

A Fine Rate Control Algorithm with Adaptive Rounding Offsets (ARO)

Qian Xu, *Member, IEEE*, Xiaoan Lu, *Member, IEEE*,

Yali Liu*, *Student Member, IEEE*, and Cristina Gomila, *Member, IEEE*

Thomson, Princeton, NJ 08540 USA

*Department of Electrical and Computer Engineering, University of California,
Davis, CA 95616 USA.

Abstract

Rate control plays an important role in regulating the bit rate to meet the bandwidth and storage requirement. Most existing video encoders regulate the bit rate by adjusting the quantization step size. We propose to incorporate a new dimension: the quantization rounding offset into rate control. In this paper, we present a rate control algorithm with adaptive rounding offsets (ARO) that jointly adjusts the quantization step size and the rounding offset for high bit rate accuracy. Different from the quantization step size that has a limited number of choices, the rounding offset is a continuously adjustable variable that allows the rate control algorithm to reach any precision in principle. Our extensive experimental results show that the proposed ARO algorithm significantly improves the rate control accuracy at almost no extra computational complexity. Compared with the ρ -domain rate control, the ARO algorithm reduces the rate control errors from about 2% to 0.5% for INTRA frames, and 5% to 1.5% for INTER frames. Our experiments also demonstrate that ARO provides with the extra benefit of smoother visual quality.

Index Terms

Video coding, rate control, H.264/AVC, quantization, rounding offset.

I. INTRODUCTION

Most video coding applications constrain the encoder to compress the input video at a given target bit rate. A rate control algorithm is often deployed to meet the rate constraints while providing high perceptual quality [1–5]. Rate-distortion models and quality metrics have been developed in order to operate the encoder at the best quality given the bit rate constraint [6, 7]. When high computational complexity is affordable, multiple encoding passes facilitate the encoder in learning the video contents and refining the encoding procedures, usually on bit allocation and quantization parameter calculation [8–10]. In general, a good rate control algorithm allocates the available bits to the video data such that the quality is optimal; and provides a mechanism to encode the video data at approximately the target number of bits. In this paper, we propose an algorithm that controls the encoder on a frame level to accurately obtain the target number of bits, assuming the bit allocation scheme is already available.

There has been a large amount of research activities on bit rate regulation. One common approach is to develop a rate-distortion model and then select the best quantization step size q based on the target bit rate [11–19]. In most modern video coding standards, the quantization step size q is chosen from a set of pre-defined values and its index, denoted as QP , is embedded in the bitstream. The bit rate R is calculated from a power function in [12] and a polynomial formulation in [13–15]. In [16], the authors employ the empirical entropy of quantized coefficients to model $R(QP)$. In the ρ -domain rate control algorithm [17–20], $R(\rho)$ and $QP(\rho)$ models are developed for more accurate rate control, where ρ is the percentage of zero transform coefficients. The common feature of these algorithms is that they all attempt to accurately characterize the relationship between the bit rate R and QP , assuming other quantization parameters such as the rounding offset and quantization matrix are constant. We denote such algorithms in this paper as QP -based rate control. With a limited number of QPs to choose from, existing rate control algorithms usually resort to macroblock-level (MB-level) QP adjustment to obtain high accuracy. However, the QP variation among different MBs may cause quality inconsistency under certain circumstances.

In recent video coding standards, the *rounding offset* s , together with the quantization step

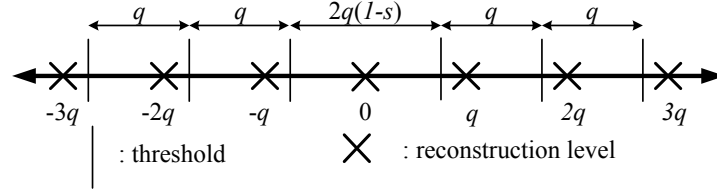


Fig. 1. An example quantizer.

size q , are used to quantize the transformed coefficient W . For example, in H.264/AVC encoding [21], W is quantized as

$$Z = \left\lfloor \frac{|W|}{q} + s \right\rfloor \cdot \text{sgn}(W), \quad (1)$$

where Z is the quantization level of W . The function $\lfloor \cdot \rfloor$ rounds a value to the nearest integer that is less than or equal to its argument, and $\text{sgn}(\cdot)$ returns the sign of the input signal. If a quantization matrix is used, W is scaled with the corresponding matrix element before quantization. The range of W where it is quantized to zero is called the *deadzone*. At the decoder, the quantization level Z is reconstructed to W' by inverse quantization:

$$W' = q \cdot Z, \quad (2)$$

where s is not involved. Therefore the rounding offset has the advantage of regulating the quantization process without the need to transmit additional parameters to the decoder. The quantization and inverse quantization processes are also illustrated in Fig. 1. Given the quantization step size q , the deadzone increases as s decreases and more transform coefficients will be quantized to zeros, resulting in a lower bit rate.

In the reference model [22] of H.264/AVC, s is by default set to $1/3$ for INTRA and $1/6$ for INTER to utilize the non-uniform probability distribution of the transform coefficients. The values are different for INTRA and INTER since the probability distributions are different between them. In [23], the author proposed a method of adaptive adjusting s to minimize the distortion, and coding efficiency improvement was reported for tests at high peak signal-to-noise ratio (PSNR).

In this paper, we present a novel fine rate control algorithm with adaptive rounding offsets, denoted as ARO, that jointly adjusts both QP and s to achieve high rate accuracy. Unlike QP

that has a limited number of choices, s is a continuously adjustable variable and it enables the rate control algorithm to reach any precision in principle. More importantly, we propose a linear rate model in which $\ln(R)$ is related to s in a linear fashion at a given QP . Our proposed algorithm adaptively adjusts s in addition to QP based on the linear $R(s|QP)$ model. Simulations with numerous video sequences show that our algorithm provides much higher bit rate accuracy than QP -based algorithms. This method can be applied to improve any QP -based rate control algorithm, and it can also be easily implemented on a MB level to achieve even higher control accuracy.

The paper is organized as follows. In Section II, we propose a joint linear source model between $\ln(R)$ and s . The novel ARO rate control algorithm that adjusts QP and s simultaneously is presented in Section III. In Section IV, we implement the ARO algorithm in a H.264 video encoder and present the experimental results. Section V concludes the paper with discussions.

