ρ-DOMAIN SOURCE MODELING AND RATE CONTROL FOR VIDEO CODING AND TRANSMISSION

Zhihai He, Yong Kwan Kim, and Sanjit K. Mitra

Department of Electrical and Computer Engineering
University of California, Santa Barbara, CA 93106
{zhihai, ykkim, mitra}@iplab.ece.ucsb.edu

ABSTRACT

In this work, the coding bit rate R is considered as a function of ρ which is the percentage of zeros among the quantized DCT coefficients. We discover that the rate function $R(\rho)$ has some very interesting properties in the ρ-domain. By introducing the new concepts of *characteristic rate curves* and *rate curve decomposition*, a novel framework for source modeling is proposed. Using the proposed source model, we can estimate the rate-quantization (R-Q) curve before quantization and coding with relative error less than 5%. Based on the estimated R-Q curve, the output bit rate of the video encoder can be accurately controlled. Our extensive simulation results show that the proposed algorithm outperforms TMN8 [1] and VM7 [5] rate control algorithms by providing more accurate and robust rate control.

1. INTRODUCTION

In real-time low-delay visual communication, an efficient rate control algorithm at the encoder is very important to assure the successful transmission of the coded video data [1]. The output bit rate of the video encoder is controlled by adjusting its quantization stepsize. The core problem in rate control is to model and estimate the R-Q behavior of the video encoder.

The classical rate-distortion (R-D) formula [2] does not work for transform coding at low bit rates. Several source models and rate control algorithms have been proposed in the literature [1, 4, 5]. These algorithms either have very high computational complexity [4], or suffer from large rate estimation or control error due to their inaccurate source models, especially at scene changes [1, 5].

This work was supported by a University of California MICRO grant with matching support from Lucent Technologies, National Semiconductor, Tektronix Corporation, and Xerox Corporation

To our best knowledge, all the source models proposed in the literature [1, 4, 5] consider the bit rate R as a function of the quantization stepsize q. In another word, the rate function is developed in the q-domain. In this work, we observe that the percentage of zeros among the quantized transform coefficients, denoted by ρ, has a very critical effect on the coding bit rate R. Note that ρ monotonically increases with q, which implies there is a one-to-one mapping between them. Therefore, R is also a function of ρ, denoted by $R(\rho)$. In this work, we propose to study the rate function in the ρ-domain instead of the q-domain. We discover that in the ρ-domain, the rate function $R(\rho)$ has some very interesting properties. Based on these properties, a novel source model for DCT video coding is proposed.

It is well known that coding performance depends on both the characteristics of the input source data and the coding algorithm. In the proposed source model, we define two characteristic rate curves $Q_{nz}(\rho)$ and $Q_z(\rho)$ to characterize the DCT coefficients. To model the coding algorithm, we introduce the new concept of rate curve decomposition, where the actual rate curve in the ρ-domain is represented by a linear combination of $Q_{nz}(\rho)$ and $Q_z(\rho)$. Based on the proposed source model, a low-delay rate control algorithm for DCT video coding is proposed. Our extensive simulation results show that the proposed algorithm outperforms TMN8 [1] and VM7 [5] rate control algorithms by providing more accurate and robust rate control.

2. CHARACTERISTIC RATE CURVES

In this section, we define two characteristic rate curves for the DCT coefficients. They are defined during the following "pseudo-coding" process. First, the DCT coefficients are uniformly quantized with a 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 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. Likewise, let Q_{nz} be the sum of the sizes of all non-zero coefficients. Obviously, both Q_{nz} and Q_z are functions of q. As discussed in Section 1, they are also functions of ρ, denoted by $Q_{nz}(\rho)$ and $Q_z(\rho)$, respectively. $Q_{nz}(\rho)$ and $Q_z(\rho)$ are called the characteristic rate curves. Following the above procedures, we can generate these two curves for the DCT coefficients of any input picture.

