PSNR-BASED OPTIMIZATION OF JPEG BASELINE COMPRESSION ON COLOR IMAGES

Linfeng Guo

BrainMedia LLC
New York, New York 10016, USA
lguo@brainmediaco.com

Yan Meng

Stevens Institute of Technology
Hoboken, New Jersey 07030, USA
ymeng1@stevens.edu

ABSTRACT

A fast rate-distortion analytical framework for the PSNR-based optimization of JPEG baseline compression on color images is proposed in this paper. The analysis is conducted in the luminance-chrominance color space and based upon the ρ-domain analysis method [1]. The simulation results show that the proposed scheme can dramatically improve the performance of the JPEG baseline compression on color images in terms of peak signal-to-noise ratio (PSNR). Meanwhile, while the complexity of the proposed scheme is kept low, its results are very close to that from exhaustive searching method.

Index Terms**—**Image coding, quantization, rate distortion theory

1. INTRODUCTION

The JPEG baseline image encoder [2] comprises the following stages: color space conversion, chroma sub-sampling (optional), 2-D 8×8 DCT, linear quantization, zigzag scan, run-length encoding (RLE), and Huffman coding. The decoder is just the step-wise inverse procedure of the encoder.

Normally the input to the JPEG baseline encoder is a color image represented by a $M\times N$ matrix, in which each pixel has three components (RGB), and each component is represented as an 8-bit integer. Since the RGB components are not completely independent, and human visual system has different acuities on luminance and chrominance details, the image can be transformed from RGB space to another color space. The most widely adopted space in compression is YCrCb, in which Y denotes the luma component, and Cr and Cb the chroma components. The forward and inverse transformations between RGB and YCrCb are defined in [3].

Next, the Cr and Cb components may be sub-sampled. Most JPEG pictures we see are 4:2:0 subsampled [4]. In most digital cameras, the Cr and Cb are 4:2:2 sub-sampled [5]. When the sub-sampling step is skipped, it may be called 4:4:4 subsampled.

During the linear quantization, two 8×8 quantization tables for luminance and chrominance are defined respectively. These quantization tables reflect the human visual sensitivities to different DCT bases under certain viewing condition, and can be scaled to achieve different qualities and bit rates of the compressed image, i.e., a quality scale parameter is multiplied to these two quantization tables to generate the actual quantization tables. Each coefficient in every 8×8 DCT block in the image is divided by the corresponding quantization step value in the quantization table, and the result is rounded up to the nearest integer. Although JPEG baseline standard suggests two 8×8 quantization tables [2] for luminance and chrominance, in reality they are seldom adopted by real applications. Rather, most real applications come up with their own quantization tables.

After the quantization, the DCT coefficients are zigzag scanned and entropy coded.

In this paper, instead of using a single quality scale to control the compression quality, we propose a fast rate-distortion searching method with different quality scales for luma and chroma components (Cr and Cb will share the same quality scale, since they will share the same quality table, which is required by the JPEG baseline standard) to achieve an optimal PSNR result. Furthermore, the optimal choice of chroma subsampling is also determined through the proposed fast searching algorithm.

Our analysis and proposed fast searching method will be based on the following configuration and assumptions:

- YCrCb color space conversion
- Recommended quantization tables in JPEG standard
- Recommended Huffman tables in JPEG standard
- 4:2:0 and 4:4:4 chroma subsampling options
- PSNR quality metric

However, the proposed method can be extended to JPEG compression applications with alternative color space conversion methods, customized quantization tables and/or Huffman tables, and other quality metrics, such as weighted PSNR.

2. ANALYTICAL FRAMEWORK

2.1. Previous Work

Rate-distortion analysis based optimization of image compression has been actively studied for many years, which is targeting at maximizing the coding performance under given bit rate constraint [4]. Most of these works have been limited to gray level images. Very few works on the optimal bit allocation between luma and chroma components in color image compression were reported. Although Robert E. Van Dyck etc. [5] investigated the bit allocation amongst luma and chroma components in different color spaces in the context of subband/VQ compression of color images, their research was focused on the optimal perceptual weighting at different subbands of either luma or chroma component. As for the issue of optimal sub-sampling in image coding, to our knowledge, the only analytical work was done by Alfred M. Bruchstein etc. [6]. In their work, an analytical model was first established to explain the better transform coding of images with down-scaling at low bit rate. Subsequently, a simple algorithm was developed to derive the optimal down-scaling factor of a gray-level image at given bit rate. Apparently, their work was still limited to the compression of gray level images. Furthermore, their method only provides the optimal down-scaling factor, and could not accurately predict the actual rate-distortion performance.

