ISUALLY LOSSLESS CODING FOR COLOR AERIAL IMAGES USING PEG

Graduate Students: Han O [1] and Yookyung Kim[1]
Advisors: Mic ael W. Marcellin[1] and Ali ilgin[2,1]
[1]Department of Electrical and Computer Engineering, [2] iomedical Engineering
T e University of Ari ona

A STRACT

This paper describes a psychophysical experiment to measure visibility thresholds (VT) for quantization distortion in JPEG2000 and an associated quantization algorithm for visually lossless coding of color aerial images. The visibility thresholds are obtained from a quantization distortion model based on the statistical characteristics of wavelet coefficients and the deadzone quantizer of JPEG2000, and the resulting visibility thresholds are presented for the luminance component (Y) and two chrominance components (Cb and Cr). Using the thresholds, we have achieved visually lossless coding for 24-bit color aerial images at an average bitrate of 4.17 bits/pixels, which is approximately 30% of the bitrate required for numerically lossless coding.

Keywords: visually lossless coding, JPEG2000, color aerial image, perceptual image coding

1. INTRODUCTION

The last few decades have seen tremendous increases in the size of aerial images due to advances in sensor and imaging technology. For efficient storage and transmission of large images, image compression has become an essential part of many telemetry systems [1]. To achieve high compression ratios, images are encoded in a lossy fashion. In general, this is done based on Mean Squared Error (MSE) or Peak Signal to Noise Ratio (PSNR) distortion metrics. However, since these metrics do not directly consider the visual importance of data, visually important information is often lost during compression. Therefore, the human visual system (HVS) should be incorporated in a distortion metric for a consistent visual quality.

In JPEG2000 [2], an image which has three color components (red, green, and blue) is transformed to an image with one luminance component (Y) and two chrominance components (Cb and Cr) for improved compression performance. Then, a discrete wavelet transform (DWT) decomposes each component into several subbands which have different frequencies and orientations. The human eye has different sensitivities to different subbands. This phenomenon occurs because the primary visual cortex called V1 in the human eyes responds to visual stimuli differently according to location, frequency, and orientation [3]. Based on the sensitivities, quantization step sizes are determined. The variance of sensitivity by subband is represented by the contrast sensitivity function (CSF), or visibility threshold (VT), which is the inverse of the CSF. Watson *et al* measured VTs for individual wavelet subbands based on randomly generated uniform quantization distortion [4]. These VTs have been applied to many wavelet-based codecs and they have shown superior visual quality at the same bitrate, as compared to the conventional MSE/PSNR metrics. However, since their assumption of the uniform quantizer is different from the deadzone quantizer of JPEG2000, a direct use of these VTs in JPEG2000 may result in inaccurate bit allocation on subbands. In [5], Chandler *et al* determined VTs through actual quantization for wavelet coefficients obtained from a few natural images. These VTs provide more accuracy than Watson's VTs because they take into account spatial correlation between wavelet coefficients. However, since these VTs are obtained with only a limited number of images, its performance with very different images is not guaranteed. Moreover, these VTs also assume the uniform quantizer, and are only available for the luminance component, and so are not usable for color image coding.

In this paper, we describe a new method for measuring VTs and an associated quantization strategy for visually lossless coding of color aerial images. A visual distortion model is developed based on the statistical characteristics of wavelet coefficients and the dead-zone quantization of JPEG2000. VTs obtained from the model are image-adaptive and ensure visually lossless coding for color aerial images.

. UANTI ATION DISTORTION IN PEG

uantization is the major step in lossy compression where most compression is performed and artifacts are generated. uantization step sizes for visually lossless coding are found through psychovisual experiments on quantization distortion. Therefore, an accurate modeling of quantization distortion is important to find an appropriate set of quantization step sizes.