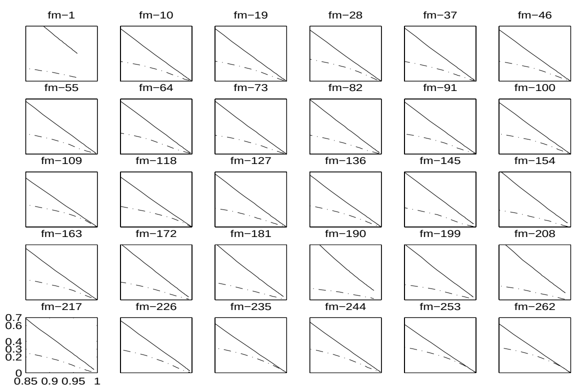

Figure 1: The plots of $Q_{nz}(\rho)$ (solid line) and $Q_z(\rho)$ (dotted line) for the 30 sample difference pictures from Foreman. All the subplots have the same coordinate system as the one at the bottom-left corner.

3. PROPERTIES OF THE CHARACTERISTIC RATE CURVES

Next we study the statistical properties of these two rate curves. To this end, we run the H.263 video encoder [3] on Foreman at $QP = 8$ to generate 30 sample motion compensated difference pictures. The sample difference pictures are taken at every 9 frames. We plot $Q_{nz}(\rho)$ and $Q_z(\rho)$ for each sample picture in Fig. 1. Two observations can be made on these plots. First, the two rate functions of each sample picture have almost the same pattern although the pictures are quite different from each other. Second, $Q_{nz}(\rho)$ is almost a straight line passing through [1.0, 0.0]. (When ρ is 1.0, Q_{nz} becomes zero because all the coefficients are zeros.) Let the slope of the straight line be θ. According to our extensive simulations on various video sequences, surprisingly, we find out that these two observations always hold [1]. This makes it possible for us to develop an accurate and robust source model in the following sections.

[1]Due to the page limitation, we can not list all the simulation results and give the theoretical explanation here.

4. ESTIMATION OF $Q_{NZ}(\rho)$ AND $Q_Z(\rho)$

Since $Q_{nz}(\rho)$ is modeled as a straight line passing through [1.0, 0.0], we only need to compute another point on it to determine the straight line $Q_{nz}(\rho)$. Let the distribution of the DCT coefficients be $\mathcal{D}(x)$. For any quantization stepsize q, after uniform quantization, the corresponding percentage of zeros is given by

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

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

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

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

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

In Section 3, we observe that for different pictures $Q_z(\rho)$ has almost the same pattern. Therefore, in the proposed source model, we approximate it by a second order polynomial

$$Q_z(\rho) = \begin{cases} -1.0\rho^2 + 0.67\rho + 0.33 & (I-frame), \\ -5.0\rho^2 + 7.55\rho - 2.55 & (p-frame), \end{cases} \tag{4}$$

where $\rho \in [0, 1]$.

5. RATE CURVE DECOMPOSITION

In digital signal processing, we often represent a signal to be studied by a weighted sum of the basis signals with well-known properties. This method is referred as *signal analysis or decomposition*. In this paper, we apply this decomposition method to study the behavior of a video coding system.

Suppose $R(q)$ is the actual R-Q curve of the video encoder. Let us map it into the ρ-domain and denote it by $R(\rho)$. In our rate curve decomposition scheme, $R(\rho)$ is represented by a linear combination of $Q_{nz}(\rho)$ and $Q_z(\rho)$,

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

where $\{\Gamma_k(\rho)\}$ are chosen to minimize the approximation error. Instead of determining $\{\Gamma_k(\rho)\}$ for the whole range of ρ, we only determine their values at ρ_i = 0.89, 0.91, 0.92, 0.93, 0.94, 0.95, 0.96, 0.97 and

Table 1: The values of $\Gamma_1(\rho)$, $\Gamma_2(\rho)$ and $\Gamma_3(\rho)$ at ρ_i.

ρ_i	$\Gamma_1(\rho_i)$	$\Gamma_2(\rho_i)$	$\Gamma_3(\rho_i)$
0.89	0.6151	-0.2438	0.4505
0.91	0.4089	-0.2530	0.4601
0.92	0.3480	-0.1931	0.4249
0.93	0.2763	-0.1749	0.3956
0.94	0.2531	-0.1141	0.3408
0.95	0.2043	-0.0789	0.2945
0.96	0.2577	0.0882	0.2140
0.97	0.4762	0.3792	0.1145
0.98	0.4567	0.5018	0.0745

Table 2: Comparison of the number of frames skipped and average PSNR for TMN8 and the proposed algorithm in H.263.

