

SUBJECTIVE IMAGE QUALITY TRADEOFFS BETWEEN SPATIAL RESOLUTION AND QUANTIZATION NOISE

Soo Hyun Bae, Thrasyvoulos N. Pappas[†], Biing-Hwang Juang

Center for Signal and Image Processing, Georgia Institute of Technology, Atlanta, GA 30332

[†]EECS Department, Northwestern University, Evanston, IL 60208

{soohyun, juang}@ece.gatech.edu, pappas@ece.northwestern.edu

ABSTRACT

The importance of tradeoffs between spatial resolution and quantization noise has been examined in our previous work. Subjective experiments indicate that as the bitrate decreases, human observers generally prefer to reduce image resolution in order to maintain image quality, but the amount of distortion they are willing to accept increases with decreasing resolution. In this paper, we conducted further experiments with several images, different encoders, and a finer set of bitrates to determine the preferred resolution at each bitrate, and also the resolution at which there are no visible coding artifacts. Analysis of the subjective results using a wavelet-based perceptual quality metric verifies our earlier conclusion that human observers tend to reduce resolution in order to maintain image quality, but are willing to accept more artifacts as image size decreases.

1. INTRODUCTION

The proliferation of display and capture devices with varying characteristics and spatiotemporal resolution necessitates a scalable approach to image/video communication. The spatiotemporal resolution of the signal should depend on the transmission bandwidth and display device of each user, and should be determined with the help of an objective measure of image quality that takes into account the visibility of both the compression artifacts and the image/video signal. To gain an understanding of the tradeoffs between spatial resolution and quantization noise, we conducted subjective experiments [1]. We found that as the bitrate decreases, human observers generally prefer to reduce image resolution in order to maintain image quality, but the amount of distortion they are willing to accept increases with decreasing resolution. In this paper, we conducted further experiments with several images, different encoders, and a finer set of bitrates to obtain more precise results across image resolutions and bitrates. Four test images are employed in these experiments, two of which contain complicated image details. For the JPEG encoding, we introduced the perceptually tuned visibility threshold for the discrete cosine transform (DCT) at six image heights, which was proposed by Watson [2]. Also, we carefully determined a wider and finer set of bitrates for image coders so as to avoid remarkable difference of perceived quality between each bitrate. In our previous work, the *absolute perceived quality* assessment was designed to obtain both numeric expression of the subjective image quality and the perceptually transparent noise level, then the highest level of subjective quality was compared to the most preferable spatial resolution. In this work, we designed a new experiment named the critical noise

perception assessment. Throughout the experiments, the subjects were asked to distinguish the original image from the coded one at various resolutions and bitrates for the four images. We precisely drew the critical bitrates at which the human eye cannot or can recognize compression artifacts, then compared it to result of the relative perceived quality assessment. For obvious understanding of the tendency of the most preferable resolutions, we selected an image quality metric different from the previous one, the wavelet-based metric by Watson *et al.* [3], which gives a numeric expression of image quality computed at the same spatial resolution. Several viewing conditions were also refined in the experimental environment.

Analysis of the subjective results using perceptual quality metrics verifies our earlier conclusion that human observers tend to reduce resolution in order to maintain image quality, but are willing to accept more artifacts as image size decreases. We are in progress of development of image quality metric incorporating both signal visibility and noise visibility.

2. IMAGE QUALITY METRICS

Media signal processing inevitably involves distortion on the signal. A measure that provides an evaluation of the incurred distortion finds many applications in compression, transmission, and enhancement. The measure is conventionally termed a quality metric or a quality measure and can be formulated within two extreme perspectives. On the one end of the spectrum is the subjective measure, in which the evaluation is accomplished through a process that reflects the human assessment. On the other end of the spectrum is the objective measure(s), which are customarily defined on the mean squared error between corresponding signals. In between, there are a number of hybrids that attempt to establish a measure, which can be computed from the signal directly and yet draw a very close approximation of the subjective result without any cumbersome procedure in administering the human assessment process.