II. LINEAR RELATIONSHIP BETWEEN $\ln(R)$ AND s

A. Linear Relationship

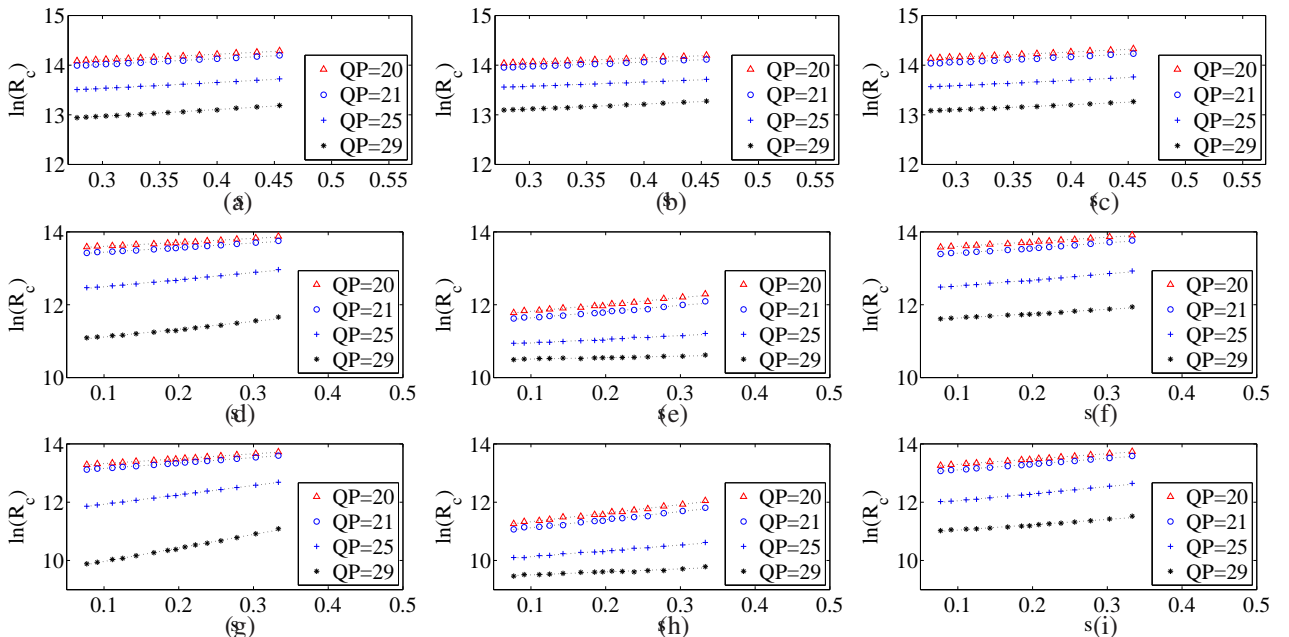


Fig. 2. $\ln(R_c)$ vs. s for (a) (d) (g) “erin”; (b) (e) (h) “man”; and (c) (f) (i) “royal4”, when encoded as an I, P, and B frames, respectively. The resolution is 1920×1080 .

We have performed extensive simulations to discover the effect of the rounding offset s on the bit rate R . Denote R as $R_c + R_h$, where R_c refers to the coefficient bits and R_h the header bits. It is observed that R_h is almost constant over different s . The plots of $\ln(R_c)$ vs. s for three different frames at different picture types (I, P, and B) are displayed in Fig. 2. Three high-definition (HD) sequences “erin brockovich” (erin), “man in restaurant” (man) and “royal wedding clip4” (royal4) are tested. The frames are INTRA or INTER encoded with $QP = \{20, 21, 25, 29\}$, $s = 0.28 \sim 0.45$ for INTRA or $s = 0.08 \sim 0.32$ for INTER. We can see from Fig. 2 that with a fixed rounding offset s , only a limited set of R_c can be obtained with discrete values of QP . However, with proper manipulation of s , any intermediate number of bits can also be achieved since s is a continuous variable. This motivates us to include the rounding offset s in the rate control algorithm to further improve the bit rate accuracy.

A closer look at Fig. 2 reveals that there is a linear relationship between $\ln(R_c)$ and s within a certain range although the sample images vary significantly from each other. More simulations with other sequences have been performed and their rate curves share the same pattern. Mathematically, the linear relationship between $\ln(R_c)$ and s can be described as:

$$\ln(R_c(QP, s)) = k_s(QP) \times (s - s_d) + \ln(R_c(QP, s_d)), \quad (3)$$

where k_s is a model parameter, s_d is the default rounding offset, and $R_c(QP, s_d)$ is the resulting bit rate (of coefficient bits) when encoding at QP and s_d .

B. The Parameter k_s

As shown in Eq. (3), k_s models how the bit rate changes with s when it differs from s_d . To study how k_s varies with QP for various video contents, we assemble in Table I the values of k_s at $QP = 21 \sim 29$ for the rate curves in Fig. 2. We observe that k_s is content specific and its value also depends on the picture type and QP . However when the range of QP is limited, which is a reasonable assumption when encoding most videos, we can approximate k_s as a constant for a given image and picture type. This is important since we can reduce Eq. (3)

TABLE I
THE VALUES OF PARAMETER k_s AT DIFFERENT QPs .

QP	I			P			B		
	erin	man	royal4	erin	man	royal4	erin	man	royal4
20	1.11	0.88	1.05	1.09	1.86	1.27	1.68	2.94	1.83
21	1.15	0.91	1.11	1.26	1.72	1.43	1.85	2.77	1.99
22	1.16	0.90	1.11	1.40	1.71	1.51	2.15	2.85	2.23
23	1.15	0.86	1.12	1.55	1.44	1.63	2.44	2.66	2.32
24	1.16	0.86	1.12	1.75	1.31	1.68	2.78	2.43	2.44
25	1.20	0.87	1.10	1.95	1.01	1.67	3.15	2.00	2.43
26	1.26	0.92	1.07	2.13	0.93	1.67	3.64	2.27	2.53
27	1.32	0.95	1.08	2.32	0.92	1.69	3.86	2.04	2.51
28	1.34	0.96	1.05	2.46	0.93	1.59	4.22	2.00	2.45
29	1.37	1.01	1.01	2.14	0.40	1.20	4.63	1.03	1.84

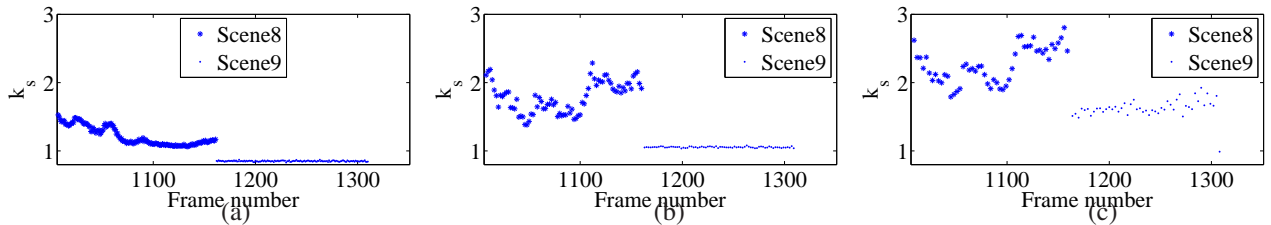


Fig. 3. The estimated k_s for frame 1006 – 1309 in “royal4” sequence for I, P, and B frames, respectively.

