A q-Domain Characteristic-Based Bit-Rate Model for Video Transmission

Chun-Yuan Chang, *Student Member, IEEE*, Cheng-Fu Chou, Din-Yuen Chan, Tsungnan Lin, *Senior Member, IEEE*, and Ming-Hung Chen

***Abstract*—For low-delay video transmission, we introduce a q-domain characteristic-based bit-rate model. Specifically, three characteristics are efficiently extracted from the quantized DCT spectra to construct the bit-rate model. Extensive experimental results show that our rate model can provide more accuracy with lower complexity than existing models.**

***Index Terms*—Rate control, rate-quantization model.**

I. Introduction

THE rate control scheme, which adjusts the quantization parameters (QPs), plays an important role in packet video transmission since the communication channel often imposes stringent constraints on the transmission bandwidth. Thus, how to construct an accurate bit-ate model has become a major challenge for the rate control designer. In classical MB-level R-D models in [1] and [2], the only characteristic that describes the input source data is the variance of source input, but this approach cannot efficiently adapt a dramatic variation of input source. To tackle this issue, an improved variance-based R-Q model in [2] is proposed. However, a big bit-rate-estimating error might happen in low-motion or low-bit-rate cases. On the other hand, Kim *et al.* [3] have first used the number of nonzero quantized transform coefficients as the main characteristic, i.e., the number of codewords, to model the bit rate for the rate controller. Furthermore, the authors in [4] defined ρ as the percentage of zeros among the quantized transform coefficients and found that there is a linear relationship between R and ρ within each frame. The latter one suggests that the linearity between R and ρ within each frame can be used to model the curve $R(\rho)$ when the slope of $R(\rho)$ is predetermined. Accordingly, He *et al.* [4] attempt to compute some control points, which are regarded as pseudobit rates, to determine the slope of $R(\rho)$, that is, they collected extensive the actual bit-rate points of different frames and classified those bit-rate points according to different pseudo ρ. Then, within the same cluster of bit-rate points, the "pseudocoding" process[1] is applied to model the pseudobit rates $R(\rho)$. After that, we can apply the *linear rate regulation* approach [5] to the pseudo bit rates to determine the slope of $R(\rho)$. Finally, the curve $R(\rho)$ is obtained and the R-Q model is constructed by one-to-one mapping between ρ and q. Nevertheless, there are two main concerns for ρ-domain-based R-Q model:

1) if the accuracy of the "pseudocoding" process for a specified ρ is general for all different types of videos;
2) if the linearity between R and ρ could still hold on within each frame.

To cope with the above issues, in this letter we provide discussions and arguments about ρ-domain model and derive a new q-domain characteristic-based bit rate model. In addition, we also demonstrate a two-level greed-based rate controller [6] using the proposed rate model. Predictably, experimental results show that the improvement of accuracy of our rate model is able to meet rate control concerned targets, e.g., the ability of meeting target buffer delay and visual quality enhancement, better than existing rate controllers [1], [4].

II. From ρ-Domain Based R-Q Model To Characteristic-Based R-Q Model

A. Problems of ρ-Domain Based R-Q Model

In [4], He *et al.* took use of a finite set of ρ as control points to perform the "pseudocoding" process. For each ρ_j, there are two characteristics, i.e., the average of the sizes of all nonzero coefficients and the average of the sizes of all run length numbers, denoted as $Q_{NZ}(\rho_j)$ and $Q_Z(\rho_j)$, respectively; they are used to measure "the finite set of pseudo bit rates," which is $\Gamma = \{\hat{R}(\rho_j)|1 \le j \le 9\}$. Herein, we write the estimated "pseudobit rate" as follows:

$$\hat{R}(\rho_j) = \bar{W}_\rho(\rho_j) \cdot \bar{Q}(\rho_j) \tag{1}$$

where the characteristic vector is defined as $\bar{Q}(\rho_j) = [Q_{NZ}(\rho_j) Q_Z(\rho_j) 1]^T$, and $\bar{W}_\rho(\rho_j)$ is the set of model coefficients obtained through the offline regression method. Furthermore, two characteristics $Q_{NZ}(\rho_j)$ and $Q_Z(\rho_j)$ in (1) are computed by the linear and cubic function [4], respectively, as follows:

