# A LOW COMPLEXITY RATE-DISTORTION SOURCE MODELING FRAMEWORK

*Chun-Yuan Chang*<sup>1</sup>, *Tsungnan Lin*<sup>2</sup>, *Din-Yuan Chan*<sup>3</sup> and *Shih-Hao Hung*<sup>1</sup>

<sup>1</sup>Graduate Institute of Networking and Multimedia, National Taiwan University <sup>2</sup>Graduate Institute of Communication Engineering, National Taiwan University <sup>3</sup>Department of Computer Science and Information Engineering, National Chiayi University tsungnan@ntu.edu.tw

### ABSTRACT

An accurate rate-distortion model, which characterizes the relationship among bitrate, distortion, and quantization parameter (QP), is very desirable for real time video transmission. It has been reported that the actual coding bitrates can be estimated by a linear combination of two characteristic rate curves in  $\rho$ -domain where  $\rho$ is defined as the percentage of zeros among the quantized transform coefficients. The process is referred as the "pseudocoding" process. Unfortunately, since  $\rho$  values are real numbers, an interpolation process must be required for the domain transformation from  $\rho$ -domain to q-domain. Thus, the prediction inaccuracy can not be avoided and even be "propagated" due to the interpolation uncertainty error. Hence, three parameters are addressed in this paper to support a more accurate and direct estimate of encoding bit rates based on q-domain They include the number of nonzero coefficients, the count of zeros before the last nonzero coefficient in the zigzag-scan order, and the sum of absolute quantized nonzero coefficients, respectively. We surprisingly find that the estimation accuracy in q-domain is better than currently well-known  $\rho$ -domain based R-Q model. In addition, a quantization-free extraction method, which only involves some additions and a few multiplications, is developed. That is, the implementation complexity of the proposed mechanism is very low. Consequently, the proposed R-Q model is very suitable for real time applications.<sup>1</sup>

### **1. INTRODUCTION**

Rate control scheme plays a significant role in real time video transmission. It regulates the quantization parameter (QP) of either frame or macroblock (MB) layer to achieve better picture quality under limited bandwidth. Therefore, an accurate rate-distortion model, which characterizes the relationship among bitrate, distortion, and quantization, is very desirable. Some R-D models [1]-[5] derive a set of mathematical formulas to estimate the R-D model according to entropy theory. These source models suffer from large estimation error, especially in low-motion/activity cases. Even in high-motion/activity cases, there are still large mismatches between theoretical entropy and actual coding bitrate with increase of QP [8]. Those large estimation errors are owing to the insufficient expression of the distribution of quantized DCT coefficients. Other researchers [6], [7] preencode the same frame

for several times to construct the whole R-D model. Obviously, these two models have very high computational complexity and can not be used in real time applications. Kim et al. [8] define  $\rho$  as the percentage of zeros among the quantized transform coefficients and discover that there is a linear relationship between  $\rho$  and the actual coding bitrate R. Subsequently, He et al. [9] define two characteristic rate curves to express the distribution of quantized DCT coefficients in  $\rho$ -domain. Actual coding bitrates R can be estimated by a linear combination of two characteristic rate curves in  $\rho$ -domain, which is regarded as a "pseudocoding" process. The whole R-Q model is then interpolated from some sampled  $R-\rho$ pairs. Since  $\rho$  is over real domain, the interpolation process results in error propagation during the domain transformation from  $\rho$ domain to q-domain. To avoid the phenomenon of error propagation, we begin to study whether the concept of "pseudocoding" can be also applied to q-domain as the appropriate characteristic rate curves are chosen. In this paper, we propose a new combination of characteristic rate curves to parameterize the distribution of quantized DCT coefficients. The parameters we used are the number of nonzero coefficients [12], the sum of zeros before the last nonzero coefficient after zigzag scan and quantization [9]-[11], and the sum of absolute quantized nonzero coefficients, respectively. We surprisingly find that the estimation accuracy using this combination to construct R-O model in q-domain is better than current well-known  $\rho$ -domain based R-Q model. In addition, a quantization-free extraction method, which only involves some additions and a few multiplications, is developed. That is, the implementation complexity of the proposed mechanism is very low. Consequently, the proposed R-Q model is very suitable for real time applications. This paper is organized as follows. In Sec. 2, we first briefly review  $\rho$ -domain based R-Q model. Then, in Sec. 3, we will propose a new source modeling framework in q-domain and present a novel extraction mechanism. Finally, in Sec. 4, extensive experimental results show that proposed R-Q model performs better than  $\rho$ -domain based R-Q model in terms of the accuracy.

