

Wavelet-Based Quality Constrained Compression Using Binary Search

Yizhen Huang

Computer Sciences Department, University of Wisconsin-Madison

1210 W. Dayton St., Madison, WI, 53706, USA, huang.yizhen@gmail.com, 001-608-262-5068

ABSTRACT

Recently Gao et al. proposed a wavelet-based image quality metric, based on which, a quality constrained image subband quantization scheme is derived for lossy compression. But this scheme needs entropy coding to find the compression ratio decrement for tuning the final quantization step of each subband, which is time consuming and wasteful. In addition, computing the initial quantization steps also relies too much upon parameter tuning. This paper is a follow-up study, and proposes a simple binary search to determine the quantization step, according to the observation that, the quality metric decreases monotonically as the quantization steps increase. Experiments show that, the proposed quantization method outperforms its old counterpart in terms of computing expense, image quality, compression ratio and mathematical elegance.

Keywords— Lossy compression, quantization steps, Image Quality Metric, binary search, inverse function

1. INTRODUCTION

Quantization [1-4] plays an important role in lossy image compression. It is the primary contributor to high compression ratio, and likewise the major source of errors. Image Quality Metrics (IQMs) are necessary and important to assess and control the impact of quantization errors.

Image quality assessment is an important problem in the field of image processing, and has been heavily investigated by the image processing community (see [5-7] for a partial survey). Image distortion evaluation models have very wide applications, ranging from automatic adaptive imaging system, to the establishment of benchmark of image processing algorithms, and it even has significant values by providing novel IQMs and steering for some traditional image processing problems such as image interpolation, enhancement, and compression.

From the perspective of bionics, a great deal of effort has been made to develop IQMs that fit the Human Visual System (HVS). Although some metrics yield decent results, most are not always consistent with HVS and are sometimes limited to very specific applications. Furthermore, some metrics tend to be quite complicated for incorporating into other ingenious image processing algorithms. Watson et al. [8] developed a Discrete Cosine Transform (DCT) based video quality metric considering many characteristics of human visual sensitivity, and a simple IQM was proposed by Sendashonga et al. [9]

for both DCT and Discrete Wavelet Transform (DWT). Damera-Venkata et al. [10] developed two complementary quality metrics to measure the impact of frequency distortion and noise injection to HVS separately. The purpose was for quantifying the degradation of the compressed image as compared with the original image.

Gao et al. [11] proposed a new IQM in the wavelet domain, referred to as the Weighted Normalized Mean Square Error (WNMSE). It is claimed to be consistent with the human judgment of visual quality as well as able to perform quality estimation during the compression process. A quantization algorithm is developed according to the relationship between quantization steps, the WNMSE, and 5 statistical features defined. The limitation of this quantization algorithm is twofold: firstly it relies too much upon parameter tuning, which is a formidable challenge to be thoroughly motivated or justified; secondly there is a single uniform quantization step for each wavelet subband, which does not full exploit the HVS redundancy and consider the fact that, the sensitivity of the HVS is different for different type of image patches.

As a continuation of [11], we regard the WNMSE as a function $q(\cdot)$ of quantization steps and experimentally verify the approximate monotonicity of $q(\cdot)$. Given the desired WNMSE, the problem of computing the most appropriate quantization step for each wavelet subband becomes how to obtain the inverse of $q(\cdot)$. This naturally brings about the idea of using binary search to numerically approximate the desired quantization steps. It depends on no extra parameters and is elegant from the perspective of mathematics. Experiments demonstrate its superiority over its counterpart.

The rest of the manuscript is arranged as follows: Section 2 briefly introduces the definition and theoretical explanation of the WNMSE; Section 3 details the proposed binary search quantization method; Section 4 gives experimental results; Section 5 points out future work.

2. WEIGHTED NORMALIZED MSE

It is modeled in [12] that there are multiple layers of cells in the retina of human eyes. Each layer plays different functions of sampling and filtering in both temporal and spatial domains. Among all the cells, the P-type Ganglion cells and the M-type Ganglion cells compose the last layer of cells before the sensed signal enters the brain through the optical nerve. It is found that the P-type cells respond to higher spatial frequencies and the M-type cells to lower. This filtering property indicates that the HVS naturally decomposes the image into multiple subbands and process them differently.