$$\hat{Q}_{NZ}(\rho) = \kappa \cdot (1 - \rho) \tag{2}$$

$$\hat{Q}_Z(\rho_j) = \bar{W}_Z(\rho_j) \cdot \bar{K} \tag{3}$$

Therefore, assuming that the linearity between R and ρ within the same frame holds, the ρ-domain based R-Q model is proposed to model the source bit rate, which is shown as follows:

$$\hat{R}(\rho) = \theta \cdot (1 - \rho). \tag{4}$$

Specifically, we can compute the finite set of pseudobit rates Γ by (1). Afterward, Γ is fed into a *linear rate regulation* [5] to determine θ_ρ. Accordingly, the $R(\rho)$ curve can be constructed using the estimated θ_ρ. Finally, the R-Q model is obtained via the one-to-one mapping between q and ρ.

Manuscript received August 8, 2006; revised March 4, 2007 and July 17, 2007. First published April 30, 2008; current version published October 8, 2008. This paper was recommended by Z. He.

C.-Y. Chang is with the Graduate Institute of Networking and Multimedia, National Taiwan University, Taipei 106, Taiwan, R.O.C.

C.-F. Chou and M.-H. Chen are with the Department of Computer Science and Information Engineering, National Taiwan University, Taipei 106, Taiwan, R.O.C. (e-mail: ccf@csie.ntu.edu.tw).

D.-Y. Chan is with the Department of Computer Science and Information Engineering, National Chayi University, Chayi 600, Taiwan, R.O.C.

T. Lin is with the Graduate Institute of Communication Engineering, National Taiwan University, Taipei 106, Taiwan, R.O.C.

Digital Object Identifier 10.1109/TCSVT.2008.924103

[1]The "pseudocoding" process is to extract some useful characteristics from frequency domain to estimate the actual bit rate.

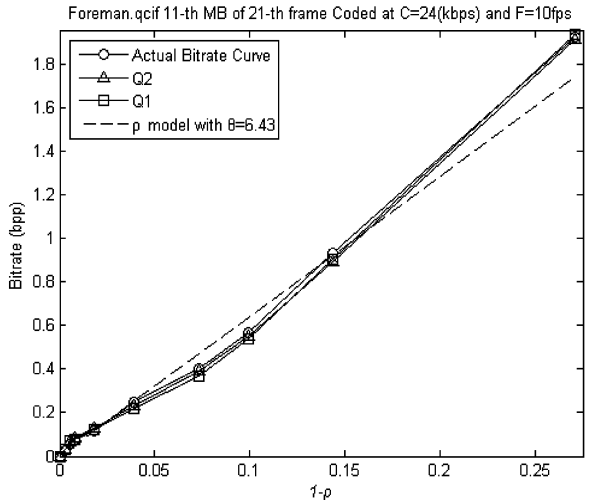

Fig. 1. Relationship $R-(1-\rho)$ at MB-level among the proposed two characteristic vectors Q1 and Q2, The ρ-domain-based R-Q model with optimal slope θ_{opt} and actual rate curve.

We can see that the ρ-domain-based R-Q model relies on a strong assumption of linearity between R and ρ within each frame. Once the linearity between R and ρ is not strong enough, the estimated R-Q model could lose its accuracy. In Fig. 1, we plot some actual rate points $R(\rho(q_i)))$, which are generated by encoding the same MB for several times. We can observe that the $R(\rho)$ curve is not always a straight line. In addition, we define an optimal slope θ_{opt}[2] and use θ_{opt} to plot the estimated curves used in Fig. 1. Fig. 1 indicates that it is insufficient to just use a single optimal slope θ_{opt} to predict the actual rate point $R(\rho(q_i))$. In other words, the assumption of linearity between R and ρ does not always hold. Therefore, the estimated $R(\rho)$ curve could lose its accuracy, even if the optimal slope θ_{opt} is employed to model the rate curve $R(\rho)$. Intuitively, one of feasible solutions to this problem is to perform the "pseudocoding" process directly in the q-domain instead of in the ρ-domain to avoid the unnecessary estimation error caused by assumption of linearity between R and ρ within each MB/frame. Therefore, we argue that θ should be a function of q_i as follows:

$$\hat{R}(q_i) = \theta(q_i) \cdot (1 - \rho(q_i)). \tag{5}$$

It also implies that ρ itself could be an important characteristic for "pseudocoding" process in the q-domain.