to

$$\ln(R_c(QP, s)) = k_s \times (s - s_d) + \ln(R_c(QP, s_d)). \quad (4)$$

To illustrate how k_s changes across different frames in a sequence, we plot in Fig. 3 the estimated k_s for frame 1006 – 1309 in “royal4”, where scene cut occurs at frame 1161. In Fig. 3(a), all frames are encoded as I frames (denoted by “III”) and k_s is plotted for each frame; In Fig. 3(b), one I frame is followed by sequential P frames in each GOP (denoted by “IPP”) and k_s is plotted for each P frame; In Fig. 3(c), one I frame is followed by multiple B and P frames in each GOP (denoted by “IBP”) and k_s is plotted for each B frame. The GOP length is 24 in all cases. For the two scenes in the figure (scene 8: frame 1006 – 1160, scene 9: frame 1161 – 1309), there is large motion at the beginning of scene 8 and the video content turns more stationary for the following frames in both scenes. As a result, k_s varies a lot at the beginning of scene 8 and becomes much smoother later. To accommodate its dynamic nature, the parameter

k_s will be estimated for each frame in the proposed ARO algorithm. Only information from previous frames of the same type will be used to estimate k_s for a new frame as k_s also varies significantly among different picture types.

With the estimated k_s based on previous frames and the QP from a QP -based rate control, the ARO algorithm needs to compute s that approaches the target (coefficient) bit rate R_c^T . Suppose the encoding bit rate is $R_c(QP, s_d)$ from a given QP and an initial s_d , the rounding offset that approaches R_c^T can be calculated as

$$s^T = \frac{1}{k_s} \ln \frac{R_c^T}{R_c(QP, s_d)} + s_d \quad (5)$$

according to Eq. (4).

III. ARO: RATE CONTROL WITH ADAPTIVE ROUNDING OFFSETS

A. QP -based Rate Control

As we have seen above, the rounding offset s , in addition to the QP , should be incorporated in rate control for better accuracy. Since we aim to show the prominent effect of the rounding offset on bit rate in this work, any regular QP -based rate control algorithm that uses a constant rounding offset s can be used to optimize QP . We need to emphasize here there is no constraint to any specific QP -based rate control scheme as our proposed scheme does not require a closed-form solution of $R_c(QP, s_d)$ as a function of QP in the source model of Eq. (3).

Among the currently available rate control schemes, we choose to use the ρ -domain rate control [19, 20] due to its superior performance. This algorithm adjusts the QP based on the linear rate model:

$$R_c = \theta(1 - \rho), \quad (6)$$

where R_c is the number of coefficient bits, ρ is the percentage of zeros of the DCT coefficients, and θ is the model parameter. To perform ρ -domain rate control in a H.264 encoder, the two-loop encoding framework in [20] is employed with the first loop (transform plus quantization) collecting global statistics to determine QP before the second loop encoding. More specifically,

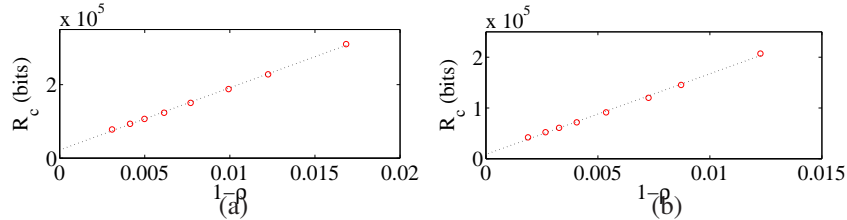


Fig. 4. The plot of R vs. $(1 - \rho)$ for “royal4” when encoded as P and B frames, respectively. Each curve is generated by encoding the same frame with $s = 0.19$ and different QP s.

after the first loop, a one-to-one mapping [20] between ρ and QP is calculated, which indicates the value of ρ at each QP . With Eq. (6) and the ρ - QP mapping table, we can identify the optimal QP from the target rate R .

B. Improve ρ -domain Rate Control

We have improved the ρ -domain rate control beyond [20] in two aspects. In [20], θ is initialized with an empirical value for the first frame and is only updated after encoding. It causes rate control inaccuracy of the first few frames, which may significantly degrade the visual quality. To adapt to different video contents and target bit rates, we enable entropy coding during the first encoding loop [20] of the first I/P/B frames after a scene change to obtain a more accurate θ . This greatly improves the rate control accuracy for these frames and the following frames as well with almost negligible increase of complexity for medium to long video segments.

In addition, for INTER frames, we refined the rate model in Eq. (6) to

$$R_c = \theta(1 - \rho) + c, \quad (7)$$

where c is a constant [24]. This is based on the observation that the fitted curve of R_c vs. $(1 - \rho)$ may not pass the origin. As an example, we plot in Fig. 4 the fitted curve of R_c vs. $(1 - \rho)$ for P/B frames in “royal4” sequence. So in our modified scheme, a small QP is used during the first encoding loop of the first P/B frames after a scene change to generate $(R_0, 1 - \rho_0)$. Then θ is calculated as $\frac{R_0}{1 - \rho_0}$ to select QP for those first frames. After the second loop encoding with the selected QP , the model parameter θ and c are both updated by considering the line between $(R_0, 1 - \rho_0)$ and $(R, 1 - \rho)$. This modification provides better rate control for INTER frames.

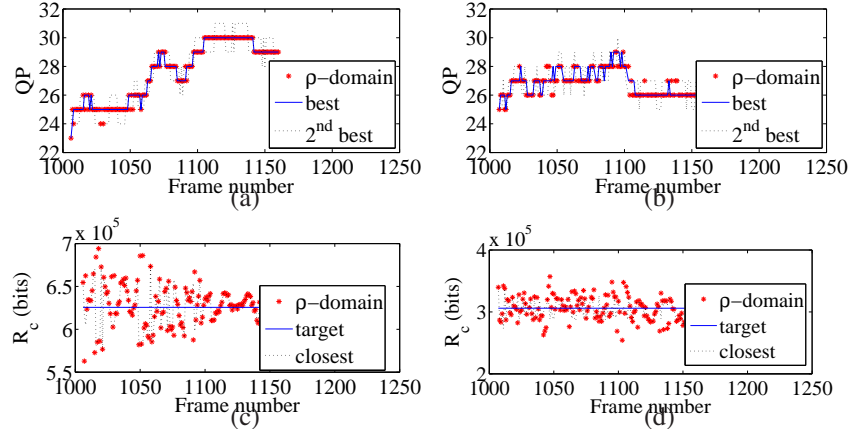


Fig. 5. QP s and bit rates for frame 1006 – 1160 in ‘royal4’ sequence with (a)(c) ‘III’ encoding and (b)(d) ‘IPP’ encoding respectively.