### **2.** BRIEF REVIEW OF $\rho$ - DOMAIN R-Q MODEL

Before introducing R-Q modeling in  $\rho$ -domain, we first review the quantizer in video coding systems. MPEG-2, MPEG-4 and H.263+ adopt uniform threshold quantizer (UTQ) to compress a huge amount of video data. Each DCT coefficient *x* can be quantized by a piecewise definition:

<sup>&</sup>lt;sup>1</sup> This work was supported in part by Taiwan National Science Council under grant 94-2213-E-002-066 and 94-2219-E-002-017

$$C(x) = \begin{cases} 0, & \text{if } |x| \le \Delta; \\ \left\lceil \frac{x - \Delta}{q} \right\rceil, & \text{if } x > \Delta; \\ \left| \frac{x + \Delta}{q} \right|, & \text{if } x < -\Delta. \end{cases}$$
(1)

ſ

, where C(x) is the quantized coefficients,  $\Delta$  is the dead zone threshold and q is the quantization step size. In this study, we will take H.263+ as an example. Hence, dead zone threshold  $\Delta$  is set to 2q and 2.5q for intra- and inter-macroblock, respectively. Quantization step is set to 2q. Consequently,  $\rho$ -"the percentage of zeros among the quantized transform coefficients" in [8] can be defined as follows:

$$\rho(q) = \frac{1}{M} \sum_{|x| < \Delta} D_F(x) \qquad (2)$$

, where  $D_F(x)$  is the histogram of DCT coefficients in frame-layer. M is total pixels in a frame. Following, we will summarize  $\rho$ domain based R-Q model from the concept of "pseudocoding" to the linear interpolation. First, two quantized characteristic rate curves,  $Q_{nz}(\rho)$  and  $Q_{z}(\rho)$ , [9] are used to estimate the actual coding bitrate  $R(\rho)$  in  $\rho$ -domain by following linear combination:

$$(\rho) = A(\rho) \cdot Q_{nz}(\rho) + B(\rho) \cdot Q_z(\rho) + C(\rho) \quad (3)$$

, where  $Q_{nz}(\rho)$  and  $Q_{z}(\rho)$  are defined as  $\frac{1}{M} \cdot \sum_{C(x)} (\lfloor \log_2 C(x)_{LEVEL} \rfloor + 2)$  and  $\frac{1}{M} \cdot \sum_{C(x)} C(x)_{RUN}^2$ , respectively. Then,

through the observation of the linearity between R and  $\rho$ , an interpolated R-Q model [9]-[11] is constructed.

#### **Pseudocoding:**

Step1(a): Quantize the frame once to construct the function:

 $\hat{Q}_{nz}(\rho) = \kappa \cdot (1 - \rho)$ (4) ,where  $\kappa = Q_{nz}(\rho(q))/(1-\rho(q))$  and  $\rho(q) \in [0.9, 0.95]$ 

**Step1(b):** Comput  $\hat{Q}_{nz}(\rho_i)$  for each  $\rho_i$  using the function  $\hat{Q}_{nz}(\rho)$  in Eq. (4).

**Step1(c):** Compute  $\hat{Q}_{r}(\rho_{i})$  for each  $\rho_{i}$  by

$$\hat{Q}_{z}(\rho_{i}) = a_{3}(\rho_{i}) \cdot \kappa^{3} + a_{2}(\rho_{i}) \cdot \kappa^{2} + a_{1}(\rho_{i}) \cdot \kappa + a_{0}(\rho_{i})$$
(5)

,where  $a_3(\rho_i)$ ,  $a_2(\rho_i)$ ,  $a_1(\rho_i)$  and  $a_0(\rho_i)$  are weighting coefficients obtained by statistical linear regression.

