International Journal of Wavelets, Multiresolution and Information Processing
Vol. 7, No. 5 (2009) 1–19


World Scientific
www.worldscientific.com

# AN EFFICIENT RATE CONTROL ALGORITHM FOR A WAVELET VIDEO CODEC

M. TUN*, K. K. LOO† and J. COSMAS‡

*School of Engineering and Design*
*Brunel University, Uxbridge*
*UB8 3PH, UK*
**Myo.Tun@brunel.ac.uk*
*†Jonathan.Loo@brunel.ac.uk*
*‡John.Cosmas@brunel.ac.uk*

Received 15 May 2008
Revised 26 August 2008

Rate control plays an essential role in video coding and transmission to provide the best video quality at the receiver's end given the constraint of certain network conditions. In this paper, a rate control algorithm using the Quality Factor ($QF$) optimization method is proposed for the wavelet-based video codec and implemented on an open source Dirac video encoder. A mathematical model which we call Rate-QF ($R-QF$) model is derived to generate the optimum $QF$ for the current coding frame according to the target bitrate. The proposed algorithm is a complete one pass process and does not require complex mathematical calculation. The process of calculating the $QF$ is quite simple and further calculation is not required for each coded frame. The experimental results show that the proposed algorithm can control the bitrate precisely (within 1% of target bitrate in average). Moreover, the variation of bitrate over each Group of Pictures (GOPs) is lower than that of H.264. This is an advantage in preventing the buffer overflow and underflow for real-time multimedia data streaming.

*Keywords*: Rate control; wavelet; Dirac.

AMS Subject Classification: 22E46, 53C35, 57S20

## 1. Introduction

In real-time visual communication, an efficient rate control algorithm at the encoder is important to assure the successful transmission of the coded video data. Essentially, the rate control part of the encoder tries to regulate the varying bitrate characteristics of the coded bitstream in order to produce high quality decoded frame at the receiver's end for a given target bitrate so that the compressed bitstream can be delivered through the available channel bandwidth without causing buffer overflow and underflow. In other words, without rate control, any video encoder would be practically hard to use for real-time end-to-end video communication.

Nowadays, rate control has become one of the important research topics in the field of video compression and transmission. To achieve the constant bitrate, most of the rate control algorithms dynamically adjust the encoder parameters in order to produce high quality decoded frames at a given target bitrate. In a novel rate control algorithm for H.264,[1] bit allocation is performed on both frame level and Macroblocks (MBs) level, and the Quantization Parameter ($QP$) is calculated from the allocated number of bits. The algorithm gives the bitrate much closer to the target bitrate. The joint source-channel rate control strategy[2] considers the end-to-end distortion caused by source quantization and channel error. However, the consideration in Refs. 1 and 2 is only for the base line profile of H.264 which consists of $IPPP$ coding and there is no bits allocation procedure for the $B$ frames. In mathematical model based rate control scheme for MPEG-2,[3] the model enables to predict the bits and the distortion generated from an encoded frame at a given quantization parameter and vice versa. Even though the scheme achieves 0.52–1.84 dB PSNR gain over MPEG-2 Test Model 5 (TM5), the prediction error of generated bits and the distortion are still too high.