Note QP selection of the first P/B frames in a scene may be sub-optimal as we assume $c = 0$.

To demonstrate the advantage of the modified ρ -domain rate control algorithm, frames 1006 – 1160 of the video sequences ‘royal4’ are encoded as follows: (1) ‘III’ with an average bit rate of 15 Mbps; (2) ‘IPP’ with an average bit rate of 8 Mbps and GOP length 24. In Fig. 5(a)(b), we plot the QP s selected by the improved ρ -domain algorithm, the set of ‘best QP s’ that provide the bit rates closest to the target ones, and the set of QP s that are the second best. we observe that in most cases, the ρ -domain algorithm is capable of selecting the best candidate QP s, with the second best alternatives being selected in all other cases. Therefore, we argue ρ -domain provides optimal or near-optimal QP selection. This is also evident in Fig. 5(c)(d) where the encoding bits obtained from ρ -domain are closely around the target.

C. Proposed ARO Rate Control Algorithm

Based on Eq. (4), an intuitive solution to find the optimal combination of QP and s is to first select QP at a default s_d (with QP -based rate control), then estimate $R_c(QP, s_d)$ through actual encoding or with a rate model before calculating s^T using Eq. (5). The former approach becomes a multi-pass design [25], while in this paper, we focus on the latter single-pass approach [26] which adds little extra computation complexity to the QP -based rate control.

Specifically, when ρ -domain algorithm is used as the QP -based rate control, it selects a QP using Eq. (7) assuming a constant θ and fixed s . However, it has been observed that θ varies with

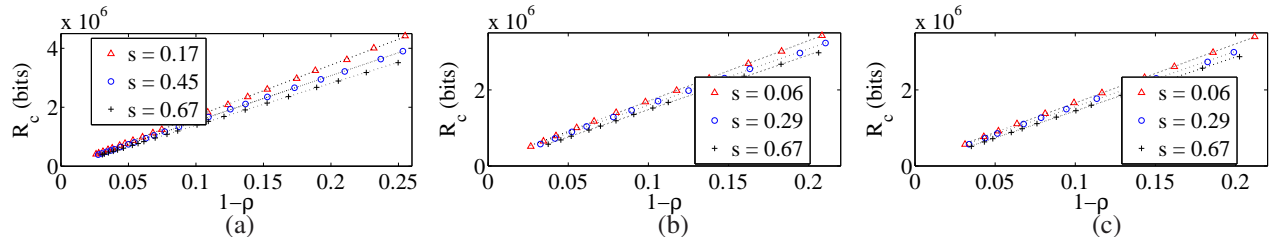


Fig. 6. The plot of R_c vs. $(1-\rho)$ at three different rounding offsets for “royal4” when encoded as (a) I, (b) P, and (c) B frames, respectively. Each curve is generated by encoding the same frame with fixed s and different QP s.

s [19], which is also shown in Fig. 6. For each picture type I, P or B, the ρ -domain rate model still holds for a fixed s , but the model parameter θ changes with different s . With our ARO algorithm where s is expected to be adaptive across frames, the assumption that θ is a constant is no longer valid. In order for the ρ -domain rate control to be functional in this context, we need to ensure it sees an encoder that operates at a fixed rounding offset s . Besides the model parameter θ in ρ -domain, we also need to estimate the parameter k_s for rounding offset update in the ARO algorithm.

We address the above mentioned issues while describing our ARO algorithm in the following subsections.

1) *Compute s_n^T* : We initialize the rounding offset s to a default value of s_d , and use it during the first loop to build the ρ - QP mapping table [20] (denoted by “ ρ -(QP, s_d) table”). Suppose QP_n is selected by the ρ -domain algorithm to encode the n^{th} frame, we need to estimate the encoding (coefficient) bit rate $R_n(QP_n, s_d)$ at QP_n and s_d before updating s^T using Eq. (5). According to the linear source model in Eq. (7), a reasonable estimate of $R_n(QP_n, s_d)$ is

$$\tilde{R}_n(QP_n, s_d) = \tilde{\theta}(QP_n, s_d) \cdot [1 - \rho(QP_n, s_d)] + c, \quad (8)$$

where $\rho(QP_n, s_d)$ can be obtained by looking up the ρ -(QP, s_d) table. In general, $\tilde{\theta}(QP_n, s_d)$ can be set as θ of the previously encoded frames of the same picture type. To achieve the target rate R_c^T , s_n^T is updated with Eq. (5) using $\tilde{R}_n(QP_n, s_d)$.

Due to the small dynamic range of linearity between $\ln(R_c)$ and s as shown in Fig. 2, after QP_n is selected by ρ -domain rate control, we may need to increase or decrease QP_n until s_n^T is within (s_L, s_U) . It is known that the modification of the rounding offset is directly associated

with the intensity of the remaining film grain in the reconstructed video [27]. Therefore the restriction of s is also desirable for smooth visual quality without introducing visible film grain strength variation.

2) *Model parameters update:* After the actual encoding with the selected QP_n and s_n^T , the obtained coefficient bit rate $R_n(QP_n, s_n^T)$ from actual encoding will be converted to $\hat{R}_n(QP_n, s_d)$ by

$$\hat{R}_n(QP_n, s_d) = [R_n(QP_n, s_n^T) - c] \cdot e^{k_s \cdot (s_d - s_n^T)} + c, \quad (9)$$

where k_s is the model parameter in Eq. (4). It is then used to update the source model parameter, which is

$$\theta(QP_n, s_d) = \frac{\hat{R}_n(QP_n, s_d) - R_0}{\rho_0 - \rho(QP_n, s_d)} \quad (10)$$

in the case of ρ -domain rate control. For INTER frames, (R_0, ρ_0) is generated as described in section III-B, and $R_0 = 0$, $\rho_0 = 1$ for INTRA frames. In this way, the QP -based rate control operates as if s is fixed.

As to the model parameter k_s , a typical value is used for the first I/P/B frames in each scene. After encoding the n^{th} frame, it is updated using the linear regression between $\left\{ \ln \frac{R_j(QP_j, s_j^T)}{\hat{R}_j(QP_j, s_d)} \right\}$ and $\{s_j^T - s_d\}$ for $j = 1, \dots, n$.

3) *Algorithm summary:* Fig. 7 depicts the block diagram of the proposed ARO algorithm, which proceeds by the following steps:

Step 1 Set $n = 1$. Initialize k_s and s_d . Encode the 1st frame using $s_1^T = s_d$ at QP_1 that is determined by the ρ -domain rate control algorithm.

Step 2 Set $n = n + 1$. Encode the n^{th} frame with the following steps:

Step 2.1 Preprocess the frame and build the ρ -(QP, s_d) table. Initialize QP_n .