It is important to note that these conventional measures are designed to quantify the error sensitivity between the original signal and the distorted one, while keeping most of the signal characteristics intact. For example, in image processing, conventional quality metrics are mostly defined over a squared difference between corresponding pixel values; the sampling rate of the image remains the same. Since the human visual system (HVS) involves perception along several dimensions (i.e. visual area, viewing angle, viewing distance, etc.), a new class of quality measures should involve the various perceptions as well in order to incorporate these



(a) Critical noise perception assessment



(b) Relative perceived quality assessment

Fig. 1. Test images presented to observer at each test. Bank coded by JPEG at 0.2 bits/pixel (a), Lena by JPEG2000 at 0.1 bits/pixel (b)

additional factors.

To obtain a better understanding of the displaying and viewing parameters with an ultimate goal of designing image quality metrics for scalable image coding applications, we conducted various subjective experiments upon tradeoffs between compression artifacts and spatial resolution. First, a series of compressed images at different bitrates, which are carefully chosen for covering wider perceptual quality than the previous results [1], are generated and then downsampled by optimal sinc-function. The subject tests are designed along two aspects: *critical noise perception assessment* and *relative perceived quality assessment*. The first aspect is measurement of the critical compression noise level at which human cannot or barely recognize compression artifact. The second one is measurement of the most preferable resolution.

3. SUBJECTIVE TEST SETUP

One of psychophysical experiments for analyzing the effect of spatial resolution in image quality assessment is [4], which formed a basis of modern image quality analysis. However, the specific tradeoffs we examine in this work were not addressed in the paper. Exploring tradeoffs between spatial resolution and image compression artifacts to obtain the perceptually optimal compression conditions at a given coding algorithm and a bitrate is our goal. As we discussed above, two subjective tests were designed as follows: The critical noise perception assessment aims at pointing the noise transparent bitrate at every image and its spatial resolution. We showed an observer two images, the original image and the decoded one, then asked to differentiate the original one. A combination of two images are randomly ordered when displayed as shown in Figure 1 (a). Basically if the answer is correct, it indicates that noise at a combination of given image, coding bitrate, and spatial resolution is visible to human eyes. Otherwise, it



(a) Bank (b) Lena (c) Bike (d) Woman

Fig. 2. Images for subjective tests.

Image	Coder	Bitrates (unit: bits/pixel)
Bank	JPEG	1.0, 0.8, 0.7, 0.6, 0.5, 0.45, 0.4, 0.3, 0.25, 0.2
	JPEG 2000	1.0, 0.75, 0.6, 0.5, 0.3, 0.2, 0.18, 0.15, 0.12, 0.1, 0.05
Bike	JPEG	1.0, 0.8, 0.6, 0.5, 0.4, 0.3, 0.27, 0.23
	JPEG 2000	1.0, 0.75, 0.5, 0.4, 0.3, 0.2, 0.1, 0.08, 0.05, 0.04, 0.03
Lena	JPEG	1.0, 0.5, 0.4, 0.35, 0.3, 0.27, 0.25, 0.23, 0.2
	JPEG 2000	1.0, 0.5, 0.4, 0.35, 0.3, 0.2, 0.1, 0.08, 0.05, 0.03
Woman	JPEG	1.0, 0.6, 0.5, 0.4, 0.35, 0.3, 0.25, 0.2, 0.15, 0.14
	JPEG 2000	1.0, 0.5, 0.3, 0.2, 0.1, 0.08, 0.06, 0.04

Table 1. Coding bitrates for the image coders

means that noise at given conditions is invisible to human eyes. If the two images cannot be distinguished or narrowly distinguished, the coding bitrate is equivalent to the critical bitrate of the given conditions. For validation of the measurement, each question is given ten times repeatedly. If they are correct over eight times out of ten repetitions, the current bitrate is finally set to *distinguishable* level and another test at higher bitrate are given. On the other hand, if they are incorrect over three times out of ten repetitions, the current bitrate is set to *indistinguishable* level. Although the desired level is reached, another test at lower bitrate is given for making sure that the previous critical level is guaranteed. The repetition process provides us an active assessment which can abridge a number of unnecessary measurements. Mid-level bitrate, which is initially specified, also helps the efficient assessment process.

In the relative perceived quality assessment, which is the main goal of this paper, we presented image in seven different spatial resolutions, and asked to make a choice of the most preferred image at a fixed bitrate in terms of overall image quality, i.e., including both distortion artifacts and image size. The test environment is in Figure 1 (b).