**Step1(d):** Compute  $\hat{R}(\rho_i)$  for each  $\rho_i$  by

$$R(\rho_i) = A(\rho_i) \cdot Q_{nz}(\rho_i) + B(\rho_i) \cdot Q_z(\rho_i) + C(\rho_i)$$
(6)

, where  $A(\rho_i)$ ,  $B(\rho_i)$  and  $C(\rho_i)$  are weighting coefficients obtained by statistical linear regression.

## From *o*-domain to *a*-domain:

**Step2**: By the assumption of linearity between R and  $\rho$  in Eq. (7), the whole R-Q model can be constructed through a linear rate regulation interpolation [10]

$$R(\rho) = \theta \cdot (1 - \rho) \tag{7}$$

In [9],[11], it has been presented that the relative estimation error at given  $\rho_i$  is always less than 5%. It is very accurate. However,  $\rho$ is over real domain. Therefore, it is difficult to obtain all weighting coefficients in Eq. (5) and (6) for all  $\rho_i$  over  $\rho$ . Consequently, it inspires us to study how to apply the concept of "pseudocoding" from  $\rho$ -domain to q-domain. We believe there is still a large of improvement space in estimation accuracy and computational complexity, respectively. In the next section, we will study feasibility of "pseudocoding" in q-domain and propose our source modeling framework.

#### 3. PROPOSED SOURCE MODELING FRAMEWORK

#### 3. 1 Analysis of characteristic rate curves in q-domain

In this subsection, we try to find an appropriate combination of characteristic rate curves in q-domain to express the distribution of quantized DCT coefficients. Most video coding standards, such as MPEG-2, MPEG-4 and H.263+ adopt run length coding (RLC) to encode quantized DCT coefficients. In H.263+, a 3-dimension lookup table (LAST, RUN, LEVEL) is used. It is obvious that the level of quantized nonzero coefficient and its run zeros have a direct influence on the actual coding bitrate R. In [12], it is reported that there is a linear relationship between R and the number of nonzero coefficient in smaller QP (in high bitrate situations). However, in larger QP cases, the accuracy [12] is not satisfactory. To propose a robust R-Q model which works both for large and small QP, we exploit the influence of the number of nonzero coefficient  $Q'_{c}(q)$ , the level of quantized nonzero coefficient  $Q'_{L}(q)$  and run zeros  $Q'_{Z}(q)$  on the bitrate. They are defined as follows:

$$Q_{C}(q) = \frac{1}{M} \cdot Q'_{C}(q) = \frac{1}{M} \cdot \sum_{C(x)} n \text{, where } \begin{cases} n = 1, \text{ if } C(x)_{LEVEL} \neq 0\\ n = 0, \text{ if } C(x)_{LEVEL} = 0 \end{cases}$$
$$Q_{L}(q) = \frac{1}{M} \cdot Q'_{L}(q) = \frac{1}{M} \cdot \sum_{C(x)} C(x)_{LEVEL}$$
$$Q_{Z}(q) = \frac{1}{M} \cdot Q'_{Z}(q) = \frac{1}{M} \cdot \sum_{C(x)} C(x)_{RUN} \tag{8}$$

In Fig. 1, we plot three characteristic rate curves  $Q_{C}(q)$ ,  $Q_{L}(q)$  and  $Q_{Z}(q)$  respectively for different compensated frames. We find that the behaviors of  $Q_{C}(q)$  and  $Q_{L}(q)$  are very similar to the behavior of R(q). To further exploit their relationship with respect to R, we calculate relative correlation coefficients of  $\{(R,Q_C)\}, \{(R,Q_L)\}$  and  $\{(R,Q_Z)\}$  on average through extensive experimental data in Table. 1. Their correlation coefficients are very close to 1. It is very interesting that the correlation coefficients of  $\{(R,Q_C)\}, \{(R,Q_L)\}$ are always larger than 0.93, and correlation coefficients of  $\{(R, Q_Z)\}$ are getting larger with increase of QP. It implies that run zeros play a significant role in larger QP cases. Based on the observation, we hypothesize that R(q) could be expressed as a linear combination of  $Q_{C}(q)$ ,  $Q_{L}(q)$  and  $Q_{Z}(q)$ . That is, the "pseudocoding" phenomenon [9] can be described in q-domain as follows:

$$\overline{R}(q) \approx A(q) \cdot Q_C(q) + B(q) \cdot Q_L(q) + C(q) \cdot Q_Z(q) + D(q).$$
(9)

, where A(q), B(q), C(q) and D(q) are weighting coefficients obtained through extensive statistical regression. Surprisingly, we find that the average estimation error is about 3.6% (detailed results are shown in Sec. 4).