In the Rate Distortion Optimization (RDO) based rate control algorithm,[4] a coding mode which minimizes the cost function is chosen and the corresponding $QP$ is used for actual encoding. Even though their proposed algorithm achieves a maximum gain of 0.48 dB over H.264 current rate control scheme, the algorithm requires two pass RDO process in finding the optimum $QP$, which introduces unnecessary coding delay and complexity to the encoder. Some research has considered the coding rate, $R$ and distortion, $D$ as the percentage of $\rho$ which is the percentage of zeros among the quantized transformed coefficients for low bitrate applications especially for H.263.[5–7] In some paper, derivation of Rate-Quantization model from the Rate-Distortion function based upon the distribution of source data to be quantized is considered.[8,9] It is assumed that the data to be quantized has Laplacian distribution in Ref. 8 and Generalized Gaussian Distribution in Ref. 9. However neither of these distributions is likely to occur in all types of video sequences and transformed methods. All the rate control techniques mentioned so far are mainly for the Discrete Cosine Transform (DCT) based encoder. There are also numerous research carrying out the rate control work on the Discrete Wavelet Transform (DWT) based video encoders in the literature. Among them, rate control via bit allocation for each sub-band is proposed in Ref. 10 for baseline coder. Even though bit allocation is one of the fundamental approaches in controlling bit rate, distributing the bit budget among the sub-bands can be very complex and the complexity increases with the level of wavelet transform. So, an accurate and less computationally complex rate control algorithm which works on either DCT or DWT based encoder, on any video format (QCIF to HD) and any type of GOP structure becomes necessary. The main objective of this research is to propose a simple and efficient rate control algorithm which meets all the above mentioned requirements and to be able to apply on any type of video encoder which uses RDO Motion Estimation and Quantization.

In this paper, a rate control algorithm is proposed by deriving the Rate-QF ($R - QF$) model, where the $QF$ which is inversely related to Lagrangian Multiplier of RDO process, control the quality of the encoded video sequence. In most type of video encoder, the constant quality coding is achieved by setting $QF$ to a constant value leaving the encoded bit rate to be an arbitrary. The algorithm presented in this paper exploits this idea by considering $QF$ as a varying parameter in order to achieve constant bitrate. It has the advantage of giving stable quality while delivering the desired constant bitrate. The proposed idea is implemented on the wavelet-based Dirac video encoder[11] where $QF$ is already integrated to the encoder as the user control parameter for constant quality coding.

The organization of this paper is as follows. Section 2 provides a brief introduction to Dirac video codec. Section 3 presents the detailed procedure of the proposed rate control algorithm. The results and discussions followed by conclusions are presented in Secs. 4 and 5, respectively.

## 2. DIRAC Video Codec

Dirac is an open source wavelet-based video codec aimed at resolutions from QCIF ($176 \times 144$) to HDTV ($1920 \times 1080$) progressive or interlaced, initially developed by BBC.[11] It aims to be competitive with the other state-of-the-art standard video codecs and its performance is very much better than MPEG-2 and slightly less than H.264 even in the alpha development stage. However, the performance was not the only factor driving its design. Dirac is intended to be simple, powerful and modular. The codec can support any frame dimensions and common chroma formats (luma only, 4:4:4, 4:2:2, 4:2:0) by means of frame padding. The padding ensures that the wavelet transform can be applied properly. Frame padding also allows for any size blocks to be used for motion estimation, even if they do not evenly fit into the picture dimensions.

Dirac uses hierarchical motion estimation for faster motion estimation and Overlapped Block-based Motion Compensation (OBMC) to avoid block-edge artifacts. First the motion compensated residual frames are wavelet-transformed using separable wavelet filters and divided into subbands. Then, they are quantized using RDO quantizers. Finally, the quantized data is entropy coded using an arithmetic encoder.

The Discrete Wavelet Transform (DWT) is now extremely well known and is described in numerous references. In Dirac, it plays the same role of the DCT in MPEG-2 in de-correlating data in a roughly frequency sensitive way, whilst having the advantage of preserving fine details better. The choice of wavelet filters has an impact on compression performance, filters having to have both compact impulse response in order to reduce ringing artefacts and other properties in order to represent smooth areas compactly. The filters currently used in Dirac are the Daubechies $(9, 7)$ filter set.