B. Characteristics Analysis and Characteristic-Based R-Q Model

Now we turn the attention to the "pseudocoding" process in (1). Since each ρ_j is a constant, according to the regression theory [7], the characteristic vector in (1) can be equivalently extended from $[Q_{NZ}(\rho_j)Q_Z(\rho_j)1]^T$ to $[\rho_j Q_{NZ}(\rho_j)Q_Z(\rho_j)\ 1]^T$. This implies that the actual bit rate could be represented via the three characteristics Q_C,[3] Q_{NZ}, and Q_Z.

To study the relationships of $R-Q_C$, $R-Q_{NZ}$, and $R-Q_Z$, we collect lots of rate points $R(Q_C)$, $R(Q_{NZ})$, and $R(Q_Z)$ from different video sequences and plot them in Fig. 2. We calculate the correlation coefficients of $C(R,Q_C)$, $C(R,Q_{NZ})$, and $C(R,Q_Z)$. We can see that the correlation coefficients of

[2] $\theta_{\text{opt}} = \arg\min_\theta \sum_{i=1}^{31} |R(\rho(q_i)) - \hat{R}(\rho(q_i))/R(\rho(q_i))|$

[3] Q_C means nonzeros among the quantized transform coefficients.

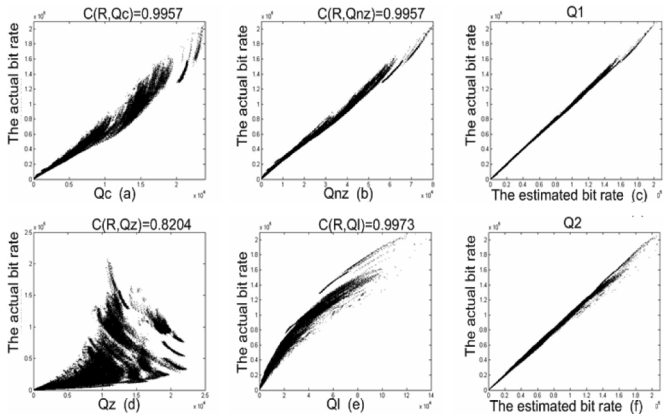

Fig. 2. (a)–(f) Plots of the relationships between the actual bit rate in the y-axis and the observed characteristics QC, QNZ, QZ,and QL, in (a), (b), (d), and (e), respectively. The relationships between the actual bit rate and the estimated bit rate using Q1 and Q2 are shown in (c) and (f), respectively.

$C(R,Q_C)$ and $C(R,Q_{NZ})$ are larger than 0.99 on average and of $C(R,Q_Z)$ is 0.82 on average. Hence, it is adequate to model the actual coding bit rate R as a linear combination of Q_C, Q_{NZ}, and Q_Z. We denote the first characteristic vector $[Q_C Q_{NZ} Q_Z 1]^T$ as Q1.

Particularly, in this work, we study another characteristic, i.e., the sum of levels of quantized nonzero coefficients, denoted as Q_L. We note that the correlation coefficient of $C(R,Q_L)$ is also more than 0.99. Hence, it could also be possible to model the actual coding bit rate R using the second characteristic vector $[Q_C Q_L Q_Z 1]^T$, named Q2. Fig. 2(c) and (f) plots the relationship between the actual bit rate and the estimated bit rate modeled by Q1 and Q2, respectively. It is clear that such multivariable modeling framework performs better than single variable modeling framework.

The main idea of this work is to perform the pseudocoding process in the q-domain. Thus, we collect extensive rate points from different frames and classify those rate points according to different q_i. Within each cluster of rate points, we study the correlation coefficient between the actual bit rate and each mentioned characteristic. They are denoted as $C_{qi}(R,Q_C)$, $C_{qi}(R,Q_{NZ})$, $C_{qi}(R,Q_L)$, and $C_{qi}(R,Q_Z)$. We can observe that correlation coefficients of $C_{qi}(R,Q_C)$, $C_{qi}(R,Q_{NZ})$, and $C_{qi}(R,Q_L)$ are very close to 1 and $C_{qi}(R,Q_Z)$ is more than 0.88. Recall that we have argued that the "pseudocoding" process should directly be performed in the q-domain instead of the ρ-domain. Therefore, we make a reasonable hypothesize—the "pseudocoding" process can be performed in the q-domain, i.e.,

$$\hat{R}(q_i) = \bar{W}_k(q_i) \cdot \bar{Q}_k(q_i)$$

with

$$\begin{aligned} \bar{Q}_{k=1}(q_i) &= [Q_C(q_i)Q_{NZ}(q_i)Q_Z(q_i)1]^T \\ \bar{Q}_{k=2}(q_i) &= [Q_C(q_i)Q_L(q_i)Q_Z(q_i)1]^T. \end{aligned} \tag{6}$$