We first introduce some basics about wavelet transform and subbands. Let $D_l(z), D_h(z), R_l(z)$ and $R_h(z)$ be the low-pass and high-pass analysis filters, the low-pass and high-pass synthesis filters, respectively, satisfying the perfect reconstruction condition $D_l(z)R_l(z) + D_h(z)R_h(z) = 1$. Denoting $d_l(\cdot), d_h(\cdot), r_l(\cdot)$ and $r_h(\cdot)$ as the impulse responses of $D_l(z), D_h(z), R_l(z)$ and $R_h(z)$, the four subbands of a 2D signal $S(x, y)$ is computed as below:

$$S_{LL1}(x, y) = d_l(x) * [d_l(y) * S(x, y)]$$

$$S_{LH1}(x, y) = d_l(x) * [d_h(y) * S(x, y)]$$

$$S_{HL1}(x, y) = d_h(x) * [d_l(y) * S(x, y)]$$

$$S_{HH1}(x, y) = d_h(x) * [d_h(y) * S(x, y)]$$

According to the perfect reconstruction condition, the wavelet reconstruction is computed as:

$$S(x, y) = r_l(x) * [r_l(y) * S_{LL1}(x, y)] + r_l(x) * [r_h(y) * S_{LH1}(x, y)] + r_h(x) * [r_l(y) * S_{HL1}(x, y)] + r_h(x) * [r_h(y) * S_{HH1}(x, y)]$$

As an undecimated transform, the LL subband can be decomposed further into level 2 subbands LL_2, LH_2, HL_2 and HH_2 . Then LL_2 can be decomposed into level 3 subbands etc. To be consistent with [11], we also use 3-level wavelet decomposition.

For better illustration, we also introduce the definition of *frequency index* in [11] of a subband sb :

$$fq(sb) = NL(sb) - NH(sb)$$

where NL and NH are the number of low-pass and high-pass filters that the subband sb passed through respectively. The frequency indexes of the ten subbands {HH₁,HL₁,LH₁,HH₂,HL₂,LH₂,HH₃,HL₃,LH₃,LL₃} are {-2,0,0,0,2,2,2,4,4,6}.

The Normalized MSE of a subband sb is defined as

$$NMSE(sb) = \frac{\sum_{x,y} (S_{sb}(x,y) - S'_{sb}(x,y))^2}{\sum_{x,y} S_{sb}^2(x,y)} \quad (1)$$

Where $S_{sb}(x,y)$ and $S'_{sb}(x,y)$ are the wavelet coefficients of the subband sb at (x,y) of the original and distorted images respectively.

An obvious fact is that, distortion to different subbands usually has different impact to the final image quality. So the NMSE of each subband is assigned with a weight^①:

$$w(sb) = 2^{level(sb) - 1 + fq(sb)/4}$$

Where $level(sb)$ is the wavelet transform level of the subband sb . So the Weighted NMSE is naturally written as

$$WNMSE^* = \sum_{sb} w(sb) NMSE(sb) \quad (2)$$

To make $WNMSE^*$ have similar range and monotonicity with the most familiar IQMs such as PSNR, the following one-one mapping is defined:

$$WNMSE = 20 \times \log_{10} \frac{100}{WNMSE^*}$$

3. THE PROPOSED METHOD

Quality constrained image compression is a concept that, for a user-specified quality requirement over a certain IQM, images are compressed to achieve the maximal compression ratio satisfying the quality requirement.

When applying IQMs to control the impact of quantization errors in lossy image compression, the $WNMSE$ can be regarded as a function $q(\cdot)$ of quantization steps:

$$WNMSE^* = q(\beta(HH_1), \beta(HL_1), \beta(LH_1), \beta(HH_2), \beta(HL_2), \beta(LH_2), \beta(HH_3), \beta(HL_3), \beta(LH_3), \beta(LL_3))$$

Where $\beta(sb)$ is the quantization step size for the subband sb .