We used four test images: Lena, Bank, cropped Bike, cropped Woman, shown in Figure 2. For the JPEG encoding, we used a visibility model presented by Watson [2], which computes the visibility thresholds for the discrete cosine transform (DCT) coefficients at six image heights. For the JPEG2000 encoding, the number of DWT decomposition levels is set to 5, codeblock size is set to 32. The images are first compressed by JPEG or JPEG2000 at a number of bitrates which are previously selected as shown in Table 1. Since the coders have different coding efficiencies, the bitrate table was carefully designed in consideration of no remarkable difference of perceived quality between bitrates. For each coder and bitrate, the reconstructed images were then downsampled to obtain seven different resolutions: 512, 384, 256, 192, 128, 96 and 64, using sinc-function upsampling and downsampling in integer ratios. Three observers, rsmk, jwk and shb, took part in the experiments. All are corrected myope and binocular. The images were viewed in a darkened room on flat panel liquid-crystal display (LCD) screen with blue background. Observers' eyes are to the center of the display device. The viewing distance for all six images is fixed to six image heights of the 512×512 image, so that the viewing angle for the highest resolution image is $2 \arctan(1/12) \simeq 9.53^\circ$. In the actual test, the observers were

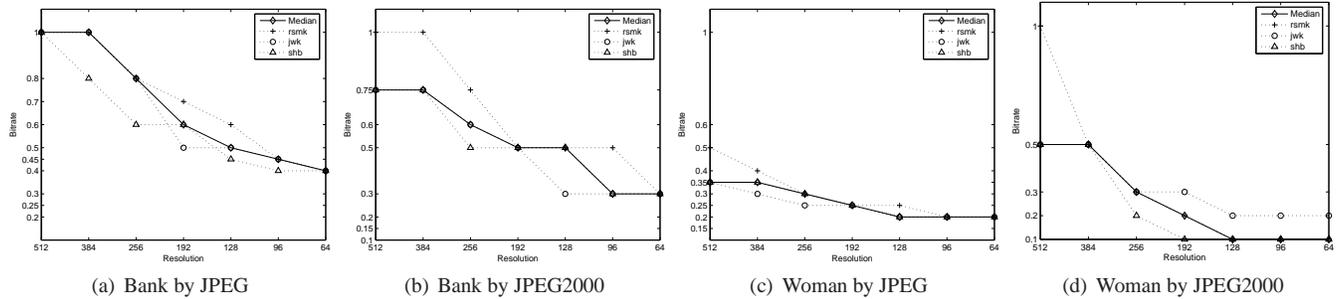


Fig. 3. Analysis of the result of the critical noise perception assessment. The solid lines are corresponding to the medians of all votes.

allowed enough time to make their decisions and to view the original test images before and during the test. The ordering of images and coders are randomized to avoid any biases, but the bitrates in the relative perceived quality assessment decreases at every test. Again, the bitrates in the critical noise perception assessment begin at the mid-level bitrates and remaining bitrates are actively determined according to the observers' answers in order to abridge a number of unnecessary measurements.

Quantitative expressions of image quality are computed between the originals and the decoded images at the same resolution by the Wavelet-based metric developed by Watson *et al.* [3]. The linear-phase 9/7 biorthogonal filters are used for signal decomposition, then the baseline sensitivity thresholds, $\tau_{i,k}$, for the wavelet decomposition were measured. Here k denotes the subband index and i the coefficient location in the subband image. The overall image distortion "Perceptual Masked Error (PME)" is then computed with

$$D_p = \left\{ \frac{1}{N} \sum_{i,k} \left| \frac{b_{i,k} - \hat{b}_{i,k}}{\tau_{i,k}} \right|^Q \right\}^{\frac{1}{Q}}, \quad (1)$$

where $b_{i,k}$ is the subband coefficient of the reference image, $\hat{b}_{i,k}$ is corresponding coefficient of the distorted image, and $\tau_{i,k}$ is the visibility threshold. Here we will use $Q = 2$. In order to become PME comparable to traditional error metrics, we define the "masked peak signal-to-noise ratio (MPSNR)" as

$$\text{MPSNR} = 10 \log_{10} \frac{255^2}{D_p^2}. \quad (2)$$

A detailed description of this metric can be found in [5].