Step 2.2 Compute $\hat{R}_n(QP_n, s_d)$ using Eq. (8). Compute s_n^T at QP_n based on Eq. (5). If $s_n^T > s_U$, $QP_n = QP_n - 1$; if $s_n^T < s_L$, $QP_n = QP_n + 1$. Repeat this process for a maximum of M times until $s_L \leq s_n^T \leq s_U$.

Step 2.3 Encode the n^{th} frame at QP_n and s_n^T to obtain the encoded bit rate $R_n(QP_n, s_n^T)$.

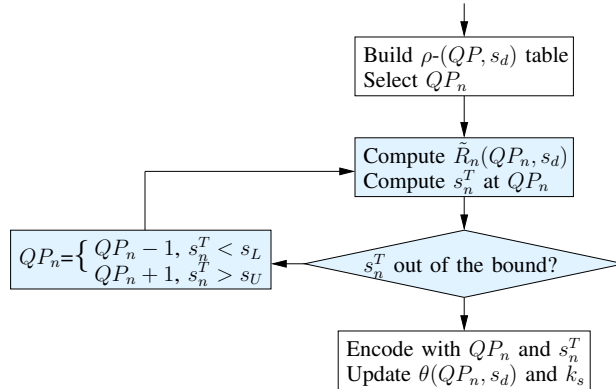


Fig. 7. The block diagram of the proposed ARO rate control algorithm. The shaded blocks are the extra computation compared to the ρ -domain rate control.

Step 2.4 Compute $\hat{R}_n(QP_n, s_d)$ using Eq. (9). Calculate $\theta(QP_n, s_d)$ with Eq. (10). Update k_s using linear regression.

Step 3 Loop step 2 until all the frames are encoded.

In the above algorithm, only step 2.2 and the computation of $\hat{R}_n(QP_n, s_d)$ and k_s in step 2.4 are the extra computation compared to the ρ -domain rate control.

IV. EXPERIMENTAL RESULTS

A. Simulation Setup

We implemented the ARO algorithm in a H.264 encoder and tested its performance with numerous video sequences. As mentioned in section III-A, we chose the ρ -domain rate control as an example of QP -based algorithms due to its high accuracy. In our ARO algorithm, the rounding offset s is restricted to be within $[0.23, 0.45]$ for INTRA frames and $[0.05, 0.32]$ for INTER frames. The maximum times of iteration in step 2.2 is set as $M = 3$. The algorithm is implemented on a frame level to guarantee consistent visual quality throughout all MBs.

Among all the sequences we have tested, only two HD and two SD sequences are presented for illustration purpose (see Table II(a)). A complete scene segment is selected for the HD sequences; For the SD sequences, two scenes are encoded independently with the scene cut occurs at frame 155. The simulation parameters are provided in Table II(b). The III, IPP, and IBP GOP structures are used to test the performance of I, P, and B frames, respectively. To

TABLE II
(A) EXAMPLE SEQUENCES (B) SIMULATION PARAMETERS.

notation	sequence	resolution	frames
erin	“erin brockovich”	1920×1080	0 – 90
royal4	“royal wedding clip4”	1920×1080	1161 – 1309
foreman	“foreman”	720 × 480	0 – 299
coastguard	“coastguard”	720 × 480	0 – 299

(a)

GOP structure	III, IPP, and IBP
GOP length	24
Target bit rate R^T (Mbps)	HD: 30 (III), 14 (IPP), 12 (IBP) SD: 8 (III), 3 (IPP)
Default k_s	1.0 (I), 1.1 (P), 2.4 (B)
Bit allocation	I:P:B = 3:1:0.5

(b)

allow fair comparisons, we fix the rate allocation of each frame in all experiments. Specifically, frames of the same type have the same target number of bits, and the bit allocation for different frame types is set as I:P:B = 3:1:0.5.

B. Performance

To measure rate control performance, we define for the m^{th} frame the relative control error as

$$\Delta_m = \frac{B_m - B_m^T}{B_m^T} \times 100\%, \quad (11)$$

where B_m and B_m^T are the actual and target number of bits of the m^{th} frame. The average control error over N frames is calculated as

$$\bar{\Delta} = \sum_{m=1}^N |\Delta_m| / N. \quad (12)$$

Denote the default rounding offset vector as $\mathbf{s}_d = (s_d^1, s_d^2)$, where s_d^1 and s_d^2 are for INTRA and INTER frames respectively. In the following, we present experimental results for the four sequences with two different settings: $\mathbf{s}_d = (1/3, 1/6)$ and $\mathbf{s}_d = (1/2, 1/4)$.

1) $\mathbf{s}_d = (1/3, 1/6)$: In Fig. 8, we compare the relative control error Δ_m of each frame between the ρ -domain rate control (with the improvements in section III-B) and our ARO algorithm for

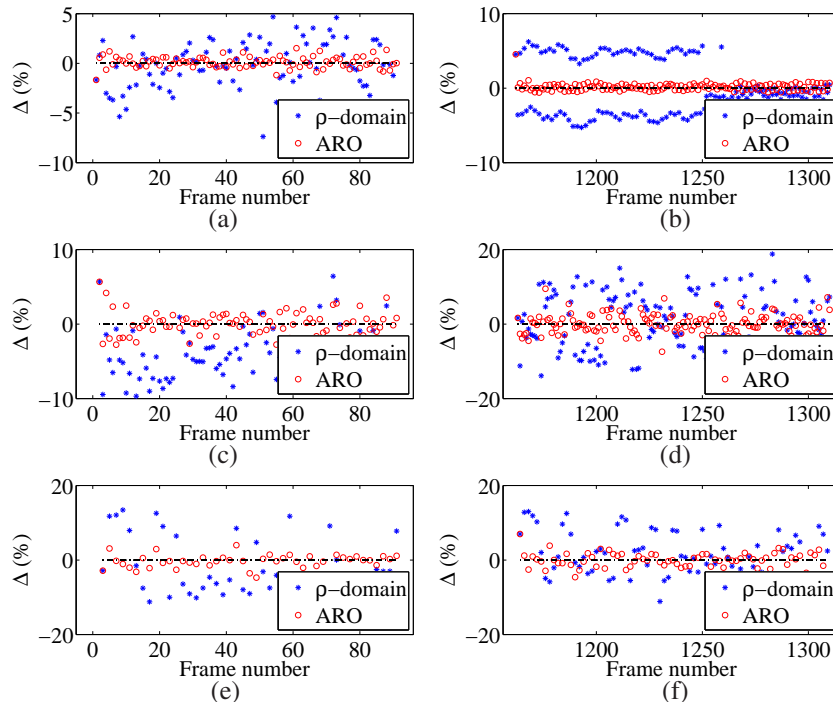


Fig. 8. Relative control errors for “erin” (left) and “royal4” (right): (a)(b) I in ‘III’; (c)(d) P in ‘IPP’; (e)(f) B in ‘IBP’. $s_d = (1/3, 1/6)$.

