A Novel Source Modeling Framework for Bit Rate and Picture Quality Control in DCT Visual Coding

Zhihai He, Yong Kwan Kim, and Sanjit K. Mitra
Dept. of Electrical and Computer Engineering
University of California, Santa Barbara
{zhihai, ykkim, mitra}@iplab.ece.ucsb.edu

Extended Abstract

(for PCS 2001)

The Major Contributions of This Work

We present a novel source modeling framework in this paper. With the proposed source model, we can estimate the rate-distortion (R-D) curve before quantization and coding for DCT-based still images and video compression. The relative estimation error is less than 5%. In addition, the estimation algorithm has very low computational complexity. Based on the estimated R-D curve, we can perform accurate bit rate and picture quality control for still images and video compression.

Related Work

For a typical transform coding system, such as the JPEG coding [1], the theoretical formula [2] does not work, especially at low bit rates. In Fig. 1, we plot the actual JPEG coding bit rates and the theoretical entropy at different quantization scales. It can be seen that the relative error between them is very large. There are several algorithms published in the literature to model the R-D curve [3, 4, 5]. These algorithms either have high computational comlexity [5] or large estimation error [4, 3]. Obviously, the variance-based rate formula [3, 4] is far insufficient to characterize the input source data and to model the coding algorithm.

The Proposed Work

The proposed work is based on the following ideas:

- **ρ-domain source modeling.** Zeros play a key role in transform coding. Let ρ be the percentage of zeros among the quantized DCT coefficients. Note that ρ monotonically increases with the quantization stepsize q. This implies that there is a one-to-one mapping between them. Therefore, any rate function in the q-domain can be mapped into the ρ-domain and vise versa. The mapping between q and ρ can be computed from the distribution of the DCT coefficients. In this work, we propose to study the rate function in the ρ-domain because the rate function has some very unique properties in this domain.

- **Characteristic rate curve.** Variance is far insufficient to describe the R-D characteristics of the input source data. In this work, we introduce the new concept of characteristic rate curves to characterize the input source data accurately and robustly.

- **Rate curve decomposition.** In functional analysis and digital signal processing, we often study the behavior of a complex function/signal by decomposing it into a series of basis functions/signals with well known properties. Fourier transform and spectrum analysis are good examples of this. In our work, we apply this method to R-D curve analysis and estimation. The actual rate curve is represented as a linear combination of the characteristic rate curves.

Definition of the Characteristic Rate Curves

The characteristic rate curves are defined during the following "pseudo coding" process. First, the DCT coefficients are uniformly quantized with stepsize q. Second, we re-arrange all the DCT coefficients in each block into a 1-D array $\mathcal{L}$ in a zig-zag scan order. For each consecutive string of zeros in $\mathcal{L}$, their run length is counted. We sum up the sizes of the run-length numbers in all the blocks, and denote the sum by Q_z. Here, the size of a non-zero integer is the number of bits for its sign-magnitude representation. For example, the size of -7 is 4. Let Q_{nz} be the sum of the sizes of all the non-zero DCT coefficients. Obviously, both of them are functions of q, denoted by $Q_{nz}(q)$ and $Q_z(q)$, respectively. Note that there is a one-to-one mapping between ρ and q. Therefore, Q_{nz} and Q_z are also functions of ρ, denoted by $Q_{nz}(\rho)$ and $Q_z(\rho)$, respectively. These two curves are called the characteristic rate curves. For any input image, following the above procedures, we can generate these two curves by varying the quantization stepsize q.

Properties of the Characteristic Rate Curves

We show that the characteristic rate curves have some very interesting properties. To this end, we randomly select 24 sample images with a wide range of R-D characteristics. The sample images are shown in Fig. 2. We plot $Q_{nz}(\rho)$ (solid line) and $Q_z(\rho)$ (dash-dot line) for each sample image in Fig. 3. Two observations can be made on these plots. First, the two curves of each sample image have almost the same pattern, although the sample images are quite different from each other. This invariance makes it possible for us to estimate the R-D curve accurately and robustly. This is the reason why we develop the source model in the ρ-domain. Second, $Q_{nz}(\rho)$ is almost a straight line passing through [1.0, 0.0]. It should be noted that when the percentage of zeros is 100%, there is no non-zero coefficients. Hence Q_{nz} becomes zero. Let θ be the slope of the straight line. We have run this experiment over many other images. The above two observations always hold.