To evaluate the modeling accuracy of different models, we transform the rate points estimated by using Q1 and Q2 from the q-domain to the ρ-domain and plot them in Fig. 1. From Fig. 1, we see that the proposed rate curves have substantial improvement compared with the ρ-domain-based R-Q model

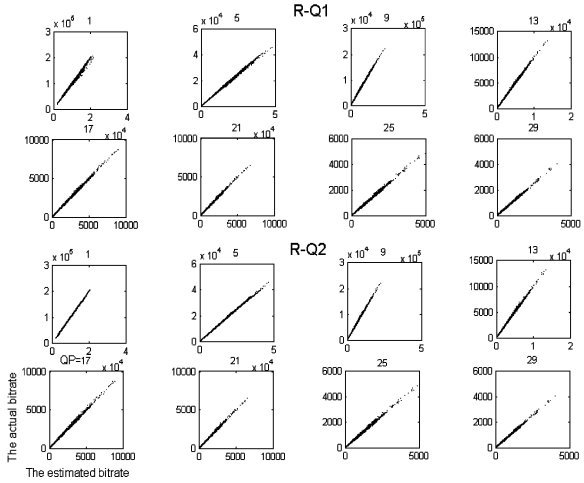

Fig. 3. Plots of the relationships between the actual bit rate for different QP and the estimated bit rate using Q1 and Q2 at frame level, respectively.

with θ_{opt}. This evidence shows that our proposed vector of characteristics can construct a better model for the actual bit rate. Further, we plot extensive pairs of $\hat{R}(q)$ and $R(q)$ for different q_i by using Q1 and Q2 in Fig. 3. Obviously, the relationships between $R(q)$ and $\hat{R}(q)$ estimated by (6) using Q1 and Q2 nearly converge to a straight line. The relative average estimation errors are about $\pm 3.3\%$. Therefore, all of these results support our argument—the "pseudocoding" process can be completely transplanted from the ρ-domain to the q-domain.[4]

III. Fast Extraction Framework

Here, we provide a fast extraction framework for computing $Q_C(q_i)$, $Q_L(q_i)$, and $Q_Z(q_i)$. In the following extraction framework, two characteristics $Q_C(q_i)$ and $Q_Z(q_i)$ are extracted in actual calculations and $Q_L(q_i)$ is obtained by a fast approximation method.

According to the definition of $Q_L(q_i)$, we only pay attention to the coefficients out of dead zone Δ_i. To speed up the computation of $Q_L(q_i)$, we introduce another temporary characteristic $Q_{\text{SANZ}}(q_i)$, which means the sum of all absolute values of nonzero transform coefficients out of dead zone Δ_i, into our extraction process. For the sake of simplicity, we assume that the two characteristics $Q_C(q_i)$ and $Q_{SANZ}(q_i)$ are already known and the processing unit in question is an $M \times M$ MB. The computation of $Q_C(q_i)$ and $Q_{\text{SANZ}}(q_i)$ will be discussed in more detail later.

Based on the definition of *Uniform Threshold Quantizer* [8], $Q_L(q_i)$ could be calculated by

$$\sum_{n=0}^{Q_C(q_i)-1} \left\lceil \left(\frac{|x_n| - \Delta_i}{q_i} \right) \right\rceil \tag{7}$$

[4]Note that without any fast lookup-table (LUT) approach, the computation of $Q_{NZ}(q_i)$ in (6) is of high computational complexity. In the following, we choose the second characteristic vector Q2 to model the source bit rate and present a fast extraction framework for it.