In Dirac, motion estimation mode decision is carried out by using RDO Motion Estimation Matrix. The metric consists of a basic block matching metric plus some constant multiplied by a measure of the local motion vector smoothness. The basic block matching metric used by Dirac is the Sum of Absolute Difference (SAD). The smoothness measure is based on the difference between the candidate motion vector and the median of the neighboring previously computed motion vectors. The total metric is a combination of these two metrics. Given a vector $V$ which maps the current frame block $P$ to a block $R = V(P)$ in the reference frame, the metric is given by:

$$\mathrm{SAD}(P, R) + \lambda \times \max(|V_x - M_x| + |V_y - M_y|, 48). \tag{2.1}$$

where $\lambda$ is called the Lagrangian multiplier. Dirac uses a parameter called $QF$ to control the quality of the encoded frames. $QF$ plays an important role since it is involved in the RDO processes of motion estimation and quantization as an indirect representation of the Lagrangian multiplier. $QF$ is inversely related to $\lambda$ and their relation is as below.

$$\lambda = \frac{(10^{(10-QF)/2.5})}{16}. \tag{2.2}$$

In RDO Quantization as well, subband quantization is carried out by picking the best quantizer which minimize the Lagrangian combination of rate ($R$) and distortion ($D$) for a given value of $\lambda$ as expressed below.

$$D(QP) + \lambda.R(QP) \tag{2.3}$$

where $QP$ is the quantization parameter. Rate is estimated via an adaptively-corrected zeroth-order entropy measure, ($\mathrm{Ent}(QP)$) of the quantized symbols resulting from applying the quantization factor, calculated as a value of bits/pixel. Distortion is measured in terms of the perceptually weighted fourth-power error, ($E(QP, 4)$) resulting from the difference between the original and the quantized coefficients. The perceptual weighting ensures the RDO quantization process to generate a larger weighting factor for the higher subband frequencies and vice versa. The fourth-power error, ($E(QP, 4)$) is given by:

$$E(QP, 4) = \left( \sum_{ij} |P_{ij} - Q_{ij}|^4 \right)^{\frac{1}{4}} 2.5 \tag{2.4}$$

and the total measure becomes,

$$\frac{E(QP, 4)^2}{w} + \lambda \times C \times \mathrm{Ent}(QP) \tag{2.5}$$

where $w$ is the perceptual weight associated with the subband and $C$ is a correction factor.

Since the $QF$ controls the quality of the encoded video sequence by involving in the RDO processes of motion estimation mode decision and quantization, the accuracy of the motion estimation can be greatly reduced especially for the lower $QF$ encoding mode, which affects the subjective quality of the decoded video. So, the value of $QF$ in Dirac should be set to at least 5 for low quality encoding, even though Dirac allows the value of $QF$ to range from 1 to 10.

Dirac defines three frame types. Intra frames ($I$ frames) are coded independently without reference to other frames in the sequence. Level 1 frames ($L_1$ frames) and Level 2 frames ($L_2$ frames) are both inter frames, which are coded with reference to other previously (and/or future) coded frames. The definition of the $L_1$ and $L_2$ frames are the same with $P$ and $B$ frames in H.264. The encoder operates with standard Group of Picture (GOP) modes whereby the number of $L_1$ frames between $I$ frames, and the separation between $L_1$ frames, can be specified depending on the application. The detail explanation of the Dirac's GOP and intra frames prediction structure can be found in Ref. 12.

## 3. Proposed Rate Control Algorithm

As mentioned in Sec. 1, the current Dirac architecture is controlling constant quality rather than bitrate by using a user defined parameter, $QF$ as quality indicator to maintain the desired quality. The proposed algorithm exploits this idea by considering $QF$ as a varying parameter in order to achieve constant bitrate. Since the $QF$ plays an important role in controlling the quality of the encoded video sequence or the number of bits generated in the encoding process of Dirac video codec, finding the optimum $QF$ for a given set of target bitrates and video test sequences could lead to an algorithm which controls the output bitrate of the encoder.