Test Video	Frame Skipped		PSNR(dB)	
	This Work	TMN8	This Work	TMN8
Salesman	1	3	32.83	32.71
News	2	5	28.48	28.13
Carphone	1	6	29.72	29.38
Trevor	0	2	32.37	32.33

0.98. For each sample difference picture, we generate the data $\{Q_{nz}(\rho_i),\ Q_z(\rho_i),\ R(\rho_i)\}$. The values of $\Gamma_1(\rho_i)$, $\Gamma_2(\rho_i)$ and $\Gamma_3(\rho_i)$ are then obtained by solving the following linear regression equation

$$R(\rho_i) = \Gamma_1(\rho_i)Q_{nz}(\rho_i) + \Gamma_2(\rho_i)Q_z(\rho) + \Gamma_3(\rho_i). \quad (6)$$

The values of $\{\Gamma_k(\rho_i)\}$ are listed in Table 1. In the proposed source model, $\{\Gamma_k(\rho_i)\}$ describe the coding behavior of the data compression algorithm. According to our statistical analysis, the probability of the relative estimation error of Eq. (6) being less than 5% is 0.99. This implies the actual rate curve can be closely approximated by a linear combination of the characteristic rate curves.

6. RATE CONTROL ALGORITHM FOR DCT VIDEO CODING

Based on the proposed source model, a rate control algorithm for DCT video coding is proposed as follows.

Step 1. Compute the distribution $\mathcal{D}(x)$ of the DCT coefficients. With Eqs. (1), (2) and (3), compute the slope θ and construct the straight line $Q_{nz}(\rho)$. With Eq. (4), compute $\{Q_z(\rho_i)|1 \le i \le 9\}$. Using the rate curve decomposition formula Eq. (6), compute $\{R(\rho_i)\}$.

Step 2. Let the target bit rate of the current picture be R_t which excludes the coding bits for the motion vectors and the header information. Based on the estimated rate curve $\{R(\rho_i)\}$, the corresponding ρ_t is determined to achieve the target bit rate R_t.

Step 3. Based on the one-to-one mapping between ρ and q, the corresponding quantization step size q_t is determined. q_t is then applied to the current video frame. Note that in video coding the quantization parameter (QP) for each macroblock must be an integer between 1 and 31. However, q_t is a real number. To solve this problem, we let the QP of each macroblock vary between $[q_t - 2,\ q_t + 3]$ such that the mean QP for the whole frame is very close to q_t.

It can be seen that the proposed algorithm has very low computational complexity. The rate control is performed at the frame level instead of the macroblock level as in [1].

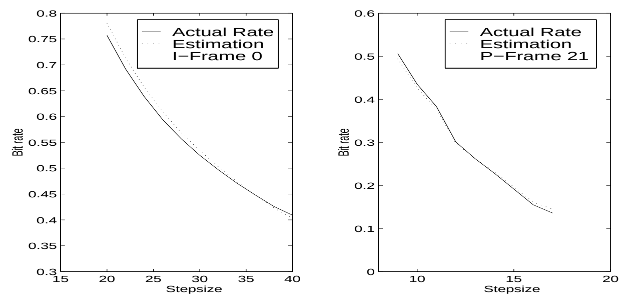

Figure 2: The R-Q curve estimation results for frames No. 0 (I-frame, left) and No. 21 (P-frame, right) in Carphone.qcif.

7. EXPERIMENTAL RESULTS

The proposed rate control algorithm is implemented in the Telenor's H.263 codec and tested for various video sequences. The estimated R-Q curves for the first frame (I-frame) and the 21st frame (P-frame) from Carphone (in QCIF format) are plotted in Fig. 2. It can be seen that the estimated R-Q curves are very close the actual ones. The relative estimation error is less than 5%. Since TMN8 [1] significantly outperforms VM7 [5], in the following tests, we only need to compare the rate control performance of the proposed algorithm and TMN8. The four test videos are Salesman (at 32 kbps), News (at 24 kbps), Carphone (at 24 kbps) and Trevor (48 kbps). The frame rate is fixed at 10 fps. In Fig. 3 (A) – (D), we plot the number of bits in the encoder buffer when the proposed algorithm (solid) and TMN8 (dotted) are applied. It can be seen that the proposed rate control algorithm maintains a much steadier buffer level (or output bit rate) than TMN8. The frame skip number and the average PSNR of the luminance component for each test are listed in Table 2. It can be seen that the proposed algorithm has better PSNR performance due to its more robust rate control and less skipped frames.