The quantization distortion in JPEG2000, d, is the difference of wavelet coefficients in the encoder and decoder generated by the deadzone quantizer and mid-point reconstruction. Since the wavelet coefficients in the LH, HL, and HH subbands are usually represented by the Laplacian distribution, the distribution of quantization distortion in the HL, LH, and HH subbands is modeled by PDF

$$f(d)=\begin{cases}\dfrac{1}{\sqrt{2}\sigma}e^{-\frac{\sqrt{2}\,d}{\sigma}}+\dfrac{1-p_1}{\Delta} & 0\le d\le\dfrac{\Delta}{2}\\ \dfrac{1}{\sqrt{2}\sigma}e^{-\frac{\sqrt{2}\,d}{\sigma}} & \dfrac{\Delta}{2}<d\le\Delta\\ 0 & \text{otherwise}\end{cases} \tag{1}$$

where $p_1=\int_{-\Delta}^{\Delta}1\Big/\sqrt{2}\sigma\, e^{-\frac{\sqrt{2}\,y}{\sigma}}dy=1-e^{-\sqrt{2}\Delta/\sigma}$. Wavelet coefficients in the deadzone $(-\Delta,\Delta)$ become quantization errors themselves since the deadzone quantizer quantizes these coefficients to zero (the first term in the first two lines of Eq. (1)). Wavelet coefficients not in the deadzone yield errors modeled as uniform distributed over the interval $(-\Delta/2,\Delta/2)$ (the second term of the first line in Eq. (1)). Figure 1 (c) shows the distribution of quantization distortion in the subbands. We can see that our distortion model using the deadzone quantizer is significantly different from the uniform quantization distortion model of previous studies illustrated by a dashed line in Figure 1 (c)

The distribution of wavelet coefficients in the LL subband approximately follows that of the original image, but we assume they have a uniform distribution with zero-mean for simplicity as illustrated in Figure 1 (c). The distribution of the quantization distortion of the LL band is then modeled by

$$f(d)=\begin{cases}\dfrac{1}{\sqrt{12}\sigma}+\dfrac{1-p_2}{\Delta} & 0\le d\le\dfrac{\Delta}{2}\\ \dfrac{1}{\sqrt{12}\sigma} & \dfrac{\Delta}{2}<d\le\Delta\\ 0 & \text{otherwise}\end{cases} \tag{2}$$

where $p_2=\Delta/\sqrt{3}\sigma$. Since the variance of the LL subband is much larger than that of the other subbands, the quantization distortion in the LL band is predominately in the interval $(-\Delta/2,\Delta/2)$, as shown in Figure 1 (d).

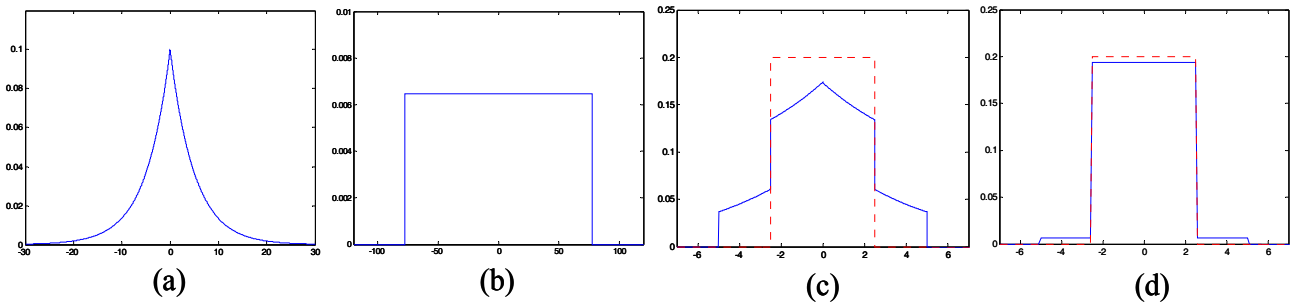


Figure 1. Distribution of wavelet coefficients in (a) HL, LH, HH bands ($\sigma^2 = 50$) and (b) LL band ($\sigma^2 = 2000$). uantization distortions for $\Delta = 5$ in (c) HL, LH, HH bands and (d) LL band.