As a consequence of the random nature of the video sequences, the complexity of each frame in the sequence could be changing all the time. So, it is practically impossible to use the constant $QF$ to encode the entire video sequence to achieve the constant bitrate over a GOP because optimum $QF$ for a previous frame would be no longer optimum for the current and the following frames. However, bitrate controlling over a GOP could be possible by adaptively changing the $QF$ of each frame according to a certain type of algorithm before encoding. Based upon this idea, a relationship between the bitrate, $R$ and the $QF$, which can be used to estimate the $QF$ for a given target bitrate, is derived. This model is known as Rate-QF ($R - QF$) model.

Figure 1 shows the overall block diagram of Dirac encoder showing proposed rate control idea with the blue color. Using the generated number of bits required to encode a frame as the feedback parameter (bit rate, $R$), $R - QF$ model adaptively calculates the optimum $QF$ to encode the following frames in order to achieve the target bitrate. Given the value of $QF$, $\lambda$ is calculated using Eq. (2.2) in the next block, $\lambda$(QF). Both $\lambda$ and $\lambda_{ME}$ which is the scaled version of $\lambda$, are used in the RDO process of motion estimation and quantization as Lagrangian multiplier. In the DWT base video encoder, instead of allocating the total bit budget among the subbands like the rate control approach in Ref. 10, the propose technique calculates only the optimum $QF$ and let the existing encoding architecture to decide the optimum $QP$ according to the chosen value of $QF$, yielding lower $QP$ for higher value of

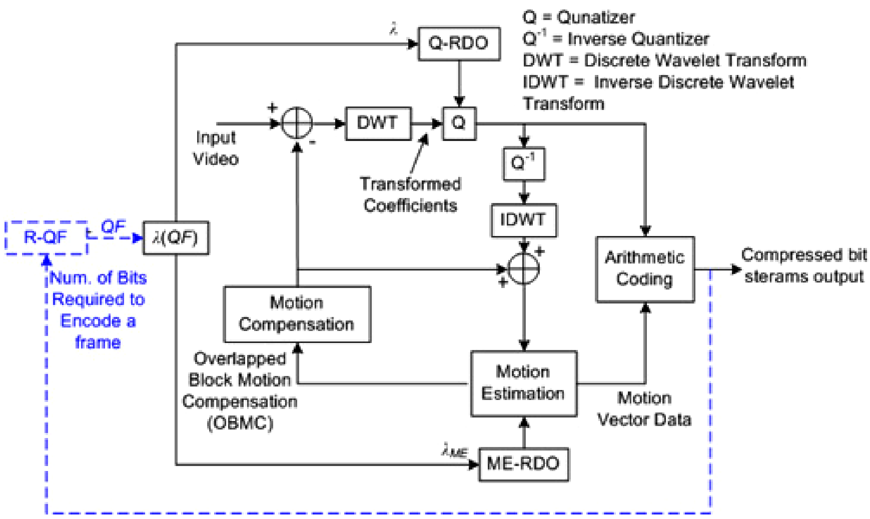


Fig. 1. The block diagram of Dirac encoder showing the proposed rate control idea using R-QF model in blue color.

$QF$ and vice versa. Again, the optimum $QP$ for each subbands is obtained by utilizing the perceptual weighting concept resulting smaller $QP$ for lower frequency subband and vice versa for a given value of $QF$. But in DCT based video encoder, calculating the optimum $QF$ (or $\lambda$) for a frame being encoded gives the optimum $QP$ for the entire frame if the RDO is enabled, reducing the computational load from calculating the $QP$ of each subband like in DWT based encoder.

The target bitrate is considered as the average bitrate over a GOP and the optimum bitrates contributed from the different types of individual frame (i.e. $I$, $L_1$ and $L_2$) in order to meet the target bitrate still need to be calculated. In order to do this, we used a modified version of test model version 5, (TM5) bit allocation procedure for MPEG-2,[13] so that calculation of bitrate contributed from different types of individual frames becomes possible by using the allocated bits to each frame type and the overall frame rate.