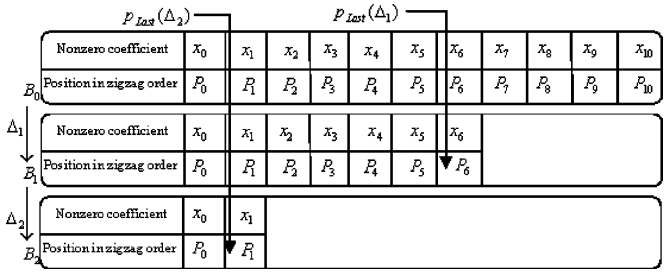

Fig. 4. 1-D array B records the status of nonzero DCT coefficients and their relative positions during applying dead zone thresholding from Δ_1 to $\Delta_{3\,1}$. $P_{Last}(q_i)$ records the position of the last nonzero coefficient after applying dead-zone thresholding Δ_i.

where x_n is a transform coefficient out of dead zone Δ_i and $Q_C(q_i)$ is the number of nonzero coefficients. Herein, we compute the item $\sum_{n=0 \text{ to } Q_C(q_i)-1} |x_n|$ in advance and denote it as $Q_{\text{SANZ}}(q_i)$. Accordingly, we can approximate (7) by following expression: (details can be found in [6] and [8]):

$$Q_L(q_i) \approx (Q_{\text{SANZ}}(q_i) - \Delta_i \cdot Q_C(q_i)) \left(\frac{1}{q_i} \right) + Q_C(q_i) \gg 1. \tag{8}$$

Now, we discuss how to compute $Q_{\text{SANZ}}(q_i)$, $Q_C(q_i)$ and $Q_Z(q_i)$. To reuse most of computed results, the proposed fast extraction process is performed recursively from q_1 to q_{31}. Herein, we involve an auxiliary status array B of an $M \times M$ MB in order to extract the three rate characteristics $Q_{\text{SANZ}}(q_i)$, $Q_C(q_i)$, and $Q_Z(q_i)$. To achieve it, after zigzag scan and DCT, we copy all absolute values of nonzero coefficients and their corresponding positions p_j into the status array B, as shown in Fig. 4. We denote the initial status of B as B_0 and its size as N. In addition, the histogram of DCT coefficients of B_0 is also built simultaneously, denoted as $D_B(x)$. The computational complexity of $D_B(x)$ roughly needs the $2N$ additive operations.

Afterwards, when we successively apply a dead-zone threshold from Δ_1 to $\Delta_{3\,1}$ to $D_B(x)$, the two rate characteristics $Q_{\text{SANZ}}(q_i)$, $Q_C(q_i)$ are generated progressively. Similarly, $Q_Z(q_i)$ is also computed recursively according to the status of B_i.

Now, we compute the initial values $Q_C(q_1)$ and $Q_{\text{SANZ}}(q_1)$ for recursive computation of $Q_C(q_i)$ and $Q_{SANZ}(q_i)$. In (9) and (11), $Q_C(q_1)$ and $Q_{\text{SANZ}}(q_1)$ represent the count of nonzero coefficients and the sum of absolute values of those coefficients, respectively. Then, the results of $Q_C(q_1)$ and $Q_{\text{SANZ}}(q_1)$ are reused to compute $Q_C(q_2)$ and $Q_{\text{SANZ}}(q_2)$, respectively. As (10) and (12) have shown, $Q_C(q_i)$ and $Q_{\text{SANZ}}(q_i)$ will be generated recursively from q_1 to q_{31}.

It is clear to see that the computational complexity of (11)[5] is about N additive operations. Obviously, the computational complexity of (9) is not more than N additive operations. Therefore, we conclude that, when we successively apply dead-zone threshold from Δ_1 to $\Delta_{3\,1}$ to $D_B(x)$, the total computational

[5]Since x and $D_B(x)$ are integer numbers, we can replace (11) fully with additive operations. For each item $D_B(x) \cdot x$, if $D_B(x) = 0$, we do nothing, otherwise $D_B(x)$ additions are performed.

complexity of the extraction of $Q_C(q_i)$ and $Q_{\mathrm{SANZ}}(q_i)$ is not more than four times of N additive operations:

$$Q_C(q_1) = \sum_{x \geq \Delta_1} D_B(x) \tag{9}$$

$$Q_C(q_i) = Q_C(q_{i-1}) - \sum_{\Delta_{i-1} \leq x \leq \Delta_i} D_B(x) \tag{10}$$

$$Q_{\mathrm{SANZ}}(q_1) = \sum_{x \geq \Delta_1} D_B(x) \cdot x \tag{11}$$

$$Q_{\mathrm{SANZ}}(q_i) = Q_{\mathrm{SANZ}}(q_{i-1}) - \sum_{\Delta_{i-1} \leq x \leq \Delta_i} D_B(x) \cdot x. \tag{12}$$