. ISI ILITY THRESHOLDS FOR COLOR IMAGES

A color image having RGB components is converted into an image with one luminance component (Y) and two chrominance components (Cb and Cr) by the irreversible color transform (ICT)

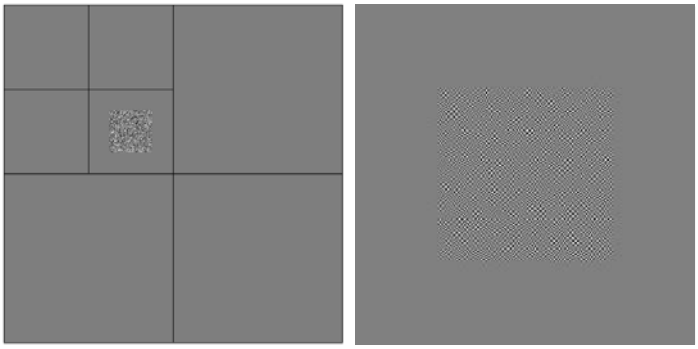
$$\begin{pmatrix} x_Y[\mathbf{n}] \\ x_{Cb}[\mathbf{n}] \\ x_{Cr}[\mathbf{n}] \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ -0.168736 & -0.331264 & 0.5 \\ 0.5 & -0.418688 & -0.081312 \end{pmatrix} \begin{pmatrix} x_R[\mathbf{n}] \\ x_G[\mathbf{n}] \\ x_B[\mathbf{n}] \end{pmatrix}. \quad (3)$$

Each component in the image is transformed by the dyadic 9/7 wavelet transform and then quantized independently. We measure visibility thresholds for quantization distortion in each subband and each component through psychovisual experiments.

A stimulus is an RGB image where the inverse wavelet transform and the inverse ICT are applied to wavelet data containing quantization distortion, as shown in Figure 2. uantization distortion is randomly generated in the subband of interest according to either Eq. (1) or Eq. (2), depending on the subband, while quantization distortions in other subbands are set to 0. In Figure 2, a stimulus appears at the center of an image of size 512×512 with a uniform gray background (Y=128, Cb=0 and Cr=0 for a 24-bit color image). In the wavelet domain, the size of the quantization distortion region is $N \times N$ with $N = \min 64, 512 \times 2^{-k}$ for wavelet transform level k, which corresponds to the size of a codeblock in JPEG2000.

To measure the visibility threshold for each stimulus, we use a three-alternative forced-choice (3AFC) method. Three images, one with a stimulus and two without, are displayed side by side for 10 seconds (the left to right order is chosen randomly) and the subject is asked to answer which image contains the stimulus. For each trial, the quantization step size is adaptively adjusted by a UEST staircase procedure [7]. Through 32 trials, the threshold is determined from the 75%

correct-point of a fitted Weibull function [4,6]. The images are displayed on a Dell E248WFP 24-in TFT LCD Monitor. The monitor has a resolution of 1920×1200 and provides a 92% color gamut. The viewing distance is 60 cm (23.6 inches) and the visual resolution is 35.62 pixels/degree. The center spatial frequency for the wavelet transform levels are then 17.81, 8.91, 4.45, 2.23, and 1.11 cycles/degree, respectively. Four subjects took part in the experiment. Two of the subjects estimated thresholds, and the other two verified these thresholds. All subjects are familiar with wavelet quantization distortion and image compression, and have normal visual acuity.



(a) (b)

Figure . A sample quantization distortion in HH2 (a) and the corresponding representation in the image domain (b).

.1. ISI ILITY THRESHOLDS FOR LUMINANCE COMPONENT

In [8], visibility thresholds were measured for the luminance component. Figure 3 shows measured visibility thresholds of two subjects for the HL, LH, and HH bands at a fixed variance of 50. Thresholds increased as the spatial frequency increases. In particular, the increase of threshold values in the HH band was more significant. Thresholds of the HL and LH bands were very similar, so we consider these bands as equivalent.