Finally, we employed our proposed rate control algorithm in order to achieve the bitrate close to the target bitrate for both types of frame coding available in Dirac which are $I$ frame only coding, where there is only intra frame type and normal coding which is $IL_2L_2L_1$ or $IBBP$ coding.

### 3.1. *The rate-QF ($R - QF$) model*

This section presents the derivation of the relation between rate and $QF$ in the $R - QF$ model. Since $R$ and $D$ are inversely proportional to each other as below:

$$R \propto \frac{1}{D}$$

$$R = \frac{K}{D}$$

where, $K$ is constant.

$$D = KR^{-1}.$$

The differentiation of $D$ with respect to $R$ gives the slope of the Rate vs. Distortion curve that is expressed as:

$$\frac{\partial D}{\partial R} = \lambda = -KR^{-2}$$

$$\therefore \lambda = \frac{K}{R^2} \tag{3.1}$$

where, the negative sign is neglected and $\lambda$ is the slope of the rate vs. distortion curve or the Lagrangian parameter of RDO processes for the motion estimation mode decision and optimum $QP$ selection processes in the Dirac encoder.[11] By substituting the value of $\lambda$ from Eqs. (2.2) to (3.1), we obtain:

$$\frac{10^{(10-QF)/2.5}}{16} = \frac{K}{R^2},$$

$$\frac{2}{5}(10 - QF) = \log_{10}\left(\frac{16K}{R^2}\right), \tag{3.2}$$

$$QF = 10 - \frac{5}{2}\log_{10}\left(\frac{16K}{R^2}\right).$$

The accuracy of the calculated $QF$ can be verified by the practical value captured from the encoding of the canal vertical pan street sequence[11] in CIF as shown in the rate and $QF$ relation curve in Fig. 2. It can be seen that the proposed $R-QF$ model given in Eq. (3.2) has a very accurate approximation to the practical results. The $K$ value can be calculated by substituting a set of rate and $QF$ in Eq. (3.2) from the practical data. In Fig. 2, it is calculated from the practical encoding of canal sequence with $QF = 7$ and its corresponding rate in kpbs.

### 3.2. *The bit allocation procedure*

The bit allocation that we used is the modification of TM5 from MPEG-2.[13] The complexity of each frame types is initialized as follows:

$X_I$ = Number of bits generated from the first $I$ frame coding,
$X_{L_1}$ = Number of bits generated from the first $L_1$ frame coding,
$X_{L_2}$ = Avg. num. of bit generated from the first two $L_2$ frames coding,

where $X_I$, $X_{L_1}$ and $X_{L_2}$ are the complexities of $I$, $L_1$ and $L_2$, respectively.

The number of frames in a GOP can be calculated as follows:

$$GOP_{\text{Len}} = (N_{L_1} + 1) \times L_{1\text{Sep}}$$

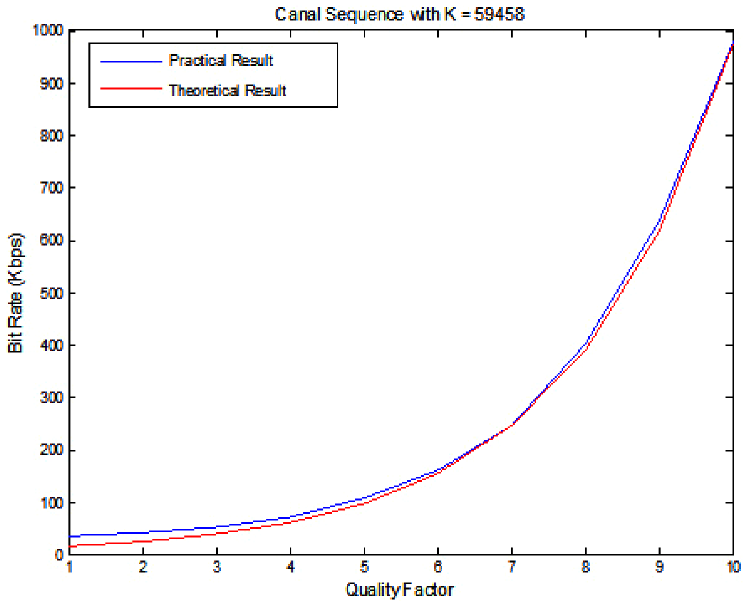


Fig. 2. The approximation of the rate and $QF$ relation with the proposed $R-QF$ model.