TABLE III
 $\bar{\Delta}$ AND PSNR COMPARISON FOR HD CONTENTS.

video	erin			royal4		
	I	P	B	I	P	B
$\bar{\Delta}$ (%)						
ρ -domain	1.97	4.78	5.51	2.98	6.50	5.15
ARO	0.44	1.32	1.50	0.39	1.96	2.32
PSNR (dB)						
ρ -domain	42.82	41.29	40.96	38.73	39.96	39.64
ARO	42.83	41.39	41.03	38.64	39.93	39.62

different frame types of the two HD sequences. The first frame of each scene segment is always encoded using the default rounding offset s_d , so its control error is the same in both algorithms. For the remaining pictures, the figures demonstrate that the rate control accuracy of ARO is consistently higher than the QP -based ρ -domain method for all tested sequences and conditions. With the QP -based method, the QP of each frame may switch up and down attempting to meet the target rate (as seen in Fig. 8(b)). The control error for each frame can be higher than 10% and the average control error $\bar{\Delta}$ around 3% for INTRA frames and 5% for INTER frames due to the limited choices of QP . By introducing the rounding offset into rate control, the control errors

for most frames are reduced to 0.5% or lower for INTRA frames and around 2% for INTER frames. As shown in Table III, the average control error $\bar{\Delta}$ with our ARO algorithm is only about 0.4% for I frames, 1.3% \sim 2.0% for P frames, and 1.5% \sim 2.3% for B frames. In other words, our ARO algorithm accomplishes 70% or higher rate control accuracy improvement over the QP -based scheme with almost no additional complexity!

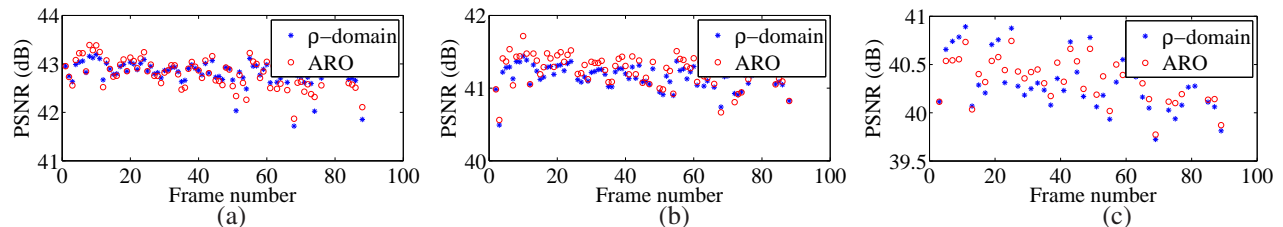


Fig. 9. PSNR for each frame of “erin” (a) I in ‘III’; (b) P in ‘IPP’; and (c) B in ‘IBP’.

Fig. 9 and Table III compare the PSNR of each frame in both schemes. We can see that the PSNRs for each frame are comparable in the two schemes and ARO has a much smoother PSNR curve in some cases. Therefore, with the bit rate closely meeting the target bit rate at each frame, ARO provides the extra benefit of smoother PSNR which can translate into more pleasant viewing experience to the human visual system.

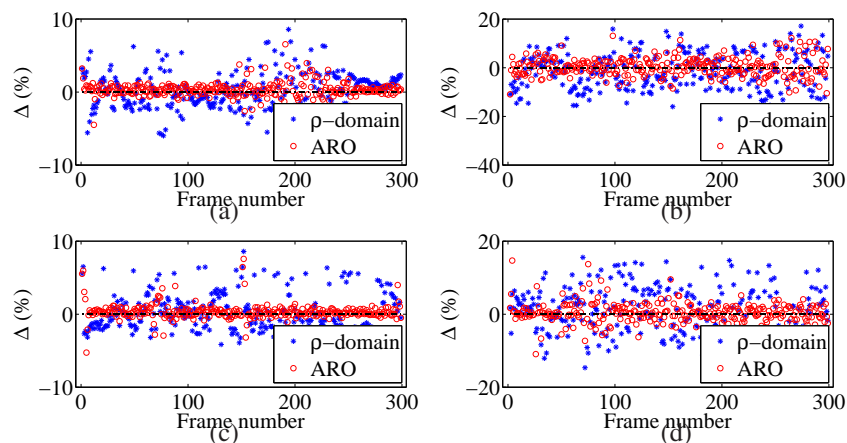


Fig. 10. Relative control error for “foreman” (left) and “coastguard” (right): (a) and (b) I in ‘III’; (c) and (d) P in ‘IPP’. $s_d = 1/3$ for INTRA frames and $s_d = 1/6$ for INTER frames.

Similar experimental results with the two SD sequences are displayed in Fig. 10. Note that the relatively large control errors for the first few frames after frame 150 are due to the starting

of a new scene segment. Compared to the QP -based method, the average control error $\bar{\Delta}$ with our ARO algorithm is 50% \sim 75% lower. The relatively large control error of ARO (but still much less than the QP -based method) at frame 60 – 80 in “coastguard” is caused by the large camera panning, which makes $\theta_n(QP_n, s_d)$ hardly predictable from previous frames. In general, ARO achieves much higher control accuracy than QP -based rate control with SD sequences.

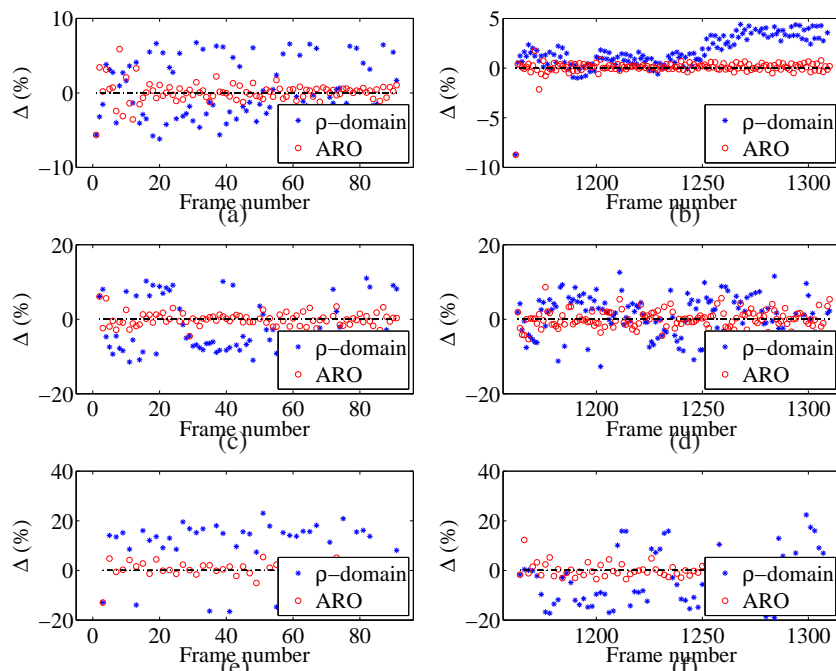


Fig. 11. Relative control error for “erin” (left) and “royal4” (right): (a)(b) I in ‘III’; (c)(d) P in ‘IPP’; (e)(f) B in ‘IBP’. $s_d = 1/2$ for INTRA frames and $s_d = 1/4$ for INTER frames.