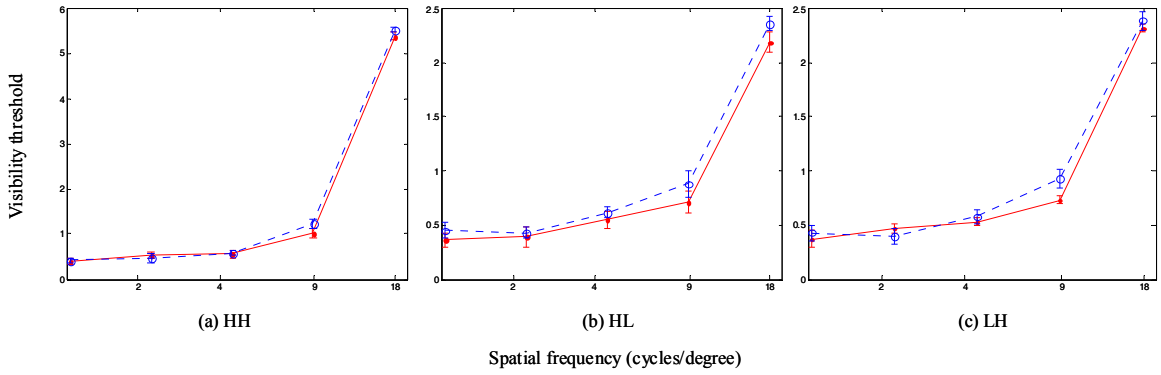


Figure . Visibility thresholds for the luminance component measured by two subjects ($\sigma^2 = 50$). The points are measured threshold values and the bars are standard deviations.

The distribution of quantization distortions modeled in Eq. (1) and Eq. (2) is a function of quantization step size Δ and variance σ^2 of wavelet coefficients. In general, the variance change for the luminance component is considerable. According to the variance, the threshold can be estimated. Figure 4 shows the change of threshold values as a function of variance on HL3/LH3. We can see that the threshold is well approximated linearly. Thus, the final threshold t_b for a given subband b is determined by

$$t_b = a_b \cdot \sigma_b^2 + c_b \tag{4}$$

where σ_b^2 is the variance of wavelet coefficients calculated in subband b. The linear parameters a_b and c_b are listed in Table 1. For the LL band, we use a fixed threshold value, 0.63 ± 0.05, since the variance of the LL band is usually much larger than that of other subbands and the distortion is less affected by the variance change.

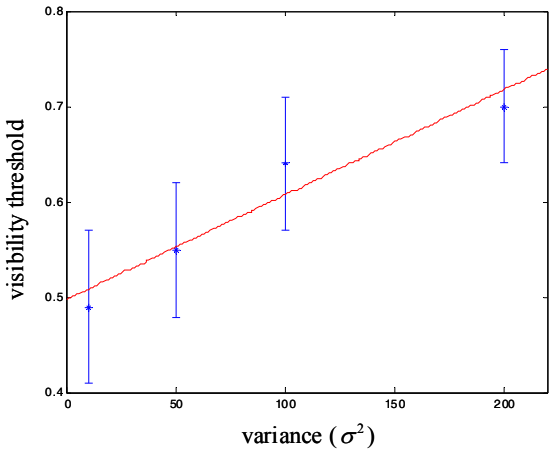


Figure . Linear elevation of visibility threshold in the HL3/LH3 bands as a function of variance.

subband	a_b	c_b	subband	a_b	c_b	subband	a_b	c_b
HH1	105.67×10^{-4}	4.85	HH3	11.04×10^{-4}	0.51	HH5	7.91×10^{-4}	0.36
HL1/LH1	46.03×10^{-4}	1.98	HL3/LH3	10.83×10^{-4}	0.50	HL5/LH5	7.16×10^{-4}	0.33
HH2	19.94×10^{-4}	0.92	HH4	10.16×10^{-4}	0.47			
HL2/LH2	13.84×10^{-4}	0.64	HL4/LH4	7.75×10^{-4}	0.36			

Table 1. Linear parameters a_b and c_b for visibility threshold of the luminance component.

. . ISI ILITY THRESHOLDS FOR CHROMINANCE COMPONETS

It is well known that the human visual system is less sensitive to quantization distortions in the chrominance components than in the luminance component. The amount of information in the chrominance components is also much less than that in the luminance, so the variance change is

not significant. We determined threshold values based on the average variances calculated from 10 natural images, as shown in Table 2.

subband	Cb	Cr	subband	Cb	Cr	subband	Cb	Cr
HH1	0.18	0.17	HH3	1.16	1.20	HH5	1.52	1.45
HL1/LH1	1.33	1.43	HL3/LH3	4.26	5.14	HL5/LH5	5.34	8.93
HH2	0.72	0.74	HH4	1.43	1.37	LL5	150.08	109.51
HL2/LH2	3.06	3.52	HL4/LH4	5.34	7.85			

Table . Average variances of wavelet coefficients for the chrominance components.