Now, we focus on the computation of $Q_Z(q_i)$. Here, we define a temporary variable $P_{\mathrm{Last}}(q_i)$ which records the position of the last nonzero coefficient of B_i. So, when we know the information of $Q_C(q_i)$ and $P_{\mathrm{Last}}(q_i)$, $Q_Z(q_i)$ can be calculated by subtracting $Q_C(q_i)$ from $P_{\mathrm{Last}}(q_i)$. The computation of $P_{\mathrm{Last}}(q_i)$ also can be recursively achieved by successively applying dead-zone threshold Δ_i to B. As Fig. 4 shows, after we apply Δ_1 to B_0, $P_{\mathrm{Last}}(q_1)$ will be obtained by moving the position of the last nonzero coefficient from P_{10} to P_6. Recursively, the results of $P_{\mathrm{Last}}(q_1)$ can be reused to compute $P_{\mathrm{Last}}(q_2)$. Clearly, when we successively apply dead-zone threshold from Δ_1 to $\Delta_{3\,1}$ to B, the total computational complexity of $P_{\mathrm{Last}}(q_i)$ is about N additive operations.

In the following, we focus on the discussion of computational complexity of existing R-Q models using QCIF format[6] with the help of a fast LUT. As analyzed above, the total computational complexity of the recursive extraction of $Q_{\mathrm{SANZ}}(q_i)$, $Q_C(q_i)$, and $Q_Z(q_i)$ is roughly $7N$ additive operations. For the construction of R-Q, it is necessary to perform (8) and (6) with Q2 for 31 times. In (6) with Q2, there are three additive operations and three multiplications. In (8), two multiplications, one shift, and two additive operations are needed. Therefore, $7N + 5 \times 31$ additions, 5×31 multiplications and 31 shifts are totally needed for the construction of the whole R-Q model.

For the ρ-domain modeling framework, we first build the one-to-one mapping table for ρ and QP. This roughly costs $4N$ additive operations for the construction the histogram of DCT coefficients and the recursive computation of one-to-one mapping between ρ and QP. To construct the rate curve $Q_{NZ}(\rho)$, at least a control point $Q_{NZ}(q)$, which is defined in (13), is needed to construct the whole characteristic curve. Here, we apply another LUT, denoted as $QNZ(|x_n|, q)$, to speed up calculation of $\log_2\lceil(|x_n| - \Delta_i) \cdot (1/q)\rceil + 2$ in (13). Therefore, the computation of $Q_{NZ}(q)$ takes about $\alpha \cdot N$, i.e., $Q_C(q)$, additive operations ($\alpha \leq 1$) as follows:

$$Q_{NZ}(q) = \sum_{n=0}^{Q_C(q)-1} \left(\log_2 \left\lceil (|x_n| - \Delta_i) \cdot \left(\frac{1}{q}\right) \right\rceil + 2 \right). \tag{13}$$

In addition, we need to consider the computation complexity of (1), (2), and (3) for the finite set of pseudo bit rates $Q_Z(\rho_j)$. In (1), two multiplications and two additive operations are required. In (2), we need to perform one multiplication. In (3), six multiplications and three additive operations are cost to compute $Q_Z(\rho_j)$. After obtaining nine pseudorate points, we use the *linear rate regulation* approach, which is shown in [5], to computing θ. Since the major computational complexity is on how to calculate the numerator of the regression expression, we use another LUT to speed up the calculation of the regression expression. Hence, there are totally 25 additive operations and 19 multiplications for the computation of the regression expression. Finally, the whole R-Q model is constructed by performing linear rate prediction for 31 times. 31×10 additive operations and 31×2 multiplications are needed.

For the variance-based R-Q model in [1], we need to perform the quadratic form to obtain the R-Q model, which is shown as follows:

$$\hat{R}(q_i) = AK\sigma^2 \cdot \left(\frac{1}{4q_i^2}\right) \tag{14}$$

where A is a constant value and K is an adaptive factor which is updated using the method in [1]. Hence, its computational complexity includes the computation of variance and (14) for 31 times. For the computation of variance, it costs 4.5 $M \times M$ additions and 1.5 $M \times M$ multiplications. In addition, to compute (14), two multiplications are needed. Therefore, there are totally $4.5M \times M + 2 \times 31$ multiplications and 1.5 $M \times M$ additive operations. Extensive experimental data shows that 1.5 $M \times M$ is three times of N, statistically. Thus, N approaches $16 \times 16 \times 1.5 \times 0.333$. Note that the above analysis of the computational complexity is in terms of an $M \times M$ MB as the processing unit. If we use a frame as processing unit, N can be easily scaled to the size of $99 \times 16 \times 16 \times 1.5 \times 0.333 = 12659.328$.