Estimate $Q_{nz}(\rho)$

$Q_{nz}(\rho)$ is modeled as a straight line passing through [1.0, 0.0]. Therefore, we only need to compute another point on it to construct the whole rate curve. Let the distribution of the DCT coefficients be $\mathcal{D}(x)$. It can be approximated by the histogram of the integer parts of the DCT coefficients, denoted by $\mathcal{H}(n)$. For any given quantization stepsize q_0, the corresponding percentage of zeros is given by

$$\rho_0 = \frac{1}{M} \sum_{|n|<0.5q_0} \mathcal{H}(n). \tag{1}$$

According to the definition of Q_{nz}, we have

$$Q_{nz}(\rho_0) = \sum_{|n|\geq 0.5q_0} \mathcal{H}(n)\,[\,\log_2 \lfloor \frac{|n|}{q_0} + 0.5 \rfloor + 1]. \tag{2}$$

The slope of the rate curve $Q_{nz}(\rho)$ is then given by

$$\theta = \frac{Q_{nz}(\rho_0)}{1-\rho_0}. \tag{3}$$

Estimate $Q_z(\rho)$

We surprisingly discover that there is a strong linear correlation between θ and the function value of $Q_z(\rho)$. To show this, we watch the values of $Q_z(\rho)$ at $\rho_i = 0.6$, 0.7, 0.75, 0.80, 0.85 and 0.90 for all the 24 sample images. For each ρ_i, there are 24 pairs of $[\theta,\ Q_z(\rho_i)]$ which are plotted in Fig. 4. The correlation coefficients between θ and $Q_z(\rho_i)$ are –0.52, –0.87, –0.87, –0.88, –0.84 and –0.81, respectively. This implies there is a strong linear correlation between $Q_z(\rho_i)$ and the slope θ. Hence, the following linear model

$$Q_z(\rho_i) = A_i\theta + B_i \tag{4}$$

is employed to estimate $Q_z(\rho_i)$. The coefficients A_i and B_i are obtained by statistical regression. The linear models are also shown in Fig. 4. The above linear correlation exists in our extensive simulations over a wide range of images.

Rate Curve Decomposition

Let the actual rate-quantization (R-Q) curve be $R(q)$. We map it into the ρ-domain and denote it by $R(\rho)$. In our rate curve decomposition scheme, we approximate $R(\rho)$ by a linear combination of the two characteristic rate curves,

$$\hat{R}(\rho) = \Gamma_1(\rho) \cdot Q_{nz}(\rho) + \Gamma_2(\rho) \cdot Q_z(\rho) + \Gamma_3(\rho). \tag{5}$$

where $\{\Gamma_k(\rho)\}$ are chosen to minimize the L_∞ norm

$$||\hat{R}(\rho) - R(\rho)||_\infty = \max_{0\le\rho\le1} |\hat{R}(\rho) - R(\rho)|. \tag{6}$$

In general, the rate curve is smooth. To simply the computation, we only need to determine the values of $\Gamma_k(\rho)$ at $\rho_i = 0.6$, 0.7, 0.75, 0.80, 0.85 and 0.90 such that $|\hat{R}(\rho_i) - R(\rho_i)|$ is minimized. This is a least mean square (LMS) problem. For each sample image, we already have $Q_{nz}(\rho_i)$, $Q_z(\rho_i)$, and the actual coding bit rate $R(\rho_i)$. $\{\Gamma_k(\rho_i)\}$ is then obtained by solving the following linear regression equation

$$R(\rho_i) = \Gamma_1(\rho_i) \cdot Q_{nz}(\rho_i) + \Gamma_2(\rho_i) \cdot Q_z(\rho_i) + \Gamma_3(\rho_i). \tag{7}$$

In our extensive simulations, the probability of the relative approximation error produced by Eq. (5) or (7) being less than 5% is 0.99. This implies it is very accurate to approximate $R(\rho)$ by a linear combination of the characteristic rate curves.

R-D Curve Estimation Algorithm

Based on the above discussions, the R-D curve estimation algorithm turns out to be very simple. After DCT, we compute the distribution $\mathcal{H}(n)$ of the coefficients. With Eqs. (1), (2), and (3), the slope of $Q_{nz}(\rho)$ is obtained. Based on Eq. (4), six points on $Q_z(\rho)$ are computed. The actual rate is then estimated by Eq. (7). With the mapping between q and ρ, the R-Q curve $R(q)$ can be obtained. Also, the distortion for each quantization step size q can be computed directly from the distribution $D(x)$.