Fig.1 Plots of  $R(q), Q_{C}(q), Q_{L}(q)$  and  $Q_{Z}(q)$  for different compensated frames.

<sup>&</sup>lt;sup>2</sup> H.263+ coding algorithm uses 3-dimension variable coding table to encode (LAST, RUN, LEVEL) of each code word C(x).  $C(x)_{LEVEL}$  and  $C(x)_{RUN}$  stand for the absolute value and the number of run zeros of the codeword x, respectively.

#### 3.2 Fast extraction method

Additionally performing quantization processes to extract characteristic rate curves needs a large of division and ceiling operations such as Eq. (1). It will burden the video coding system. However,  $\rho$ -domain based R-Q model performs an extra quantization process and an amount of logarithm operations in "pseudocoding" according to procedures in Sec. 2. Therefore, for a frame of M pixels, M division and ceiling operations are needed. In the following, we will present a fast quantization-free extraction mechanism by an appropriate calculation order to further reduce the computational complexity for Eq. (9). We first compute  $Q'_{c}(q)$  and  $Q'_{z}(q)$  from the viewpoint of a 8×8 block (denoted as  $Q'_{_{C_{_{B}}}}(q)$  and  $Q'_{_{Z_{_{B}}}}(q)$ , respectively ) and then  $Q'_{_{L}}(q)$  can be obtained efficiently after  $Q'_{C}(q)$  is calculated. In addition, the calculation is performed recursively from QP=1 to QP=31. During the recursive calculation, only additive operations are involved. Therefore, the complexity is redued substantially.

After arranging DCT coefficients into 1-D array in zigzag scan order, we store all nonzero coefficients  $x_j$  and its relative positions  $P_j$  into another array *B* as Fig. 2.



Fig. 2 Status of non-quantized coefficients and the position of last nonzero coefficients after applying dead zone thresholding  $\Delta_{i}$ .

Then, the histogram  $D_B(x)$  of all nonzero coefficients in the array *B* can be construed. Hence, the number of quantized nonzero coefficients in a 8×8 block can be computed through integrating  $D_B(x)$  in the range of  $|x| \ge \Delta_i$ .

$$Q'_{C_B}(q_i) = \sum_{|x| \ge \Delta_i} D_B(x) \qquad (10)$$

However, to compute Eq. (10) directly will result in the overlap integration problem and the redundant complexity is high. However, if we rearrange the calculation as the recursive formula of Eq. (11) to take advantage of previously computed results, the overlap integration problem can be avioded.

$$Q'_{C_{B}}(q_{1}) = \sum_{|x| \ge \Delta_{1}} D_{B}(x)$$
$$Q'_{C_{B}}(q_{i}) = Q'_{C_{B}}(q_{i-1}) - \sum_{\Delta_{i-1} \le |x| \le \Delta_{i}} D_{B}(x) \quad (11)$$

During the operation, the absolute values of these non-quantized nonzero coefficients in a 8×8 block (denoted as  $Q'_{NNZ_{B}}(q_{i})$ ) is also

computed in order to calculate  $Q'_{L}(q)$  later.

$$\begin{aligned} \mathcal{Q'}_{NNZ_{B}}\left(q_{1}\right) &= \sum_{|x| \geq \Delta_{1}, D_{B}\left(x\right)} |x| \\ \mathcal{Q'}_{NNZ_{B}}\left(q_{i}\right) &= \mathcal{Q'}_{NNZ_{B}}\left(q_{i-1}\right) - \sum_{\Delta_{i-1} \leq |x| \leq \Delta_{i}} D_{B}\left(x\right) \end{aligned} \tag{12}$$