Table I summarizes the computational complexity of different R-Q models. We can see that, with the help of the fast LUT approach, two characteristic-based source models perform better than variance-based model does in term of the computation complexity. Moreover, we could speed up the computation of the ρ-domain source model by a fast table-look-up approach but the large size of the LUT is required, e.g., table for (13). On the other hand, our proposed q-domain model; 1) only requires a small size of LUT and 2) could keep low computational complexity even without the help of an LUT. Therefore, we believe our q-domain R-Q model is more suitable to the environment, where is equipped with the general purpose processing unit or the limited hardware.

IV. Experiment Results

We implement the proposed rate model on H.263+ [12] and experiment on numerous typical QCIF format videos with 300 frames. The encoding frame is fixed at 10 fps. Frame type is set to IPPP. The first I-frame is encoded with $QP = 13$.

To explore the robustness of the R-Q model, we use the predefined QP assignment similar as [2][7]to offline encode each MB and then generate the actual bit rate $R_M(q)$ and the estimated bit rate $\hat{R}_M(q)$. The ratio of the accumulation estimation errors to the total of bit rate of a frame is used to evaluate the accuracy of the MB

[6]The number of pixels in a 16×16 MB is $16 \times 16 \times 1.5$, i.e., six 8×8 blocks.

[7]The QP assignment of each MB is progressively increased by 1 from 15 to 31 and then decreased by 1 from 31 to 15 MB by MB in each encoding frame.

TABLE I
COMPARISON OF COMPUTATIONAL COMPLEXITY OF DIFFERENT MODELS

	MUL	SHIFT	ADD
Variance-Based Model [1]	3N+62	0	9N
ρ Source Model [4]	163	0	(4+α)N+100
Proposed Model	155	31	7N+155

TABLE II
COMPARISON OF COMPUTATIONAL COMPLEXITY OF DIFFERENT MODELS

E_F and E_M QP 15-31-15	Mthr_dotr		CarPhone		Highway		Foreman	
	E_F	E_M	E_F	E_M	E_F	E_M	E_F	E_M
ρ model with θ_{opt}	0.077	0.139	0.048	0.124	0.096	0.157	0.049	0.121
Multiple Logarithmic	3.481	3.896	0.583	1.036	1.771	2.192	0.279	0.660
TMN8	3.216	3.645	0.653	1.022	1.683	2.207	0.279	0.664
Q_1	0.055	0.105	0.026	0.099	0.042	0.110	0.022	0.088
Q_2	0.037	0.096	0.028	0.098	0.048	0.108	0.025	0.090

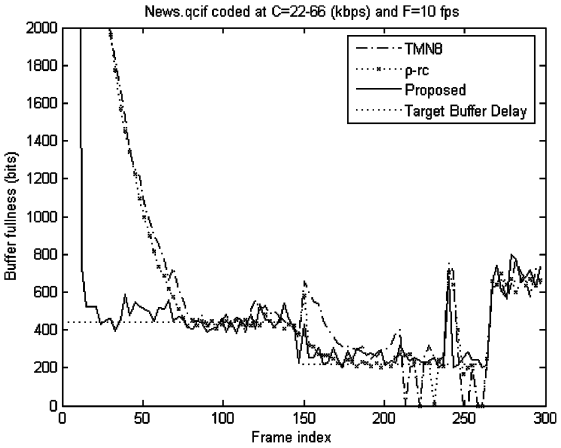

Fig. 5. Comparison of the number of bits in the encoder buffer when the proposed rate controller, TMN8rc, and ρ-rc are applied in H.263+ encoder for VBR.

level, i.e., $E_M = \sum_j |\hat{R}_{M_j}(q) - R_{M_j}(q)| / \sum_j R_{M_j}(q)$. On the other hand, for frame-level, we will adopt $E_F = |\sum_j R_{M_j}(q) - \sum_j \hat{R}_{M_j}(q)| / \sum_j R_{M_j}(q)$.