Experimental Results

The rate curve estimation algorithm proposed in this work is tested in the following experiments. The four test images are shown in Fig. 5. The actual JPEG R-D curve and the estimated one for each test images are plotted in Fig. 6. It can be seen that these two are very close to each other. The relative estimation error is less than 5%. In motion JPEG video coding, based on the estimated R-D curve, we can accurately control the bit rate of each video frame. The rate control results for Foreman, Coastguard, Carphone, and News are shown in Fig. 7. It can be seen that the frame bit rate is very close to the target rate. With the estimated R-D curve, we can also control the picture quality. The quality control results for Foreman and Coastguard are shown in Fig. 8. The target picture quality is well matched. Our simulation results show that the proposed source model is very accurate. The rate and picture quality control algorithm is very efficient. Since TMN8 [4] or VM7 [3] can not model and estimate the R-D curve accurately, they can not perform this kind of rate control for the introcoded frames, especially in active videos with frequent scene changes.

References

[1] G. K. Wallace, "The JPEG still picture compression standard," *Commun. ACM*, vol. 34, pp. 30–44, April 1991.

[2] H. Gish and J. N. Pierce, "Asymptotically efficient quantizing," *IEEE. Trans. on Inform. Theory*, vol. IT-14, pp. 676 – 683, September 1968.

[3] T. Chiang, Y. -Q. Zhang, "A new rate control scheme using quadratic rate distortion model," *IEEE Transactions on Circuits and Systems for Video Technology*, vol.7, pp. 246 – 250, February 1997.

[4] J. Ribas-Corbera and S. Lei, "Rate control in DCT video coding for low-delay communications," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 9, pp. 172 – 185, February 1999.

[5] W. Ding and B. Liu, "Rate control of mpeg video coding and recording by rate-quantization modeling," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 6, pp. 12–20, February 1996.

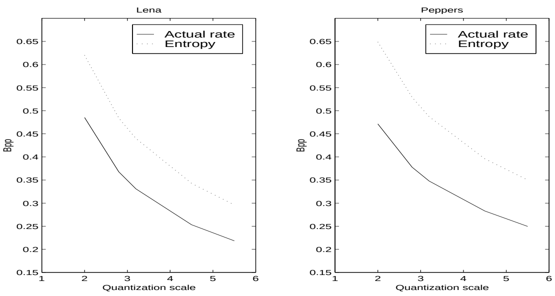

Figure 1: The theoretical entropy and the actual JPEG coding bit rate for Lena and Peppers.

Figure 2: The 24 sample images.

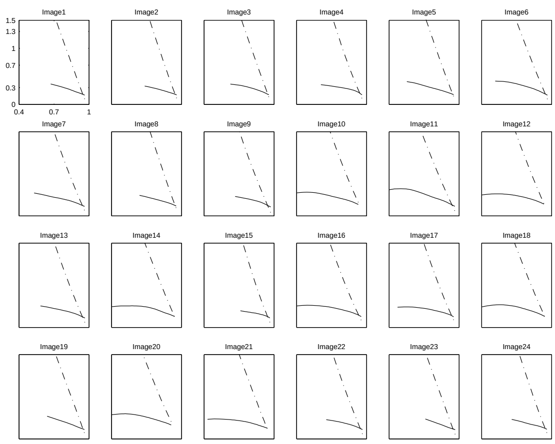

Figure 3: The plots of $Q_{nz}(\rho)$ and $Q_z(\rho)$ for the 24 sample images. The x-axis is ρ. The top curve $Q_{nz}(\rho)$ while the bottom one is $Q_z(\rho)$.

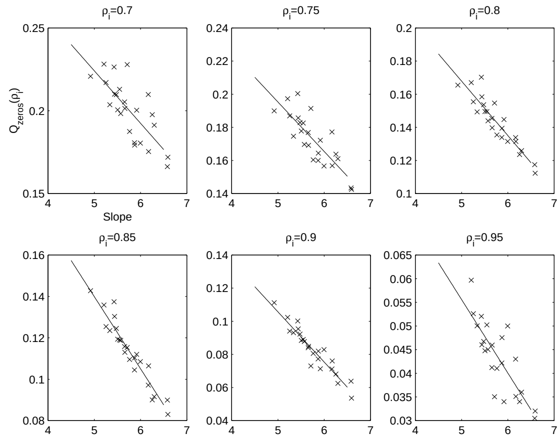

Figure 4: The linear correlation between θ and $Q_z(\rho_i)$ at 0.70, 0.75, 0.80, 0.85, 0.90 and 0.95.