When the position  $P_{Last_i}$  - the last nonzero coefficient after applying dead zone thresholding  $\Delta_i$  to array *B* is know, the sum of run zeros of nonzero coefficients (codewords) can be easily calculated as:

$$Q'_{Z_{B}}(q_{i}) = p_{Last_{i}} - Q'_{C_{B}}(q_{i})$$
 (13)

The information of  $P_{Last_i}$  in *q*-domain can be easily obtained for successive dead zone thresholding from  $\Delta_I$  to  $\Delta_{3I}$  since  $P_{Last_i}$  moves in the inverse zigzag order. As Fig. 2, we can see that after applying dead zone thresholding  $\Delta_I$ ,  $P_{Last}$  is moved from  $P_{10}$  to  $P_6$ . After calculating the results from Eq. (11) to (13) of all blocks, we can sum up these results to obtain  $Q'_C(q_i)$ ,  $Q'_Z(q_i)$  and  $Q'_{NNZ}(q_i)$ . Then, we will utilize the results of  $Q'_C(q_i)$  and  $Q'_{NNZ}(q_i)$  to obtain sum of levels of quantized nonzero coefficients  $Q'_L(q)$ . Before calculating  $Q'_L(q)$ , we review the quantization process in Eq. (1). According the definition of  $Q'_L(q)$ , it should be calculate by

$$\sum_{i=1}^{N} \left\lceil \frac{|x_i - \Delta|}{q} \right\rceil.$$
(14)

However, for each item  $\frac{|x_i - \Delta|}{q}$  in Eq. (14), a dynamic value on

[0,1) is stuffed by a ceiling operation. Hence, by assuming the stuffing value is 0.5 on average, Eq. (14) can be approximated [13] by

$$\frac{\sum_{i=1}^{N} |x_i| - \Delta \cdot N}{q} + 0.5 \cdot N. \quad (15)$$

Consequently, we can adopt previously computing result  $Q'_{C}(q)$  and  $Q'_{NNZ}(q_{i})$  to compute  $Q'_{L}(q)$  according to Eq. (15) as following:.

$$Q'_{L}(q_{i}) \approx \frac{Q'_{NNZ}(q_{i}) - \Delta_{i} \cdot Q'_{C}(q_{i})}{2 \cdot q} + 0.5 \cdot Q'_{C}(q_{i}) \qquad (16)$$

Up to now, we do not perform any quantization process yet. We just operate on nonzero DCT coefficients with addition and threshoding operations. Unlike the procedures in  $\rho$ -domain, the framework presented to extract the parameters in *q*-domain did not require any extra quantization process and logarithm computations.

#### 4. EXPERIMENTAL RESULTS

The proposed R-Q model framework is implemented in H.263+ codec. The relative prediction error is defined as

$$E = \frac{\mid R_o - R_e \mid}{R_o} \qquad (17)$$

, where  $R_o$  and  $R_e$  are the actual coding bitrate and the estimated bitrate, respectively. In our experiments, we exclude all header information such motion vector, DC value and coded block pattern, etc, because their bitrates are almost fixed during rate control scheme. Various sequences such "foreman", "carphone", "container", "suize" and "salesman" in QCIF format are exploited as our test sequences. We encode every frame 31 times and in the meanwhile construct proposed R-Q model, performing TMN8 [1] rate control scheme. Encoding frame rate is set to 10 fps and the QP in first I frame is set to 13. In Table. 2, we can see that proposed R-Q model is very accurate and robust for various test sequence from high activity sequence "foreman" to low activity sequence "salesman". In Fig. 3, we also can see that the estimated bitrate curve is very close to the actual coding bitrate curve and

better than  $\rho$ -domain based R-Q model in terms of accuracy. Extensive experiments show that relative estimation error of proposed R-D model is about 3.6% on average. Consequently, "pseudocoding" concept is feasible in *q*-domain when the appropriate characteristic rate curves are chosen as above mentioned. And, the R-Q model directly constructed in *q*-domain will avoid error propogation (about 7% estimation error) due to domain transformation from  $\rho$ -domain to *q*-domain.