4. EXPERIMENTAL RESULTS

For the analysis of the result of the critical noise perception assessment, the median of all votes were employed. The results of images, bank and woman, are given in Figure 3. At a given resolution, images coded over the bitrates above each median value are perceptually noise-transparent. Obviously, the images generated by JPEG2000 have lower critical bitrates than images by JPEG. Note that the bank image contains complicated details and the woman image has comparatively less details. The bank image thus has higher critical bitrates than woman image does. Therefore, in the woman coded by JPEG 2000, the images with resolutions lower than 128×128 are nearly noise transparent even at 0.1 bpp. For comparison of the result to an objective image quality, the critical bitrates are drawn (with shaded cells) in Table 2. Similar to the results in the subjective test, the objective quality decreases

as bitrate decreases or spatial resolution increases; conversely, it increases as bitrate increases or spatial resolution decreases. In particular, at low resolution, the bitrate, which is closely related to quantization level in the encoder, does not seriously affect the subjective and objective quality. It explains two important findings that people are able to accept more distortion for low resolution images and bitrate needs to be determined not only by a target quality but by other parameters, e.g. spatial resolution and viewing distance.

To analyze the results of the relative perceived quality assessment, the median values of the most preferable resolution are obtained. Figure 4 shows all the votes of observers and the median values for the image bike. On top of the human eye's basic preference to a higher resolution and less distorted image, Figure 4 also depicts their tradeoffs substantially.

Comparison of the two sets of tests leads us to another result that people tends to maintain perceptual quality at every spatial resolution as presented in bold numbers in Table 2. A tendency to maintain quality around 59.00 dB over various bitrates can be found in both (a) and (b) of Table 2. Comparison of the noise-transparent level and the most preferable resolution at a fixed resolution implies that people are willing to accept more distortion. For example, at 128×128 of Bank JPEG2000, the bitrates can be lowered from 0.5 bpp down to 0.18 bpp. Therefore, the level difference at a fixed resolution, i.e. the resolution at a noise transparent bitrate and most preferable resolutions, is equivalent to perceptual tolerance over noise-transparent condition. In this example, 3.69 dB of noise can be more added over the noise-transparent condition without sacrifice of perceptual quality.

5. CONCLUSION

Noise visibility of the compressed image over various spatial resolutions and bitrates in various types of images is studied for a framework of image quality metric. Since most of the image quality metric incorporate just the visibility of noise, not the visibility of signal itself, analysis of tradeoff between the spatial resolution and the quantization noise is highly necessary in the scalable image compression application.

We designed subjective tests along two aspects, the critical noise perception and the relative perceived quality to explore such tradeoffs. A series of compressed images at different bitrates, which are carefully determined for covering wide perceptual quality, is generated. Lower resolution images were then produced by optimal sinc-function upsampling and downsampling in integer ratios. The critical noise perception assessment is to obtain noise transparent bitrate at every image and spatial resolution. The rel-

(a) Bank JPEG2000

Resolution/bpp	1	0.75	0.6	0.5	0.3	0.2	0.18	0.15	0.12	0.1	0.05
512x512	60.97	59.88	58.60	58.22	56.31	55.53	55.17	54.90	54.69	54.56	53.78
384x384	61.45	60.36	59.11	58.61	56.89	55.82	55.31	55.00	54.57	54.29	53.13
256x256	63.79	62.92	61.36	60.75	58.60	57.85	56.92	56.49	56.01	55.67	54.23
192x192	64.26	63.46	61.91	61.25	59.31	58.42	57.55	57.16	56.38	55.99	54.26
128x128	65.72	65.62	64.08	63.00	60.82	60.22	59.31	58.80	57.56	57.52	55.26
96x96	66.33	66.10	64.77	63.67	61.50	61.00	60.20	59.71	58.20	58.02	55.74
64x64	67.94	67.80	67.34	65.30	62.85	62.69	62.28	62.07	59.90	59.83	57.10

(b) Lena JPEG

Resolution/bpp	1	0.5	0.4	0.35	0.3	0.27	0.25	0.23	0.2
512x512	61.87	58.87	57.94	57.42	56.74	56.31	56.01	55.65	55.04
384x384	63.35	59.85	58.82	58.23	57.43	56.91	56.54	56.15	55.39
256x256	66.49	62.56	61.22	60.49	59.54	58.89	58.47	57.85	57.02
192x192	67.11	63.30	61.89	61.09	60.00	59.27	58.74	58.20	57.21
128x128	68.64	64.92	63.54	62.74	61.57	60.76	60.28	59.55	58.58
96x96	68.84	65.47	63.98	63.13	61.96	61.06	60.65	59.83	58.79
64x64	69.57	66.95	65.71	64.97	63.84	62.71	62.29	61.47	60.32