2.2. ρ-domain Analysis Method

The ρ-domain analysis tool [1] has been chosen in our analysis due to its proven computational efficiency and accuracy. The motivation of ρ-domain analysis of rate and distortion performance of image coding is from the observation that the percentage of zeroed coefficients (denoted as ρ) after the compression has a close correlation with the quantizer stepsize, the bit rate, and the quality of the compressed image. Therefore the fast estimations of the relationships between ρ and bit rate, ρ and quantizer stepsize, and ρ and distortion are obtained. In ρ-domain analysis method, the relationship between ρ and bit rate is characterized as a straight line $R_i = \theta \cdot \rho_i + \beta$. In order to estimate the parameter θ and β, two image dependent parameters $Q_{nz}(\rho)$ and $Q_z(\rho)$ at various ρ values are derived first. After that, the corresponding bit rate can be derived by

$$R(\rho_i) = A(\rho_i) \cdot Q_{nz}(\rho_i) + B(\rho_i) \cdot Q_z(\rho_i) + C(\rho_i) , \quad (1)$$

where A, B, and C are also ρ dependent parameters, and can be pre-computed through multiple linear regression method over the ($R(\rho_i), Q_{nz}(\rho)$, $Q_z(\rho)$) results from the compressions of a set of test images. Finally, the slope θ is computed through

$$\theta = \frac{\sum_{i=1}^{N} \rho_i \cdot \sum_{i=1}^{N} R(\rho_i) - \sum_{i=1}^{N} \rho_i R(\rho_i)}{N \cdot \sum_{i=1}^{N} \rho_i^2 - \left(\sum_{i=1}^{N} \rho_i \right)^2} , \quad (2)$$

and the bit rate for any given ρ can be computed by

$$\hat{R}(\rho) = \frac{1}{N} \sum_{i=1}^{N} R(\rho_i) + \theta \cdot \left(\rho - \frac{1}{N} \sum_{i=1}^{N} \rho_i \right) . \quad (3)$$

In order to set up the relationships between ρ and the quantizer stepsize, and between ρ and the distortion, the histograms of all of the visually weighted DCT coefficients in either the luma or chroma component are collected, and numerical searching processes on these histograms are deployed.

2.3. PSNR Analysis in Luma-chroma Space

Given a $M \times N$ color image with 24-bit RGB components for each pixel and its reconstructed one after the JPEG compression, the compression distortion is generally evaluated by the mean square error. Let $R(i,j)$, $G(i,j)$, and $B(i,j)$ denote the R, G, B values of pixel at (i,j) in the original image, and $\hat{R}(i,j)$, $\hat{G}(i,j)$, and $\hat{B}(i,j)$ denote the R, G, B values of pixel at the same position in the compressed one. If we define

$$\mathbf{X_{i,j}} = \begin{bmatrix} R(i,j) \\ G(i,j) \\ B(i,j) \end{bmatrix} - \begin{bmatrix} \hat{R}(i,j) \\ \hat{G}(i,j) \\ \hat{B}(i,j) \end{bmatrix} = \mathbf{AY_{i,j}} - \mathbf{A\hat{Y}_{i,j}} = \mathbf{A(Y_{i,j} - \hat{Y}_{i,j})} , \quad (4)$$

where $\mathbf{A} = \begin{bmatrix} 1.0 & 0 & 1.402 \\ 1.0 & -0.344 & -0.714 \\ 1.0 & 1.772 & 0 \end{bmatrix}$, which is the color space conversion matrix from YCbCr to RGB,

$\mathbf{Y_{i,j}} = \begin{bmatrix} Y(i,j) \\ C_b(i,j) \\ C_r(i,j) \end{bmatrix}$, and $\mathbf{\hat{Y}_{i,j}} = \begin{bmatrix} \hat{Y}(i,j) \\ \hat{C}_b(i,j) \\ \hat{C}_r(i,j) \end{bmatrix}$, then we get

$$MSE = \frac{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \mathbf{X'_{i,j} X_{i,j}}}{3 \times M \times N} = \frac{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \left(\mathbf{Y'} - \mathbf{\hat{Y}'} \right) \mathbf{A'A} \left(\mathbf{Y} - \mathbf{\hat{Y}'} \right)}{3 \times M \times N} \quad (5)$$