where $N_{L_1}$ is the number of $L_1$ frames and $L_{1\text{Sep}}$ is the $L_1$ frame separation. Furthermore, the number of bits allocated to a GOP is calculated as:

$$B = \frac{R_T \times GOP_{\text{Len}}}{\text{FrameRate}}$$

where, $R_T$ is the target bitrate in bits per second.

#### 3.2.1. *I frame bit allocation*

The number of bits allocated for the first $I$ frame can be calculated as:

$$B_I = \frac{B}{1 + \frac{N_{L_1} X_{L_1}}{X_I} + \frac{N_{L_2} X_{L_2}}{X_I}} \tag{3.3}$$

where the complexity of $I$, $X_I$ is the actual number of bits required to encode $I$ frame. The parameters are updated as follows:

(i) Total number of bits left, $B = B - B_I$,
(ii) Number of frames left, $N_I = N_I - 1 = 0, N_{L_1}, N_{L_2}$.

#### 3.2.2. $L_1$ *frame bit allocation*

The number of bits allocated for the first $L_1$ frame can be calculated as:

$$B_{L_1} = \frac{B}{N_{L_1} + \frac{N_{L_2} X_{L_2}}{X_{L_1}}} \tag{3.4}$$

where the complexity of $L_1$, $X_{L_1}$ is the actual number of bits required to encode $L_1$ frame. The parameters are updated as follows:

(i) Total number of bits left, $B = B - B_{L_1}$,
(ii) Number of frames left, $N_I = 0, N_{L_1} = N_{L_1} - 1, N_{L_2}$.

#### 3.2.3. *$L_2$ frame bit allocation*

The number of bits allocated for the first $L_2$ frame can be calculated as:

$$B_{L_2} = \frac{B}{N_{L_2} + \frac{N_{L_1} X_{L_1}}{X_{L_2}}} \tag{3.5}$$

where the complexity of $L_2$, $X_{L_2}$ is the average of the actual number of bits required to encode two $L_2$ frames. The parameters are updated as follows:

(i) Total number of bits left, $B = B - B_{L_2}$,
(ii) Number of frames left, $N_I = 0, N_{L_1}, N_{L_2} = N_{L_2} - 1$,
(iii) The number of bits allocated for second $L_2$ frame is calculated using Eq. (3.5) and again update $B$ and $N_{L_2}$.

### 3.3. *The operation of $R - QF$ model based rate control*

#### 3.3.1. *I frame only coding*

In Fig. 3, the first $I$ frame is encoded by using the initial $QF$ which is set to 7 (medium quality). $R_1$ is calculated by using the number of bits required to encode the first $I$ frame, frame rate and GOP length. The resulting bitrate associated to the first GOP which has $n$ number of $I$ frames is equal to

$$R_{\text{1st GOP}} = R_1 + R_2 + \cdots + R_n. \tag{3.6}$$

$K_1$ is calculated by substituting $QF_{\text{Initial}}$ and $R_1$ in Eq. (3.2). After that, the bit allocation process generates the optimum number of bits required to encode the first $I$ frame to achieve the target bitrates by using Eq. (3.3). The target bitrate