Image wavelet decomposition has biological explanation corresponding to the HVS [13]. The weight $w(sb)$ for each subband sb also indicates the proportion of NMSE that each subband can count up to the whole $WNMSE^*$ in Eq.(2). This observation exactly adds up the following equality to decompose the multivariate function $q(\cdot)$ into 10 univariate functions denoted as $q_{sb}(\beta(sb))$ for each subband sb :

$$w(HH_1) NMSE(HH_1) = w(HL_1) NMSE(HL_1) = \dots = w(LH_3) NMSE(LH_3) = w(LL_3) NMSE(LL_3) \quad (3)$$

Despite of not being mathematically rigorous, $q_{sb}(\beta(sb))$ is monotonically decreasing as $\beta(sb)$ increases for natural image signals, e.g. Figure 1 are two plots of $q_{sb}(\beta(sb))$ when $sb=HL_1$ or HL_2 . Therefore, binary search may be applied to narrow the possible value range of the desired quantization

^① The rationality of the assigned weight of each subband is still controversial. We are currently working on this. We inherit the definition in [11] just for comparative purpose.

step $\beta(sb)$ at an exponential rate. Its main idea is to compute $q_{sb}(\cdot)$ using a certain trial $\beta^0(sb)$, if the computed $q_{sb}^0(\cdot)$ is lower than the desired $q_{sb}(\cdot)$, eliminate the value range of $\beta(sb)$ larger than $\beta^0(sb)$ and vice versa.

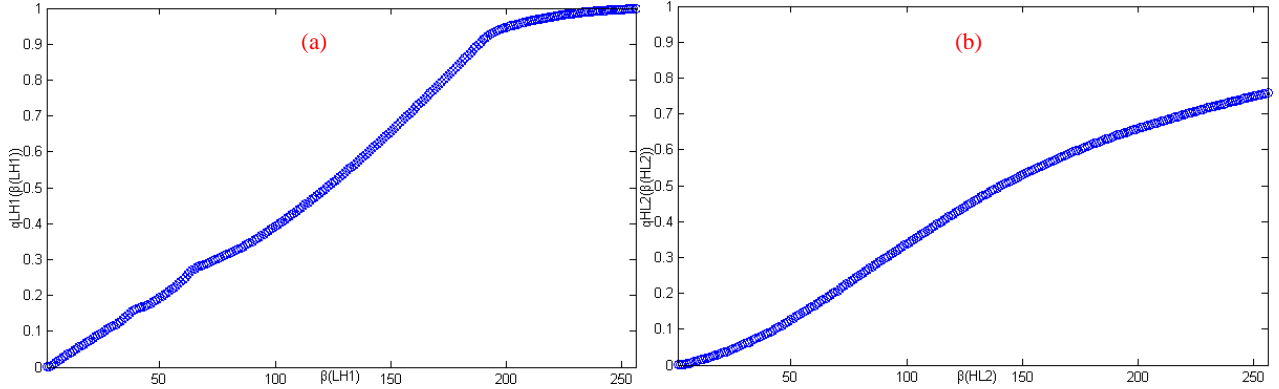


Figure 1. The plot of (a) $q_{LH_1}(\beta(LH_1))$, and (b) $q_{HL_2}(\beta(HL_2))$ as a function of the quantization step β for the red channel of the test image no. 9 used in our experiments.

The pseudo code of our binary search algorithm for a certain subband sb is outlined below:

1. Compute the desired NMSE according to Eq.(2) and (3)
2. $\beta_{low}=1, \beta_{high}=256$
3. while $low+1 < high$
4. $mid=(low+high)\div 2$
5. Compute the NMSE (denoted as $q_{sb}^0(\cdot)$) in Eq.(1) with $\beta^0(sb)=mid$
6. If $q_{sb}^0(\cdot)$ is smaller than the desired $q_{sb}(\cdot)$, $high=mid$; otherwise, $low=mid$

Compared to the Gao's method [11] for computing $\beta(sb)$, this method does not involve inverse wavelet transform and entropy encoding to tune the initial set of $\beta(sb)$.