After proper substitution and expansion, the above distortion can be represented as

$$MSE = \frac{1}{3} \left[3 \times MSE(Y) + 3.2583 \times MSE(C_b) + 2.4754 \times MSE(C_r) \right] + 2.856 \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \left(\Delta Y(i,j) \cdot \Delta C_b(i,j) \right) + 1.376 \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \left(\Delta Y(i,j) \cdot \Delta C_r(i,j) \right) + 0.4912 \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \left(\Delta C_b(i,j) \cdot \Delta C_r(i,j) \right), \quad (6)$$

where $\Delta Y(i,j) = Y(i,j) - \hat{Y}(i,j)$, $\Delta C_b(i,j) = C_b(i,j) - \hat{C}_b(i,j)$, $\Delta C_r(i,j) = C_r(i,j) - \hat{C}_r(i,j)$,

$MSE(Y) = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \left(Y(i,j) - \hat{Y}(i,j) \right)^2 / MN$,

$MSE(C_r) = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \left(C_r(i,j) - \hat{C}_r(i,j) \right)^2 / MN$,

$MSE(C_b) = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \left(C_b(i,j) - \hat{C}_b(i,j) \right)^2 / MN$.

Since $\Delta Y(i,j)$, $\Delta C_b(i,j)$, and $\Delta C_r(i,j)$ are zero mean random variables, and their PDFs can be approximated by even distribution when the quantization step size is small, the sum of the last three items in Equation (6) can be approximated by 0. Therefore, the rate-distortion optimization problem can be formulated as

$$\arg \min_{R_Y, R_{C_b}, R_{C_r}} \left[3 \times MSE(Y) + 3.2583 \times MSE(C_b) + 2.4754 \times MSE(C_r) \right], \quad (7)$$

under the constraint of $R = R_Y + R_{C_b} + R_{C_r}$, where R_Y, R_{C_b} , R_{C_r} are the bit rates of Y, Cb, and Cr, respectively.

Meanwhile, based upon the description of the JPEG quantization in session 1, the one step quantization of coefficients can be treated as the tandem process of visual weighting and subsequent uniform quantization. Let's denote the visually weighted luma and chroma components as Y^w , C_b^w , and C_r^w , our ρ-domain bit rate estimation method can be applied to Y^w , C_b^w , and C_r^w , and the distortion of each component will be derived by following the method described in session 2.2. In order to do that, we have selected a small image database and applied compression and numerical analysis methods on the results to derive A, B, C, α, β values as described in [1] at different ρ values in either luma or chroma components.

To illustrate the accuracy of the proposed fast rate-distortion estimation method, the rate-distortion curves of several images from both the actual coding and the proposed fast estimation are compared in Fig. 1. It can be seen the estimated results are very close to that derived from the real compression.

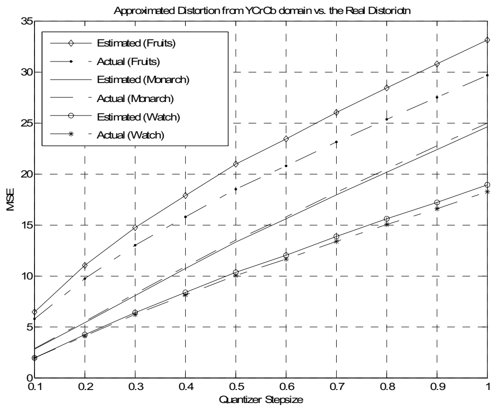

Fig. 1. Comparison of the estimated and the actual rate-distortion results.

By applying the ρ-domain analysis method to Y, C_b, and C_r, respectively, we obtain three sets of lookup tables with entries of quantizer stepsize q_i, and their corresponding bit rates and distortions. By summing up the bit rates and distortions for C_b and C_r (need to be weighted again), we get the corresponding overall bit rate and distortion for each given q_i. Therefore, given a target bit rate, we can easily figure out the optimal bit budget ratio between luma and chroma components and the corresponding quality scales by iteratively running through the following steps:

1. Given a bit budget ratio, calculate the bit rates for luma and chroma.
2. Figure out the PSNR (or MSE) of Y, C_b and C_r through the lookup tables. Here the appropriate interpolation might be required.
3. The overall MSE (or PSNR) is computed using Eq. 7.

Finally, a curve of distortions versus different bit budget ratios can be derived, and the trough of the curve will be found.

2.4. Consideration of Chroma Sub-sampling

In most JPEG baseline applications, a chroma sub-sampling operation is adopted after the color space conversion. Although this operation introduces extra distortion, the reward of increasing the available bit rate for chroma components by a factor of 4 will justify the quality loss owing to the sub-sampling. By taking this operation, the procedure of compression of chroma components becomes:

1. The chroma component is properly pre-filtered for anti-aliasing purpose.
2. The pre-filtered chroma component is decimated with certain sub-sampling factor along either one direction (horizontal or vertical) or both.
3. The sub-sampled component is compressed.

During decoding, the compressed chroma component is decompressed and interpolated to restore to its original size.