Table 3 represents measured threshold values for the chrominance components obtained through experiments, carried out in the same fashion as described above for the luminance component. From the table, we can see that threshold values of the chrominance components are larger than those of the luminance component as expected and the Cb component is the least sensitive component. Also, threshold values of the chrominance components are especially large at high spatial frequencies. This is because chrominance plays a relatively minor role in the perception of edges and fine details.

subband	Cb	Cr	subband	Cb	Cr	subband	Cb	Cr
HH1	24.72	15.49	HH3	11.55	2.57	HH5	1.05	0.56
HL1/LH1	14.50	6.35	HL3/LH3	4.03	1.23	HL5/LH5	1.05	0.58
HH2	14.77	7.40	HH4	4.95	1.25	LL5	1.32	0.66
HL2/LH2	6.36	2.60	HL4/LH4	3.26	0.69			

Table . Visibility thresholds for chrominance components.

. ISUALLY LOSSLESS CODING AND ALIDATION

The visibility threshold indicates the maximum allowable quantization step size such that quantization distortion remains invisible when one subband is quantized at a time. When all subbands are quantized simultaneously, quantization distortions can be summed up across subbands. This sum is called the compound distortion [6] and occurs because the 9/7 wavelet transform does not perfectly separate the signal into distinct frequency bands. However, compound distortions at a sub-threshold level are negligibly small and are well hidden by the background image. From this observation, we assume that there is no summation of distortions

across subbands and quantizing with these thresholds ensures visually lossless coding.

The three components are bit-plane encoded independently. For the case of the luminance component, the threshold is changed according to the variance of wavelet coefficients. In the distortion calculation stage, we calculate the variance $\sigma_{b,i}^2$ for codeblock i in subband b. Then we determine a threshold $t_{b,i}$ for the codeblock using Eq. (4). To exploit the three coding passes in bit-plane coding, we use Minkowski distance with $\beta = 0$ normalized by threshold $t_{b,i}$ as follows.

$$D^{(z)} = \frac{1}{t_{b,i}} \max_{\mathbf{n}} \left(\; y[\mathbf{n}] - y^{(z)}[\mathbf{n}] \; \right) \tag{5}$$

where $y^{(z)}[\mathbf{n}]$ denotes the reconstructed value from the quantization index $q^{(z)}[\mathbf{n}]$, which is encoded up to a truncation point z . $y[\mathbf{n}]$ is the reconstructed value from the high-precision original quantization index $q[\mathbf{n}]$. Coding is terminated for the codeblock immediately after $D^{(z)}$ is equal to or less than 1. For the chrominance components we directly use the visibility thresholds listed in Table 3 as quantization step sizes. Then all bit-planes are included in the JPEG2000 codestream.

The proposed coding scheme is implemented in Kakadu v6.1 [9] and tested with 10 512×512 24-bit color aerial images. These images are cropped from four 7200×5000 24-bit raw color aerial images provided by the Cartographic Institute of Catalonia (ICC) [10], which cover vegetation and urban areas. Figure 5 shows some example images. We compared the bitrates obtained using our visually lossless coding scheme with numerically lossless bitrates of conventional JPEG2000, as shown in Table 4. While the average bitrate for numerically lossless coding is 14.70 bits/pixel, the average bitrate for visually lossless coding is 4.17 bits/pixel at the same visual quality.

Figure 5. Examples of the test images.