Comparisons of modeling accuracy using predefined QP assignment with the other models [1], [2], and [4] are given in Table II. For the ρ-domain based R-Q model, we use an optimal slope θ_{opt} for each MB in the simulation. We can see that the proposed rate modeling function has substantial improvement compared with the other models for both frame- and MB- levels.

In the following, we compare our previous rate controller [6], which uses the proposed rate model, with other rate controllers in more detail. The CBR cases have been presented in [6]. In this study, we will demonstrate the performance of the VBR [10]. In Fig. 5, we plot the buffer fullness for each coded frame, for which our rate controllers in [6], TMN8, and ρ-rc, are applied. One can see that, when the bandwidth fluctuates from 66 to 22 kbps, our rate controller can avoid many buffer underflows compared with TMN8rc and ρ-rc. In Table III, we list the average PSNR and the number of coded frames for different sequences with different CBR. The proposed rate controller also

TABLE III
COMPARISON OF AVERAGE PSNR WITH DIFFERENT CHANNEL RATE

Video & Frames	Channel Rate (kbps)	Obtained Rate (kbps) and Coded Frames			Average PSNR (dB)		
		Our	TMN8	ρ-rc	Our	TMN8	ρ-rc
A 300f	14.4	14.41/90	14.34/90	14.41/ 90	36.07	35.73	35.86
F 400f	23.2	23.21/115	23.21/114	23.21/113	30.24	30.08	30.16
G 869f	93.6	93.67/289	93.62/289	93.63/289	42.68	41.98	42.13
B 2100f	93.6	93.61/700	93.61/700	93.61/700	39.92	39.32	39.34
S 300f	54.4	54.04/97	54.05/97	54.05/97	36.11	35.89	35.85

A: Akiyo, F: Foreman, G: Grandma, B: Bridge-far and S: Silent.

can substantially improve up to 0.70 and 0.56 dB compared with TMN8rc and ρ-rc. Consequently, the proposed rate controller can efficiently utilize the channel bandwidth, maintain the buffer fullness, and provide better visual quality compared with TMN8rc and ρ-rc.

V. CONCLUSION

In this paper, we introduce a new characteristic vector for the "pseudocoding" process to construct the bit rate model in q-domain. Experimental results show that, compared with existing models, the proposed characteristic-based bit-rate model not only has substantial improvement in accuracy of the estimated bit rate but also is beneficial for developing an efficient rate controller.

REFERENCES

[1] J. Ribas-Corbera and S. Lei, "Rate control in DCT video coding for low-delay communication," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, no. 1, pp. 172–185, Feb. 1999.
[2] J. Wei, B. H. Soong, and Z. G. Li, "A new rate-distortion model for video transmission using multiple logarithmic functions," *IEEE Signal Process. Lett.*, vol. 11, pp. 694–697, Aug. 2004.
[3] T. Y. Kim and J. K. Kim, "An accurate rate control of MPEG video by rate-codewords modeling," in *Proc. IEEE Int. Symp. Circuits Systems (ISCAS'97)*, Hong Kong, Jun. 9–12, 1997, pp. 1261–1264.
[4] 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.
[5] 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. 1121–1235, Dec. 2001.
[6] C. Y. Chang, M. H. Chen, C. F. Chou, and D. Y. Chan, "A two-layer characteristic-based rate control framework for low delay video transmission," in *Proc. IEEE Int. Conf. Commun. (ICC'07)*, Jun. 2007, pp. 2699–2704.
[7] R. J. Freund and W. J. Wilson, *Regression Analysis: Statistical Modeling of a Response Variable*. New York: Academic, 1998.
[8] M. Ghanbari, *Video Coding: An Introduction to Standard Codecs*. London, U.K.: IEE Press, 1999.
[9] C. Y. Chang, L. Tsungnan, D. Y. Chan, and S. H. Hung, "A low complexity rate-distortion source modeling framework," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP'06)*, May 2006, pp. 929–932.
[10] Z. G. Li, C. Zhu, N. Ling, X. K. Yang, G. N. Feng, S. Wu, and F. Pan, "A unified architecture for real-time video-coding systems," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 6, pp. 472–486, Jun. 2003.
[11] Z. He and S. K. Mitra, "A linear source model and unified rate control algorithm for DCT video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, no. 11, pp. 970–982, Nov. 2002.
[12] H.263+ Codec [Online]. Available: http://www.ece.ubc.ca/spmg/h263plus/h263plus.html