Although different kinds of anti-aliasing filters can be designed and implemented, quite frequently a simple averaging filter is adopted for the sake of simplicity, e.g., for the 4:2:2 sub-sampling, a 2×1 averaging filtering is executed; while for the 4:2:0 sub-sampling, a 2×2 averaging filtering is performed.

Let's assume the original chroma component and its compressed version are denoted as $x_c(i,j)$ and $\hat{x}_c(i,j)$ respectively, where $i\in[0,M-1]$, and $j\in[0,N-1]$. Similarly, the sub-sampled and its compressed version are denoted as $x_c^d(i,j)$ and $\hat{x}_c^d(i,j)$, where $i\in[0,M/2-1]$, and $j\in[0,N/2-1]$. We also denote the up-sampled version of the sub-sampled $x_c^d(i,j)$ as $\tilde{x}_c(i,j)$. Without losing generality, the sub-sampling and up-sampling are simply defined as averaging and duplicating operation. Therefore, we have

$$x_c^d(i,j)=\frac{1}{4}\sum_{m=0}^{1}\sum_{n=0}^{1}x(2\cdot i+m,2\cdot j-n)),$$

where $i\in[0,M/2-1]$ and $j\in[0,N/2-1]$. We also have $\tilde{x}_c(i,j)=x_c^d(\lfloor i/2\rfloor,\lfloor j/2\rfloor)$, and $\hat{x}_c(i,j)=\hat{x}_c^d(\lfloor i/2\rfloor,\lfloor j/2\rfloor)$, where $i\in[0,M-1]$, and $j\in[0,N-1]$.

Therefore, the coding distortion can be represented as

$$MSE=\frac{1}{MN}\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\left(x_c(i,j)-\hat{x}_c(i,j)\right)^2$$
$$=\frac{1}{MN}\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\left[\begin{matrix}\left(x_c(i,j)-\tilde{x}_c(i,j)\right)^2+\left(\tilde{x}_c(i,j)-\hat{x}_c(i,j)\right)^2+\\ 2\cdot\left(x_c(i,j)-\tilde{x}_c(i,j)\right)\left(\tilde{x}_c(i,j)-\hat{x}_c(i,j)\right)\end{matrix}\right] \quad (8)$$

Since the third item in the right hand of the above equation are the product of two zero mean random variables, the summation of this product across the whole image can be approximated as zero. Hence the distortion can be simplified as

$$MSE\approx\frac{1}{MN}\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\left(x_c(i,j)-\tilde{x}_c(i,j)\right)^2+\frac{1}{MN}\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\left(\tilde{x}_c(i,j)-\hat{x}_c^d(i,j)\right)^2$$
$$=MSE(C^d)+\frac{1}{(M/2)\cdot(N/2)}\sum_{i=0}^{M/2-1}\sum_{j=0}^{N/2-1}\left(\hat{x}_c^d(i,j)-x_c^d(i,j)\right)^2$$
$$=MSE(C^S)+MSE(C^C) \quad (9)$$

Obviously, the first item on the right hand of Eq. 9 denotes the distortion incurred by subsampling, and the second term denotes the distortion incurred by actual compression. Therefore, the overall compression distortion can be rewritten as

$$MSE_{downsampling}(ratio)=\left(\begin{matrix}3\times MSE(Y)+3.2583\times\left(MSE(C_b^S)+MSE(C_b^C)\right)+\\ 2.4754\times\left(MSE(C_r^S)+MSE(C_r^C)\right)\end{matrix}\right)/3$$

In order to decide whether the chroma down-sampling should be adopted, we need to inspect the difference of the distortions from two options, i.e.,

$$D=MSE(ratio_{optimal})-MSE_{downsampling}(ratio_{optimal})$$

When the result is positive, the subsampling of chroma components should be adopted; otherwise, no subsampling should be taken.

3. DESCRIPTION OF THE FAST RATE-DISTORITON ANALYSIS ALGORITHM

Based on the analysis in section 2, the proposed fast optimization algorithm of chroma subsampling and bit allocation between chroma and luma can be summarized as the following steps:

1. The ρ-quantizer, ρ-rate, and ρ-distortion relationships in Y, Cr, and Cb are derived by following the method in session 2.
2. The distortions due to 2×2 averaging filtering in Cr and Cb are computed directly.
3. The ρ-quantizer, ρ-rate, and ρ-distortion relationships in subsampled Cr and Cb are derived.
4. The joint ρ-rate and ρ-distortion relationships of Cr and Cb are computed by adding their corresponding ones.
5. A numerical searching method is deployed to find the optimal quality scales for luma and chroma components as well as the optimal chroma sub-sampling choice.
6. The derived quality scales and the selected chroma subsampling style are applied to the real compression to reach the optimal PSNR.