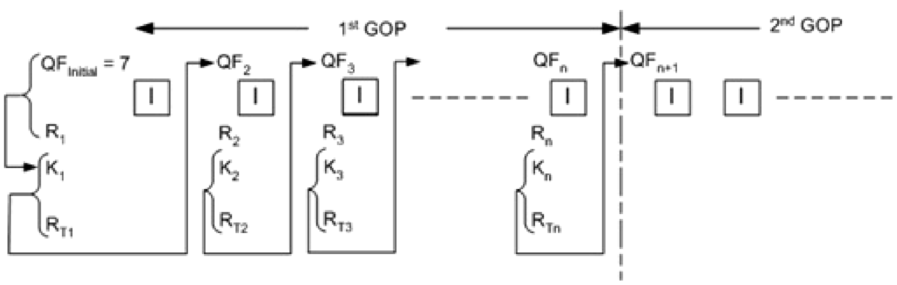


Fig. 3. Rate control procedure for $I$ frame only coding.

for a GOP which has $n$ number of $I$ frames is the combination of target bitrate of each frame and can be expressed as follows:

$$R_T = R_{T_1} + R_{T_2} + \cdots + R_{T_n}. \tag{3.7}$$

From $K_1$ and $R_{T_1}$, the optimum $QF$ for the first $I$ frame, which will be used to encode the next successive $I$ frame as $QF_2$ can be generated by using Eq. (3.2).

### 3.3.2. *Normal ($IL_2L_2L_1$ frame) coding*

In Fig. 4, the first sub-group which consist of $I$, $L_1$, $L_2$, $L_2$ frames are encoded by using the initial $QF$ which is set to 7 (medium quality). $R_1$ is calculated by using the number of bits required to encode the $L_1$ and two $L_2$ frames without including $I$, frame rate and GOP length. $K_1$ is calculated by substituting $QF_{\text{Initial}}$ and $R_1$ in Eq. (3.2). The corresponding complexities of the frames, $X_I$, $X_{L_1}$ and $X_{L_2}$ are initialized with the actual number of bits required to encode these frames but the complexity of $L_2$ frame is the average value since there are two $L_2$ frames. After that, the bit allocation process generates the optimum number of bits required to encode first sub-group frames (without including $I$) to achieve the target bitrates by using Eqs. (3.4) and (3.5), and calculates $R_{T_1}$. From $K_1$ and $R_{T_1}$, the optimum $QF$ for the first sub-group, which will be used to encode the next successive sub-group (i.e. $L_1$, $L_2$, $L_2$) as $QF_2$ can be generated by using Eq. (3.2).

The same procedure continues until the end of first GOP. The complexities of the $L_1$ and $L_2$ frames, $X_{L_1}$ and $X_{L_2}$ are updated following the encoding of each sub-group. The $QF$ of next $I$ frame, which belongs to the second GOP and is denoted as $QF_n$ in the Fig. 4, is the average of $QF$ of $I$ frame in the previous GOP and the $QF$ of previous sub-group, which is $QF_{n-1}$. The overall rate control algorithm which includes $I$ frame only and normal frame coding is illustrated in the Appendix as the flow chart.

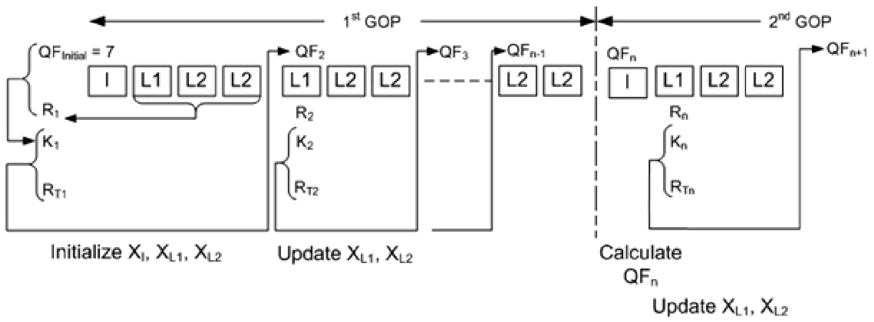


Fig. 4. Rate Control Procedure for $IL_2L_2L_1$ frame coding.