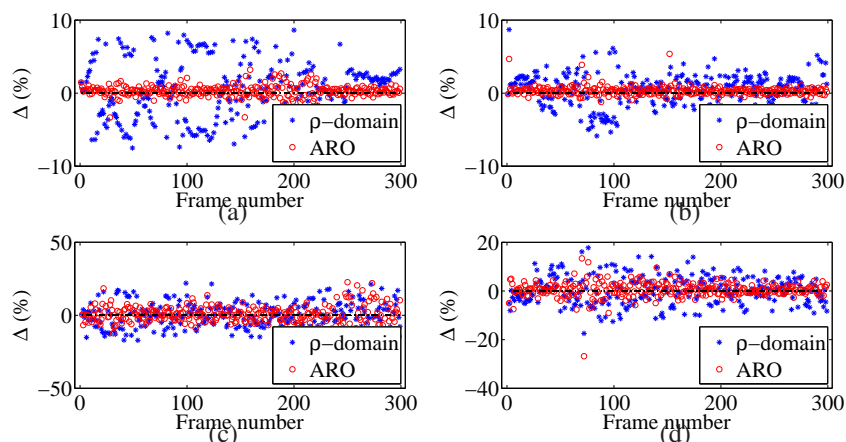


Fig. 12. Relative control error for “foreman” (left) and “coastguard” (right): (a)(b) I in ‘III’; (c)(d) P in ‘IPP’. $s_d = 1/2$ for INTRA frames and $s_d = 1/4$ for INTER frames.

2) $s_d = (1/2, 1/4)$: Experimental results with $s_d = (1/2, 1/4)$ are presented in Fig. 11 for HD sequences and Fig. 12 for SD sequences. Compared to the QP -based method, the ARO algorithm reduces the average control error by 35% to 84% as shown in the figures. Thus we conclude that our ARO scheme does not rely on any specific value of the default rounding offset s_d .

C. Parameters

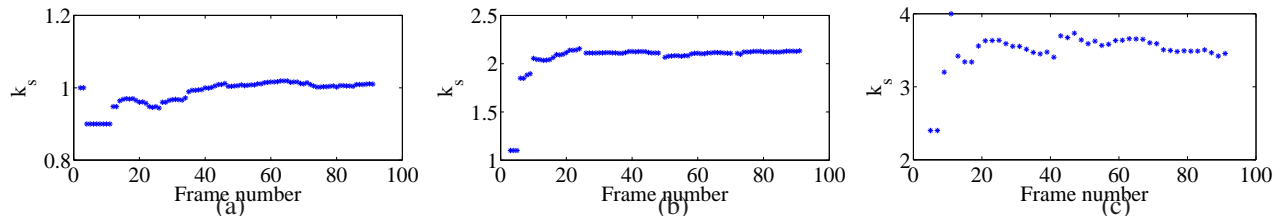


Fig. 13. k_s for each frame of “erin” (a) I in ‘III’; (b) P in ‘IPP’; and (c) B in ‘IBP’.

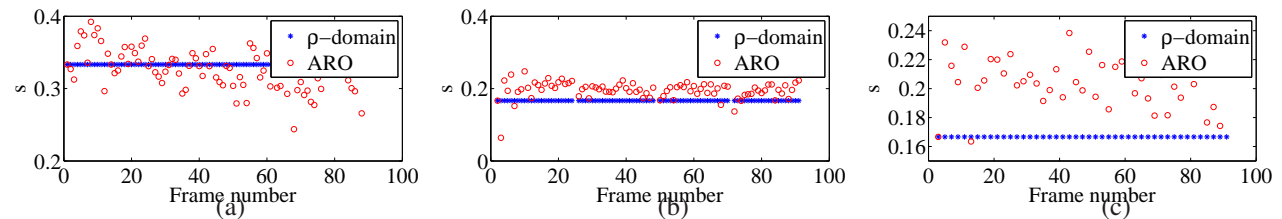


Fig. 14. s for each frame of “erin” (a) I in ‘III’; (b) P in ‘IPP’; and (c) B in ‘IBP’.

Taking the HD sequence “erin” with $s_d = (1/3, 1/6)$ as an example, the detailed model parameter k_s and rounding offset s_n^T are shown in Fig. 13 and 14. The estimated k_s starts with its default value, but soon converges after a few frames. Referring to Fig. 1, when s decreases, the deadzone and the number of coefficients that are quantized to zeros increase, generating a lower bit rate. Hence k_s can be regarded as how the coefficient distribution curve shapes around the default deadzone boundary that is defined as $\pm q(1 - s_d)$. Generally, there are more small coefficients around zeros in P and B frames and the distribution curves decay fast at the deadzone boundary, therefore the bit rates for both INTER picture types are more sensitive to the s adjustment. This explains why k_s is larger for INTER frames than INTRA ones and also varies with the contents.

As to the rounding offset plotted in Fig. 14, the QP -based approach uses the default s_d for all frames, while in our proposed scheme, the rounding offset s_n^T is adaptive to each frame according to Eq. (5) within the range where the linear relationship of Eq. (4) is valid.

D. Analysis

From the above experiments, we observe that ARO performs better for INTRA than INTER frame, and better for HD contents than for SD. In this section, we focus on understanding the natural reasons behind such behaviors.

Borrowing the concept in [28], the rate control error for a H.264 encoder using MAD prediction scheme consists of four main components, i.e.,

$$E_{total} = E_{model} + E_{QP} + E_{MAD} + E_{parameters}, \quad (13)$$

where E_{model} is caused by the model itself, E_{QP} is due to QP rounding, E_{MAD} and $E_{parameters}$ are due to inaccurate estimation. It is remarkable that by introducing the rounding offset into rate control, ARO completely removes the error caused by QP rounding through the rounding offset adjustment. For our particular implementation where ρ -domain is used as the QP -based rate control, Eq. (13) becomes

$$E_{total} = E_{model} + E_{\rho} + E_{parameters}, \quad (14)$$

where $E_{parameters}$ includes the errors from inaccurate estimation of the header bits, the parameters θ and c . For INTRA frames, the percentage of header bits is lower and the impact of $E_{parameters}$ is therefore less significant, this explains why the lowest relative control error is observed with I frames in ‘III’. On the other hand, the header bits consume a large portion of the total available bits in INTER frames [29], and change significantly with QPs and picture motions. As we assume the number of header bits of the current picture is the same as that of the previous picture, we inevitably introduce larger estimation errors in header bits counting. In addition, we inherit the problems of ρ -domain rate control when θ of previous frames is used for the current frame which may be inaccurate during high motion. Estimation inaccuracy of both header bits

and θ has adverse effects on $E_{parameters}$, causing a greater E_{total} for INTER frames. Similarly, compared with HD contents, the header bits occupy more of the total bit budget in SD, therefore $E_{parameters}$ becomes larger and leads to greater relative control errors.

