: CIEDE2000," Color Research and Application, 26, pg. 340-350 (2001).

[4] Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity," IEEE Transactions on Image Processing, 13, pg. 600-612 (2004).

[5] J. Kriesel, N. Gat, Performance Tests of True Color Night Vision Cameras, Military Sensing Symposia (MSS) Specialty Group on Passive Sensors. (2006).

[6] P. Marziliano, F. Dufaux, S. Winkler, T. Ebrahimi, "Perceptual Blur and Ringing Metrics: Application to JPEG2000," Signal Processing: Image Communication, 19, pg. 163-172 (2004).

[7] H.R. Sheikh, A.C. Bovik, L. Cormack, "No-Reference Quality Assessment Using Natural Scene Statistics: JPEG2000," IEEE Transactions on Image Processing, 14, pg. 1918-1927 (2005).

Author Biography

Lei He received his PhD in electrical engineering from University of Cincinnati (2003). Since then he has worked as a faculty member in the Department of Computer Science and Information Technology at Armstrong Atlantic State University in Savannah, GA. From 2009 to 2011, he visited the National Institutes of Health. Currently he is a digital imaging scientist at the Library of Congress. His work has focused on the image processing, computer vision and machine learning.

Steven Puglia is a Manager Digital Conversion Services at the Library of Congress. Previously he worked as a Preservation and Imaging Specialist at the US National Archives and Records Administration for over 22 years. He coordinates the Still Imaging Working Group of the Federal Agencies Digitization Guidelines Initiative.