4. EXPERIMENTAL RESULTS

The experiment[®] is conducted on classical benchmark images (as shown in Figure 2) commonly adopted by image processing researchers. All programs run on an Intel Pentium Dual-Core T3200 2.0GHz. The entropy coding module is implemented using C++; all other modules are implemented using Matlab.

Table 1 displays the performance comparison using the quantization step sizes computed by the proposed method and the method in [11] for three desired WNMSE values. To demonstrate the superiority of the proposed method, we set the desired WNMSE for the proposed method to be 1 larger than that for the Gao's method in [11]. It is seen from Table 1 that, even when its desired WNMSE is 1 larger, the proposed method achieves an averagely 2% higher compression ratio. Moreover, its computation time is about only 1/30 of that of the Gao's method. Figure 3 and 4 present the compressed images when the desired WNMSE are 28 and 27 for the proposed method and the Gao's method respectively. Images compressed with low WNMSE are displayed here for readers to compare their subjective quality, because customarily, images compressed with high WNMSE have slight quality loss and are difficult to be detected by the HVS.

[®] This paper follows the concepts of reproducible research. All the results presented in the paper are reproducible using the code and images available online at <http://pages.cs.wisc.edu/~huangyz/imageCompression.rar>.

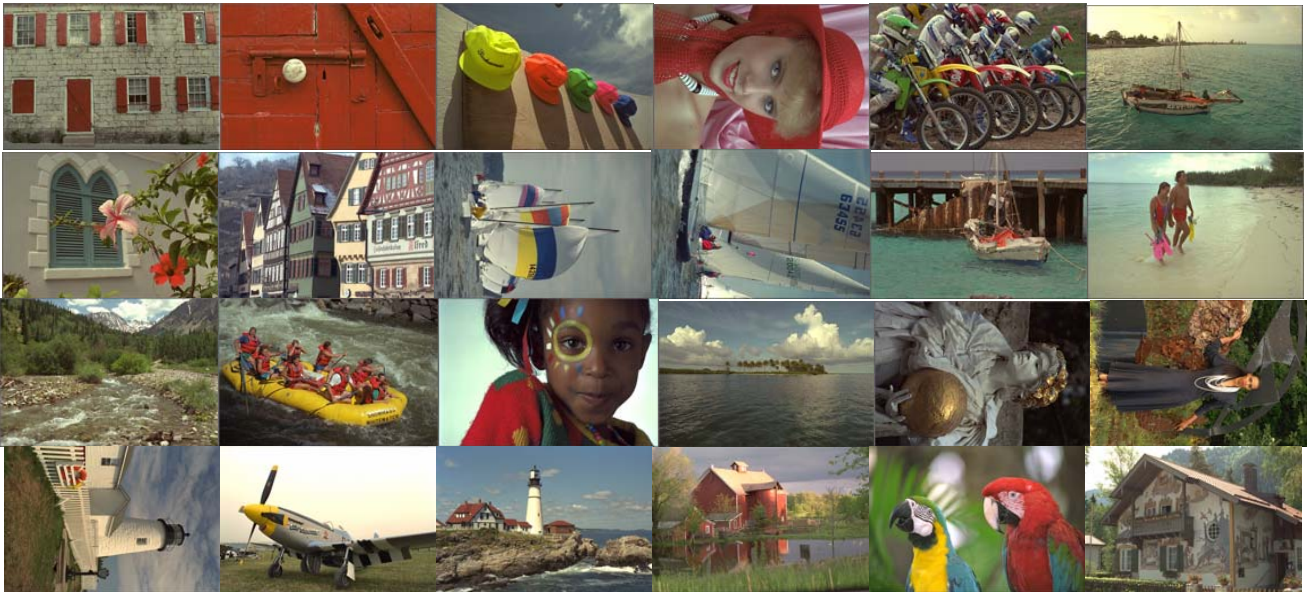


Figure 2. Twenty-four 768×512 images used in the test. These images are numbered as 1 to 24 from left to right, top to bottom.

5. CONCLUSION AND FUTURE WORK

This work presents a novel binary search method to find the quantization steps for quality constrained lossy compression. The concept of quality constrained compression and the proposed binary search framework can be extended using most image quality metrics, as long as, quality metrics decrease by and large as the quantization steps increase.