V. DISCUSSION

In this paper, we have proposed a novel ARO algorithm that jointly adjusts QP and the rounding offset s for highly accurate frame-level rate control. The experimental results demonstrate significant improvements on the rate control performance with almost no additional complexity compared to QP -based method. The ARO is based on our proposed linear source model between $\ln(R)$ and s , and it can operate on top of any QP -based rate control. In this work we choose the ρ -domain method due to its high accuracy. We have improved the ρ -domain rate control from two aspects, calculating θ for the first frame and refining the source model for INTER frames. Compared to the enhanced ρ -domain method, our ARO reduces the average rate control error from 5% to about 1.5% for HD sequences, and the percentage of reduction ranges from 35% to 84% for all test sequences and conditions.

Given the high rate control accuracy obtained, the proposed ARO algorithm still has room for additional improvements. One possible improvement for estimating the model parameter θ and k_s is to use a sliding window of previous encoded frames, with the window size being adaptive to the motion of the content. Plus, instead of assuming constant header bits, we need to provide a model for the header bits, especially for low bit rate applications. Finally, the ARO algorithm can be extended to MB-level to further improve the accuracy.

ACKNOWLEDGMENT

We would like to thank Dr. Zhihai He for his insightful discussion and valuable suggestions.

REFERENCES

- [1] A. Puri and R. Aravind, "Motion-compensated video coding with adaptive perceptual quantization," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 1, no. 4, pp. 351–361, Dec. 1991.

- [2] N. Jayant, J. Johnston, and R. Safranek, "Signal compression based on models of human perception," *Proc. IEEE*, vol. 81, no. 10, pp. 1385–1422, Oct. 1993.
- [3] W. Kim, J. Yi, and S. Kim, "A bit allocation method based on picture activity for still image coding," *IEEE Trans. Image Process.*, vol. 8, no. 7, pp. 974–977, Jul. 1999.
- [4] X. Yang, W. Lin, Z. Lu, X. Lin, S. Rahardja, E. ong, and S. Yao, "Rate control for videophone using local perceptual cues," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 4, pp. 496–507, Apr. 2005.
- [5] I. Ahmad and J. Luo, "On using game theory for perceptually tuned rate control algorithm for video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, no. 2, pp. 202–208, Feb. 2006.
- [6] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
- [7] S. Winkler, *Digital Video Quality: Vision Models and Metrics*. Wiley, 2005.
- [8] P. Westerink, R. Rajagopalan, and C. Gonzales, "Two-pass MPEG-2 variable-bit-rate encoding," *IBM J. Res. Develop.*, vol. 43, no. 4, pp. 471–488, Jul. 1999.
- [9] Y. Yu, J. Zhou, Y. Wang, and C. Chen, "A Novel two-pass VBR coding algorithm for fixed-size storage application," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, no. 3, pp. 345–356, Jul. 2001.
- [10] K. Wang and J. Woods, "MPEG motion picture coding with long-term constraint on distortion variation," in *Proc. SPIE Image and Video Comm. and Process.*, vol. 5685, 2005, pp. 284–296.
- [11] MPEG-2 Test Model 5, "ISO/IEC JTC1/SC29 WG11/93-400," Apr. 1993.
- [12] W. Ding and B. Liu, "Rate control of MPEG video coding and recording by rate-quantization modeling," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 6, pp. 12–20, 1996.
- [13] T. Chiang and Y.-Q. Zhang, "A new rate control scheme using quadratic rate distortion model," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, pp. 246–250, 1997.
- [14] Z. Li, F. Pan, K. Pang, G. Lim, X. Lin, and S. Rahardja, "Adaptive basic unit layer rate

- control for JVT,” Mar. 2003, JVT-G012, 7th meeting, Pattaya II, Thailand.
- [15] P. Yin and J. Boyce, “A new rate control scheme for H.264 video coding,” in *Proc. IEEE Int. Conf. Image Process.*, Oct. 2004, pp. 449–452.
- [16] J. Ribas-Corbera and S. Lei, “Rate control in DCT video coding for low-delay communications,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, pp. 172–185, Feb. 1999.
- [17] Z. He, Y. Kim, and S. Mitra, “Low-delay rate control for DCT video coding via ρ -domain source modeling,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, pp. 928–940, Aug. 2001.
- [18] Z. He and S. Mitra, “A unified rate-distortion analysis framework for transform coding,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, pp. 1221–1236, Dec. 2001.
- [19] Z. He and S. K. Mitra, “A linear source model and a unified rate control algorithm for DCT video coding,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, pp. 970–982, Nov. 2002.
- [20] Z. He and T. Chen, “Linear rate control for JVT video coding,” in *Int. Conf. Inf. Technol.: Research and Education*, 2003, pp. 65–68.
- [21] ITU-T Recommendation H.264, “Advanced video coding for generic audiovisual services,” Mar. 2005.
- [22] JM Reference Software 14.0, May 2008.
- [23] G. Sullivan, “Adaptive quantization encoding technique using an equal expected-value rule,” Jan. 2005, JVT-N011, 14th meeting, Hong Kong, China.
- [24] I. Shin, Y. Lee, and H. Park, “Rate control using linear rate- ρ model for H.264,” *Signal Process.: Image Communication*, vol. 19, pp. 341–352, Apr. 2004.
- [25] Y. Liu, X. Lu, C. Gomila, and Q. Xu, “A novel fine rate control algorithm with adaptive rounding offset,” in *Proc. IEEE Int. Symp. Circuits Syst.*, May 2008.
- [26] Q. Xu, Y. Liu, X. Lu, and C. Gomila, “A new source model and accurate rate control algorithm with QP and rounding offset adaptation,” in *Proc. IEEE Int. Conf. Image Process.*, Oct. 2008.
- [27] T. Wedi and S. Wittmann, “Quantization with an adaptive dead zone size for H.264/AVC

- FRExt,” Mar. 2004, JVT-K026, 11th Meeting, Munich, DE.
- [28] J. Dong and N. Ling, “A model parameter and MAD prediction scheme for H.264 MB layer rate control,” in *Proc. IEEE Int. Symp. Circuits Syst.*, May 2008, pp. 628–631.
- [29] D.-K. Kwon, M.-Y. Shen, and C.-C. J. Kuo, “Rate control for H.264 video with enhanced rate and distortion models,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 5, pp. 517–529, May 2007.