Let Ω_i be the dynamic range of the macroblock quantization parameters in the i-th video frame. In the proposed algorithm, we always have $\Omega_i \leq 6$. In TMN8, due to their inaccurate source model, the QP of each macroblock is adaptively modified according to the bits budget. In Fig. 4, we plot Ω_i for each video frame of Carphone when TMN8 (dotted) and the proposed rate control algorithm (solid) are applied. It can be seen that in TMN8 Ω_i can be as large as 30. One disadvantage of very large Ω_i is that some visually important information maybe lost due to the very coarse quantization. However, in the proposed algorithm, the rate curve can be accurately estimated. The mean quantization parameter is determined for each frame. The quantization parameter for each macroblock is regulated within a reasonably small range, which guarantees a uniform picture quality. This is a unique feature of the proposed rate control algorithm.

8. CONCLUDING REMARKS

There are two major contributions in this work. First, we have developed a novel source modeling technique to study the coding behavior of the DCT video coding system. Second, a very efficient rate control algorithm for DCT video coding has been developed. Despite its very low computational complexity, the proposed algorithm provides robust and accurate rate control for video coding and transmission.

9. REFERENCES

[1] 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.

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

[3] ITU-T, "Video coding for low bit rate communications," *ITU-T Recommendation H.263*, version 1, Nov. 1995, version 2, January 1998.

[4] 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.

[5] 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.

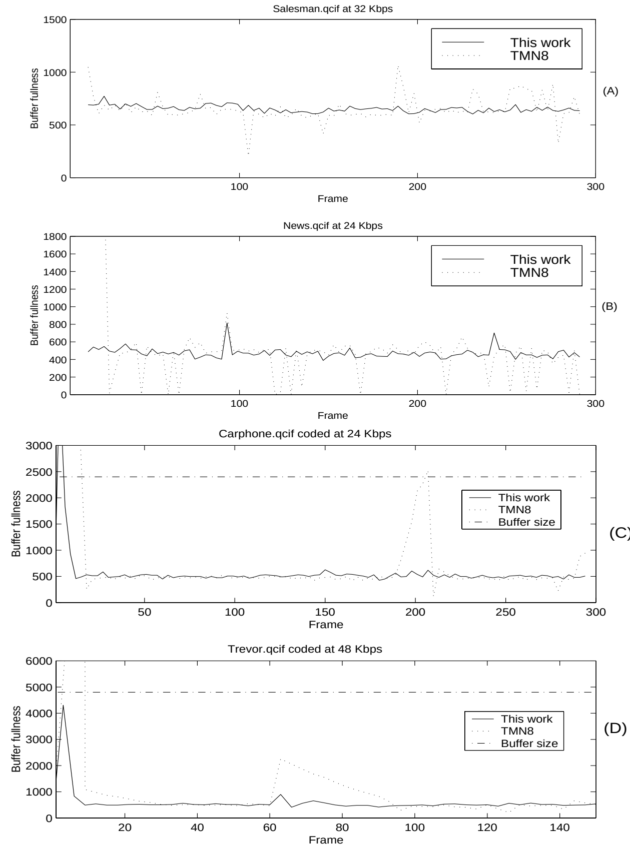

Figure 3: Comparison of the buffer fullness between TMN8 and the proposed rate control algorithm in H.263 for four test video sequences.

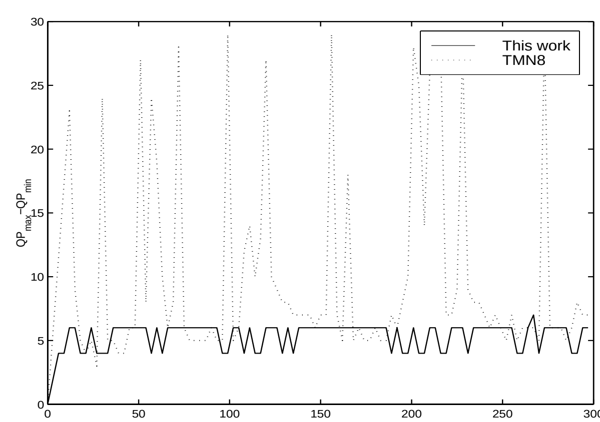

Figure 4: Dynamic range of the quantization parameters for Salesman at 32 kbps.