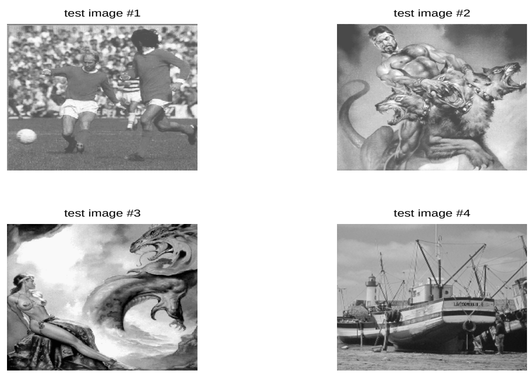

Figure 5: The four images for testing the performance of the proposed algorithm.

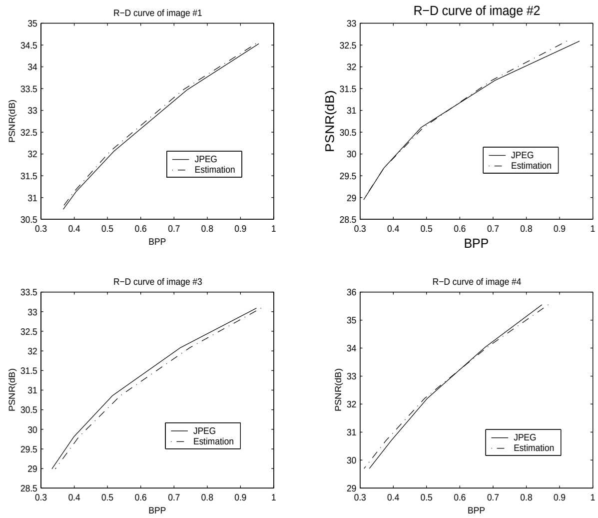

Figure 6: The estimated R-D curves and the real JPEG rate curves of the four test images.

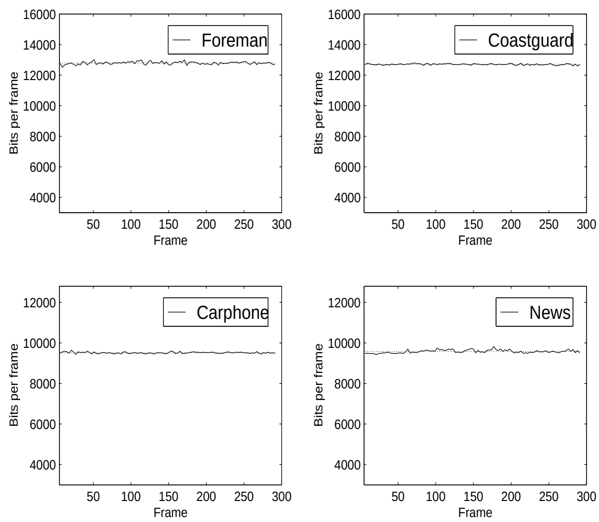

Figure 7: The bits per frame with rate control for Foreman at 384000 bps, Coastguard at 384000 bps, Carphone at 288000 bps and News at 288000 bps. The frame rate is 30 fps.

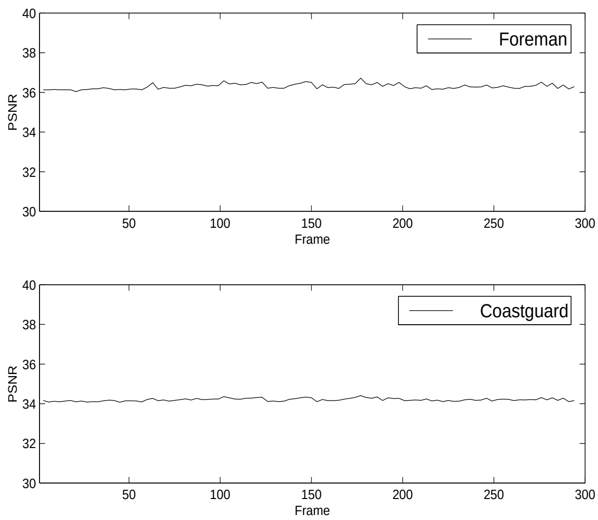

Figure 8: The PSNR for each frame with picture quality control for Foreman and Coastguard. Their target picture quality (PSNR) are 36 dB and 34 dB, respectively.