4. EXPERIMENTAL RESULTS

In order to verify the gain of the proposed optimization method, 4 well known test color images are selected, and each is compressed with the default method and the proposed optimization method at the bit rates of 2.0 bpp. The setting of default method includes the adoption of 4:2:0 chroma subsampling and JPEG recommended quantization tables. The proposed optimization method follows the fast searching scheme described in section 3. The results are tabulated in Table 1. Here the performance of the compression is evaluated by the PSNR (in dB). The results of the searched quality scales for luma and chroma components and the selected chroma subsampling style are also listed in the tables.

TABLE 1. OPTIMIZATION GAIN AT BIT RATE OF 2.0 BPP

Image	PSNR (Def)	PSNR (Opt)	Scale (Y)	Scale (CrCb)	Subsample
Airplane	33.38	37.38	0.31	0.32	4:4:4
Fruits	35.62	36.64	0.30	0.43	4:4:4
Goldhill	34.49	34.60	0.30	0.13	4:2:0
Lena	34.43	35.20	0.31	0.36	4:4:4

It can be easily noticed that the proposed optimization scheme always brings extra gain to the compression. This constant optimization gain proves the validity of the proposed optimization scheme.

The visual inspection of the image qualities from two methods shows that most of the time the visual quality of the image compressed by the proposed method is better than that by default method, but not always. Careful scrutiny of Eq. 7 suggests that additional visual weighting factors should be applied to the distortions of luma and chroma components to reflect the human vision's different acuities on Y, Cr, and Cb details. Another factor that affects the consistency of visual quality improvement might be due to the ignorance of different visual distortion visibilities in original and down-sampled chroma components.

In order to evaluate the accuracy of the proposed scheme, PSNR-based exhaustive searching of the optimal chroma subsampling and bit allocation between luma and chroma has been conducted on the compressions of two test images "Lena" and "Fruits". The test bit rates include 0.5 bpp, 1.0 bpp, and 2.0 bpp. The results are compared with those derived from the proposed fast searching scheme, and tabulated in Table 2. It can be seen that the optimal searching results are very close to that derived from the exhaustive searching.

TABLE 2. COMPARISON OF FAST AND EXHAUSTIVE SEARCHING RESULTS

Bit Rate (bpp)	Quality (PSNR)	Lena	Fruits
0.5	Exhaustive	30.90 (4:2:0)	30.80 (4:2:0)
	Fast	30.85 (4:2:0)	30.75 (4:2:0)
1.0	Exhaustive	33.34 (4:4:4)	33.73 (4:4:4)
	Fast	33.27 (4:4:4)	33.66 (4:4:4)
2.0	Exhaustive	35.29 (4:4:4)	36.79 (4:4:4)
	Fast	35.20 (4:4:4)	36.64 (4:4:4)

5. CONCLUSIONS

In this paper, owing to some reasonable approximations, a fast searching scheme for optimal choices of chroma sub-sampling style and bit allocation between luma and chroma components is established. Our simulation results confirm the validity of these assumptions as well as the validity and efficiency of the proposed fast searching scheme. However, more work on the proposed method needs to be done to achieve consistent visual quality improvement. Although our analysis is based upon the YCrCb color space, JPEG recommend quantization tables, and Huffman table, it can be easily extended to those applications which has their own color space conversion, quantization tables, and (or) Huffman tables.

6. REFERENCES

[1] Zhihai He and Sanjit K. Mitra, "A unified rate-distortion analysis framework for transform coding," *IEEE Trans. on Circuits and Systems on Video Technology*, vol. 11, no. 12, pp. 1221-1236, Dec. 2001.

[2] ITU-T Rec. T81 | ISO/IEC JTC1 10918-1: Information Technology – Digital Compression and Coding of Continuous-tone Still Images: Requirements and Guidelines, 1994.

[3] ITU-T Recommendation 601: Encoding Parameters of Digital Television for Studios, 1982.

[4] Antonio Ortega and Kannan Ramchandran, "Rate-distortion methods for image and video compression," *IEEE Signal Processing Magazine,* pp. 23-50, November 1998.

[5] Robert E. Van Dyck, Sarah A. Rajala, "Subband/VQ coding of color images with perceptually optimal bit allocation," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 4, no. 1, pp. 68-82, Feb, 1994.

[6] Alfred M. Bruckstein, Michael Elad, and Ron Kimmel, "Down-sampling for better transform compression," *IEEE Trans. on Image Processing*, vol. 12, no. 9, pp. 1132-1144, Sept, 2003.