The quantization step is uniform for each wavelet subband, which does not take full advantage of the human visual redundancy: the sensitivity of the HVS varies for different type of image patches. So another future work is to refine the WNMSE with a local weight value of each image patch according to the sensitivity of the HVS to the patch. Based on such localized WNMSE, we are currently working on its corresponding local adaptive quantization method.

6. REFERENCES

- [1] S. Kasaei, M. Deriche, and B. Boashash, “A novel fingerprint image compression technique using wavelets packets and pyramid lattice vector quantization”, *IEEE Trans. Image Process.*, vol. 11, no. 12, pp. 1365–1378, Dec. 2002.
- [2] Y. Huh, J. J. Hwang, and K. R. Rao, “Block wavelet transform coding of images using classified vector quantization”, *IEEE Trans. Circuits Syst. Video Technol.*, vol. 5, no. 2, pp. 63–67, Feb. 1995.
- [3] P. C. Cosman, R. M. Gray, and M. Vetterli, “Vector quantization of image subbands: A survey”, *IEEE Trans. Image Process.*, vol. 5, no. 2, pp. 202–225, Feb. 1996.
- [4] X. Wang, L. Chang, M. K. Mandal, and S. Pancharathen, “Wavelet-based image coding using nonlinear interpolative vector quantization”, *IEEE Trans. Image Process.*, vol. 5, no. 3, pp. 518–522, Mar. 1996.
- [5] H. R. Sheikh, M. F. Sabir, and A. C. Bovik, “A statistical evaluation of recent full reference image quality assessment algorithms”, *IEEE Trans. Image Process.*, vol. 15, no. 11, pp. 3441–3452, Nov.

2006.

- [6] I. S. Jacobs, and C. P. Bean, “Are existing procedures enough? Image and video quality assessment: review of subjective and objective metrics”, SPIE vol. 6808, Image Quality and System Performance V, pp. 68080Q.1-68080Q.11, Jan. 2008.
- [7] H.M. Al-Otum, “Qualitative and quantitative image quality assessment of vector quantization, JPEG, and JPEG2000 compressed images”, J. Electronic Imaging, vol. 12, no. 3, pp. 511–521, Jul. 2003.
- [8] A. B. Watson, J. Hu, and J. F. McGowan, III, “DVQ: A digital video quality metric based on human vision”, J. Electronic Imaging, vol. 10, pp. 20–29, Jan. 2001.
- [9] M. Sendashonga, and F. Labeau, “Low complexity image quality assessment using frequency domain transforms”, Proc. IEEE Int. Conf. Image Process., pp. 385–388, Oct. 2006.
- [10] N. Damara-Venkata, T.D. Kite, W.S. Geisler, B.L. Evans, and A.C. Bovik, “Image quality assessment based on a degradation model,” IEEE Trans. Image Process., vol. 9, no. 4, pp. 636–650, Apr. 2000.
- [11] Z.G. Gao, and Y.F. Zheng, “Quality constrained compression using DWT-based image quality metric”, IEEE Trans. Circuits Syst. Video Technol., vol. 18, no. 7, pp. 910–922, Jul. 2008.
- [12] S. Shah, and M.D. Levine, “Visual information processing in primate cone pathways - Part I: A model,” IEEE Trans. Systems, Man, Cybern. B, vol. 26, no. 4, pp. 259–274, Apr. 1996.
- [13] S. G. Mallat, “Multi-frequency channel decompositions of images and wavelet models”, IEEE Trans. Acoustic, Speech & Sig. Process., vol. 37, no. 12, pp. 2091-2110, Dec. 1989.