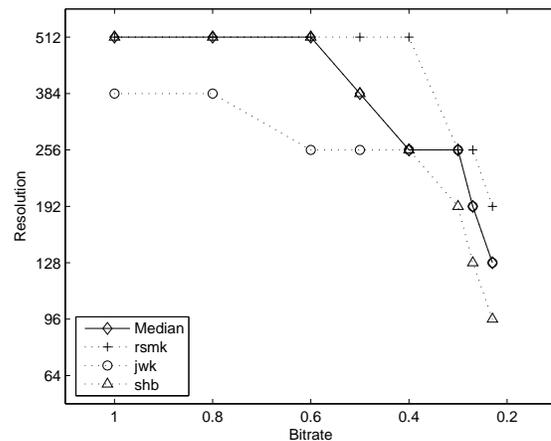
Table 2. MPSNR values over different coding rates and spatial resolutions. The critical bitrates are in shaded cells, and the most preferable resolutions are specified in bold numbers. (unit: decibels)

ative perceived quality assessment is to gain the most preferable resolution of each image with consideration of distortion artifacts and image size. Conducting the subjective tests with observers, we have found how image characteristic affects the critical bitrates, and the most preferable resolution at each noise level of every image.

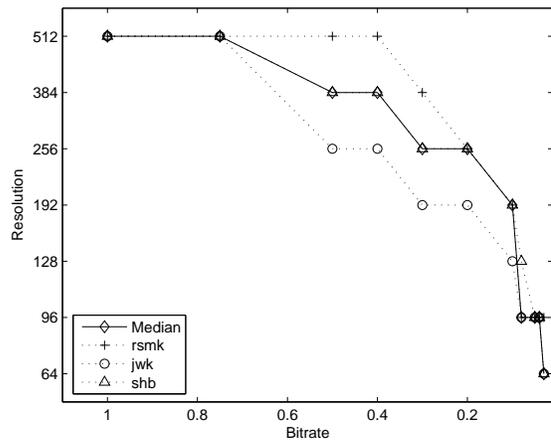
By the comparison of two different levels, a tendency that human eyes try to maintain the subjective quality as image size decreases is observed. We have also gained that when various viewing parameters are considered, there exists perceptual noise tolerance so that observers are willing to accept more artifacts as image size decreases. Estimation of the two results obtained by the subjective tests will provide a more important framework for image quality metric yielding objective perceptual quality over spatial resolution and quantization noise, which will be essential to the scalable image compression. We are in progress of developing the image quality metric incorporating signal visibility as well as noise visibility.

6. REFERENCES

- [1] S. H. Bae, T. N. Pappas, and B.-H. Juang, "Spatial resolution and quantization noise tradeoffs for scalable image compression," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, May 2006, vol. 1, pp. 945–948.
- [2] A. B. Watson, "DCT quantization matrices visually optimized for individual images," in *Proc. SPIE, Human Vision, Visual Proc., and Digital Display IV*, Jan P. Allebach and Bernice E. Rogowitz, Eds., 1993, vol. 1913.
- [3] A. B. Watson, G. Y. Yang, J. A. Solomon, and J. Villasenor, "Visibility of wavelet quantization noise," *IEEE Trans. on Image Processing*, vol. 6, no. 8, pp. 1164–1175, Aug. 1997.
- [4] J. H. D. M. Westerink and J. A. J. Roufs, "Subjective image quality as a function of viewing distance, resolution, and picture size," *SMPTE Journal*, vol. 98, pp. 113–119, Feb. 1989.
- [5] T. N. Pappas, R. J. Safranek, and J. Chen, "Perceptual criteria for image quality evaluation," in *Handbook of Image and Video Processing*, Alan C. Bovik, Ed., pp. 939–959. Academic Press, second edition, 2005.



(a) JPEG



(b) JPEG2000

Fig. 4. Tradeoffs between spatial resolution and compression artifacts. Image is bike.