To verify that the images encoded with the proposed coding scheme are truly visually lossless, we used another three-alternative forced-choice (3AFC) method. In this case, two of the images are original and one is the compressed image. The images are displayed for 10 seconds and the subject is asked to select one which looks different. Seven observers participated in this validation experiment. For each image, 70 trials were conducted and the resulting data is shown in Figure 6. When the compressed image is indistinguishable from the original, the correct response should be obtained 33.3% of the time. Our coding scheme results in correct selection of the compressed image 35.7% of the time with a 95% confidence interval of 29.48-37.15%. Thus, we claim that our scheme is visually lossless with this image set.

image	Numerically Lossless	Visually Lossless	image	Numerically Lossless	Visually Lossless
1	14.31	3.80	6	14.94	4.04
2	15.69	4.54	7	14.08	4.34
3	13.30	3.35	8	14.47	4.73
4	14.74	4.10	9	13.46	4.09
5	14.64	4.03	10	14.36	4.65

Table . Bitrate comparison for 24-bit color aerial images (bits/pixel).

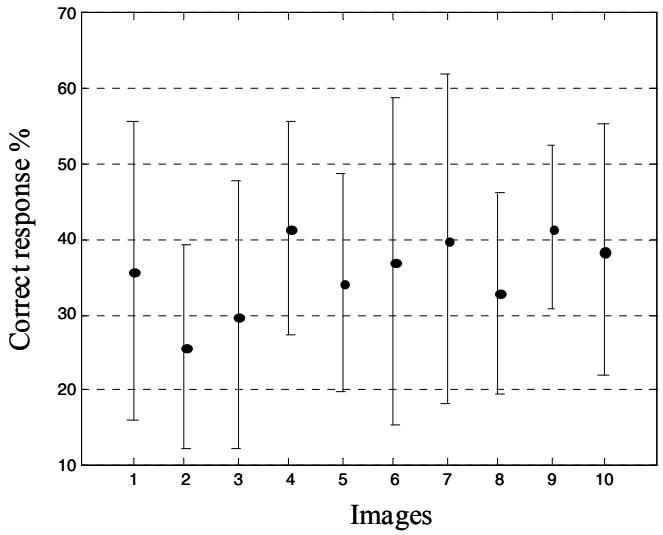


Figure . Rate of correct responses to the validation experiment for each test image. The points denote the average correct responses and the error bars give 95% confidence intervals.

5. CONCLUSIONS

We measured visibility thresholds for quantization distortion in JPEG2000. While the thresholds for the luminance component were significantly affected by the variance of wavelet coefficients, the thresholds for the chrominance components were less affected. Compared with numerically lossless coding, visually lossless coding using our visibility thresholds achieves a 70% bitrate saving on average without any visual quality degradation. Also, since this coding algorithm does not use the post-compression rate-distortion algorithm of JPEG2000, faster encoding will be possible without violating the JPEG2000 Part-I standard.

REFERENCES

[1] Yeh, P. S., Moury, G. and Armbruster, The CCSDS Data Compression Recommendations: Development and Status, Proc. SPIE Application of Digital Image Processing, Jul 2002.

[2] Taubman D. S. and Marcellin, M. W., JPEG2000: Image Compression Fundamentals, Standards and Practice, Kluwer Academic, 2002.

[3] Blakemore, C. and Campbell, On the existence of neurons in the human visual selectively sensitive to the orientation and size of retinal images, Journal of Physiology, vol. 203, pp. 237-260, 1969.

[4] Watson, A. B., Yang G. Y., Solomon J. A. and Villasenor J., Visibility of wavelet quantization noise, IEEE Trans. Image Processing, vol. 6, no. 8, pp. 1164-1175, Aug 1997.

[5] Liu, . and Karam, L. J., JPEG2000 Encoding with Perceptual Distortion Control, IEEE Trans. Image Processing, vol. 15, no 7, pp. 1763-1778, Jul 2006.

[6] Chandler, D. M. and Hemami, S. S., Effects of natural images on the detectability of simple and compound wavelet subband quantization distortions, Journal of the Optical Society of America A, vol. 20(7), pp. 1164-1180, 2003.

[7] Brainard, D. H., The psychophysics toolbox, Spatial Vision, vol. 10, pp. 433-436, 1997.

[8] Oh, H., Bilgin, A. and Marcellin, M. W., Visibility thresholds for quantization distortion in JPEG2000, Proc. of oMEX, Jul 2009.

[9] http://www.kakadusoftware.com

[10] http://www.icc.cat, Institut Cartografic de Catalunya, Barcelona, Spain, 2009.