image no.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
$\beta(HH_1)$	117	101	113	103	149	167	123	155	105	137	161	105	175	129	121	83	179	191	101	151	105	129	125	183
$\beta(HL_1)$	135	71	71	67	121	103	85	191	85	83	111	137	103	109	111	53	95	97	139	153	91	109	89	109
$\beta(LH_1)$	119	137	137	123	125	117	99	151	149	143	119	149	139	117	145	85	139	119	155	195	117	123	175	145
$\beta(HH_2)$	41	27	37	35	75	43	49	63	37	51	45	27	67	45	43	27	47	53	39	53	53	37	59	69
$\beta(HL_2)$	49	19	29	25	51	27	35	89	43	39	39	39	39	33	41	15	31	37	43	51	39	33	27	43
$\beta(LH_2)$	9	3	5	5	9	9	7	11	7	5	7	5	11	7	5	5	5	7	9	7	9	5	5	7
$\beta(HH_3)$	1	1	1	1	3	1	3	3	1	1	1	1	3	3	1	1	1	3	1	3	3	1	1	3
$\beta(HL_3)$	1	1	1	1	1	1	1	3	1	1	1	1	1	1	3	1	1	1	1	1	1	1	1	1
$\beta(LH_3)$	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
$\beta(LL_3)$	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	3	1	1	1
R1(bpp)	0.54	0.84	0.98	0.97	0.88	0.75	0.76	0.68	0.82	0.68	0.54	0.86	0.94	0.96	0.87	0.72	0.79	0.62	0.80	0.68	0.58	0.85	0.95	0.95
R2(bpp)	0.41	0.66	0.83	0.72	0.66	0.56	0.57	0.50	0.62	0.53	0.44	0.63	0.66	0.74	0.67	0.50	0.63	0.51	0.59	0.50	0.47	0.70	0.81	0.73
WNMSE	31.0	31.0	31.0	31.0	31.0	31.0	31.0	31.0	31.0	31.0	31.0	31.0	31.0	31.0	30.9	31.0	31.0	31.0	31.0	31.0	31.0	31.0	31.0	31.0
T(second)	0.21	0.22	0.21	0.22	0.21	0.22	0.22	0.21	0.20	0.22	0.21	0.22	0.22	0.22	0.22	0.21	0.21	0.22	0.22	0.21	0.22	0.20	0.21	0.22

(a) when the desired WNMSE is set to be 31 using the proposed binary search method.

image no.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
R1(bpp)	0.58	0.89	0.97	0.97	0.89	0.75	0.77	0.66	0.84	0.71	0.58	0.89	0.97	0.97	0.89	0.75	0.78	0.66	0.84	0.71	0.58	0.89	0.97	0.97
R2(bpp)	0.49	0.68	0.80	0.81	0.64	0.57	0.60	0.54	0.66	0.54	0.45	0.62	0.74	0.73	0.66	0.63	0.64	0.57	0.66	0.49	0.46	0.73	0.83	0.71
WNMSE	29.9	29.9	29.9	30.0	29.9	30.0	30.0	29.9	30.0	29.9	30.0	30.0	30.0	30.0	29.9	29.9	30.0	30.0	29.9	30.0	30.0	29.9	30.0	29.9
T(second)	6.30	6.47	6.35	6.42	6.32	6.27	6.23	6.23	6.41	6.27	6.30	6.47	6.35	6.42	6.32	6.27	6.23	6.23	6.41	6.27	6.30	6.47	6.35	6.42

(b) when the desired WNMSE is set to be 30 using the Gao’s method [11].

image no.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
R1(bpp)	0.57	0.77	0.92	0.90	0.83	0.64	0.66	0.55	0.74	0.62	0.49	0.79	0.87	0.83	0.86	0.64	0.70	0.53	0.70	0.55	0.50	0.81	0.88	0.86
R2(bpp)	0.46	0.60	0.78	0.73	0.60	0.49	0.47	0.47	0.55	0.43	0.40	0.67	0.64	0.61	0.73	0.49	0.59	0.43	0.59	0.40	0.38	0.57	0.73	0.68
WNMSE	28.0	28.0	28.0	28.0	28.0	28.0	28.0	28.0	28.0	28.0	28.0	28.0	28.0	28.0	28.0	28.0	28.0	28.0	28.0	28.0	28.0	28.0	28.0	28.0
T(second)	0.20	0.21	0.21	0.21	0.22	0.21	0.21	0.20	0.22	0.21	0.20	0.21	0.21	0.21	0.22	0.21	0.21	0.20	0.22	0.21	0.20	0.21	0.21	0.21

(c) when the desired WNMSE is set to be 28 using the proposed binary search method.