#### **5. CONCLUSION**

In this paper, we discover the "pseudocoding" phenomenon in qdomain can be also justified by a linear combination of appropriately chosen features. The parameters we used are the number of nonzero coefficients, the sum of zeros before the last nonzero coefficient after zigzag scan and quantization, and the sum of absolute quantized nonzero coefficients, respectively. We find that the estimation accuracy in q-domain has significant improvement over currently well-known  $\rho$ -domain based R-Q model. In addition, a quantization-free extraction method, which only involves some additions and a few multiplications, is developed. That is, the implementation complexity of the proposed mechanism is very low. Consequently, the proposed R-Q model is very suitable for real time applications.

### 6. REFERENCES

[1] J. Ribas-Corbera and S. Lei, "Rate control in DCT video coding for low-delay communications," IEEE Trans. Circuit Syst. Video Technol., vol. 9, no. 1, pp. 172-185, Feb. 1999.

[2] K. H. Yang, A. Jacquin, and N. S. Jayant, "Normalised rate-distortion model for quantisation in MC-DCT video coders," IEE Electronics Letter, vol. 33, no. 25, pp. 2099-2100, Dec, 1997.

[3] T. Chiang and Y.-Q. Zhang, "A novel rate control scheme using quadratic rate distortion model," IEEE Trans. Circuits Syst. Video Technol., vol. 7, no.1, pp. 246–250, Feb. 1997.

[4] J. Wei, B. H. Soong and Z. G. Li, "A New Rate-Distortion Model for Video Transmission Using Multiple Logarithmic Functions," IEEE Signal Processing Letters, vol. 11, no.8, pp. 694-697, Aug. 2004.

[5] H. –M Hang and J. –J. Chen, "Source model for transform video coder and its application-Part I : Fundamental theory," IEEE Trans. Circuit Syst. Video Technol., vol. 7, no. 2, pp. 287-298, Apr. 1997.

[6] L.-J. Lin and A. Ortega, "Bit-rate control using piecewise approximated rate-distortion characteristics," IEEE Trans. Circuits Syst. Video Technol., vol. 8, no. 4, pp. 446–459, Aug. 1998.

[7] W. Ding and B. Liu, "Rate control of MPEG video coding and recording by rate-quantization modeling," IEEE Trans. Circuit Syst. Video Technol., vol. 6, no. 1, pp. 12-20, Feb. 1996.

[8] Y. K. Kim, Z. He and S. K. Mitra, "A novel linear source model and a unified rate control algorithm for H.263/MPEG-2/MPEG-4," IEEE International Conference on Acoustics, Speech, and Signal Processing, Salt Lake City, UT, May 2001.

[9] Z. He and S. K. Mitra, "Low-Delay Rate Control for DCT Video Coding via *ρ*-domain Source Modeling," IEEE Trans. Circuits Syst. Video Technol., vol. 11, no. 8, pp. 928–940, Aug. 2001.

[10] Z. He and S. K. Mitra, "A Unified Rate-Distortion Analysis Framework for Transform Coding," IEEE Trans. Circuits Syst. Video Technol., vol. 11, no. 12, pp. 1221–1236, Dec. 2001.

[11] Z. He and S. K. Mitra, "Novel rate-distortion analysis framework for bit rate and picture quality in DCT visual coding," IEE Proc.–Vis. Image Signal Process., vol. 148, no. 6, pp. 398–406, Dec. 2001.

[12] T. Y. Kim and J. K. Kim, "An Accurate Rate Control of MPEG Video by Rate-Codewords Modeling," IEEE International Symposium on Circuits and Systems, Hong Kong, June 9-12, 1997.

[13] Y. Wang, J. Ostermann and Y. Q. Zhang,"Video Processing and Communications," Prentice-Hall, 2002, pp. 241-248.

Table. 1 Correlation coefficients on average for  $\{(R,Q_c)\}$ ,  $\{(R,Q_L)\}$ ,  $\{(R,Q_2)\}$  and  $\{(R,Eq.(9))\}$ , respectively.

	R,Q <sub>C</sub>	R,QL	R,Qz	R, Eq. (9)
Corrrelation coefficient	0.992	0.975	0.756	0.999

Table. 2 Estimation error on average for various sequences and QPs under 24K bps, performing TMN8 rate control.





Fig. 3 Bits number comparison among original H.263+ encoder actual coding bitrate, proposed model and  $\rho$ -domain based model for continuous p-frames of "suzie.qcif".