image no.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
R1(bpp)	0.58	0.81	0.90	0.91	0.83	0.67	0.66	0.57	0.72	0.62	0.50	0.80	0.88	0.85	0.84	0.68	0.72	0.56	0.73	0.58	0.49	0.80	0.90	0.87
R2(bpp)	0.43	0.61	0.76	0.69	0.58	0.49	0.46	0.45	0.51	0.48	0.35	0.65	0.67	0.71	0.59	0.54	0.62	0.41	0.62	0.41	0.34	0.59	0.76	0.61
WNMSE	26.9	27.0	26.9	27.0	26.9	26.9	27.0	27.0	26.9	26.9	27.0	27.0	26.9	27.0	26.9	26.9	26.9	26.9	26.9	27.0	27.0	30.0	26.9	27.0
T(second)	6.36	6.21	6.42	6.30	6.48	6.48	6.38	6.34	6.50	6.47	6.36	6.21	6.42	6.30	6.48	6.48	6.38	6.34	6.50	6.46	6.36	6.21	6.42	6.30

(d) when the desired WNMSE is set to be 27 using the Gao's method [11].

image no.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
R1(bpp)	0.87	0.94	0.99	1.02	0.96	0.80	0.85	0.74	0.87	0.75	0.66	0.89	1.02	1.02	0.97	0.76	0.82	0.73	0.89	0.76	0.83	0.95	1.01	0.99
R2(bpp)	0.64	0.81	0.72	0.79	0.71	0.64	0.72	0.61	0.70	0.54	0.47	0.62	0.75	0.84	0.83	0.53	0.68	0.58	0.70	0.56	0.69	0.70	0.82	0.80
WNMSE	33.0	33.0	33.0	33.0	33.0	33.0	33.0	33.0	33.0	33.0	33.0	33.0	33.0	33.0	33.0	33.0	33.0	33.0	33.1	33.0	33.0	33.0	33.0	33.1
T(second)	0.22	0.21	0.21	0.22	0.21	0.21	0.22	0.22	0.20	0.22	0.22	0.21	0.21	0.22	0.21	0.21	0.22	0.22	0.20	0.22	0.22	0.21	0.21	0.22

(e) when the desired WNMSE is set to be 33 using the proposed binary search method.

image no.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
R1(bpp)	0.86	0.94	0.99	1.03	0.97	0.83	0.87	0.75	0.89	0.77	0.66	0.92	1.05	1.04	0.96	0.80	0.84	0.75	0.88	0.78	0.84	0.95	1.04	1.03
R2(bpp)	0.62	0.74	0.79	0.82	0.70	0.67	0.62	0.58	0.71	0.62	0.53	0.77	0.73	0.77	0.79	0.59	0.71	0.61	0.62	0.54	0.58	0.69	0.81	0.79
WNMSE	31.9	32.0	31.9	31.9	31.9	32.0	32.0	32.0	31.9	32.0	31.9	32.0	32.0	31.9	32.0	32.0	31.9	32.0	31.9	32.0	32.0	32.0	32.0	31.9
T(second)	6.40	6.47	6.41	6.28	6.26	6.28	6.36	6.43	6.27	6.47	6.40	6.47	6.41	6.28	6.26	6.28	6.36	6.43	6.27	6.47	6.40	6.47	6.41	6.28

(f) when the desired WNMSE is set to be 32 using the Gao's method [11].

Table 1. Subband quantization step size $\beta(sb)$, compression ratio (in bit per pixel) and achieved WNMSE of the twenty-four testing images. R1 and R2 correspond to the compression ratio achieved using Huffman Coding and WinRAR respectively. T is the time solely spent for computing quantization step sizes. Due to the limited space, only the quantization steps for the green channel in SubTable (a) are displayed.







Figure 3. The twenty-four test images compressed by the proposed method when the desired WNMSE is 28. Due to the limited space, the images displayed here are scaled down, and the best scale for viewing these images is in 600%.





Figure 4. The twenty-four test images compressed by the Gao's method [11] when the desired WNMSE is 27. Due to the limited space, the images displayed here are scaled down, and the best scale for viewing these images is in 600%.