## 4. Results and Discussions

In order to evaluate the performance of the proposed algorithm, several test sequences in QCIF, CIF and HD formats were used. As for the test platform, Dirac version 0.6[11] and H.264/AVC JM11 reference software[14] have been employed. The proposed rate control algorithm is applied to both inter frame and intra frame-only coding in Dirac. The rate control in H.264 JM11[15] is used only for the verification and justification of our work since there is no previous work in Dirac as far as the rate control mechanism is concerned. Unfortunately, performance comparison with H.264 for intra frame-only coding cannot be done since current rate control algorithm of H.264 does not support this mode even though the codec supports intra frame-only coding in their high profile. The GOP length is set to 36 which means the number of $L_1$ frames is 11 and $L_1$ frame separations is 3, for both Dirac and H.264 in normal coding. The GOP length is set to 10 for $I$ frame only coding which is applicable only to Dirac. The parameters in configuration file of H.264 were carefully chosen in order to have fair comparison with Dirac for normal coding. Table 1 shows the list of configuration parameters used in H.264 encoding.

### 4.1. *PSNR performance*

#### 4.1.1. *I frame only coding*

Figure 5 shows the PSNR performance of Canal sequence in CIF format with $I$ frame only coding for the different bitrates. The proposed algorithm tries to adjust its parameters in order to get the proper $QF$ and becomes stable after encoding 30 frames or 3 GOP durations. The PSNR deep fading, which occurred in the 256 kbps target rate coding, is the result of the initial $QF$ setting which is too high for that particular target bitrate. The problem could be solved by setting the proper initial $QF$ approximated from the coding mode whether using $I$ frame only or normal, target bitrate and frame rate, instead of setting constant initial value which is set to 7 currently.

Figure 6 shows the average PSNR results for the target bitrates over the range from 256 to 2048 kbps. The average PSNR increases gradually with the target bitrate and the maximum value corresponds to 2048 kbps is 36.73 dB.

Table 1. H.264 configuration file parameters.

| Parameter description | Set value |
|---|---|
| Profile | Main |
| Frame Rate | 10(QCIF), 15(CIF), 24(HD) |
| Intra Period | 12 |
| Number of Reference | 2 |
| Inter Search Block Sizes | 16×16, 8×8, 4×4 |
| Number of $B$ frames | 2 |
| CABAC Mode | Disabled |
| Rate Control | Enabled |
| Use of FastME | 3(EPZS) |

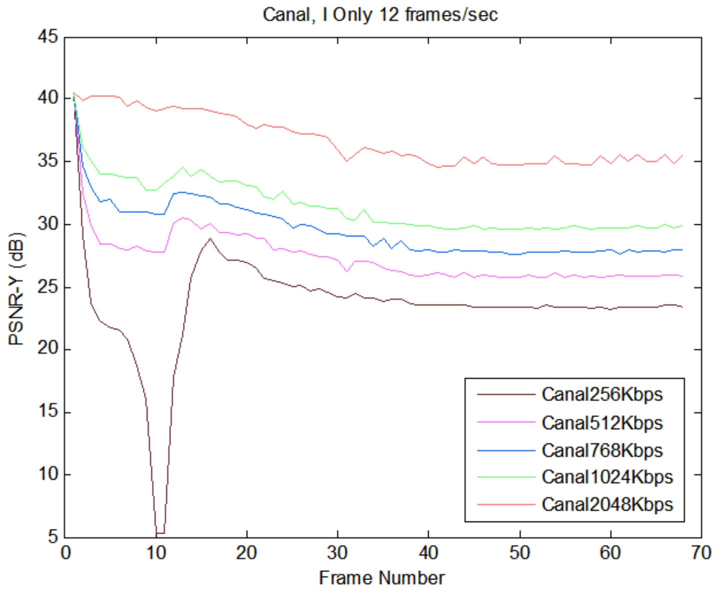


Fig. 5. PSNR performance of the Canal sequence with different bitrates for *I* frame only coding.

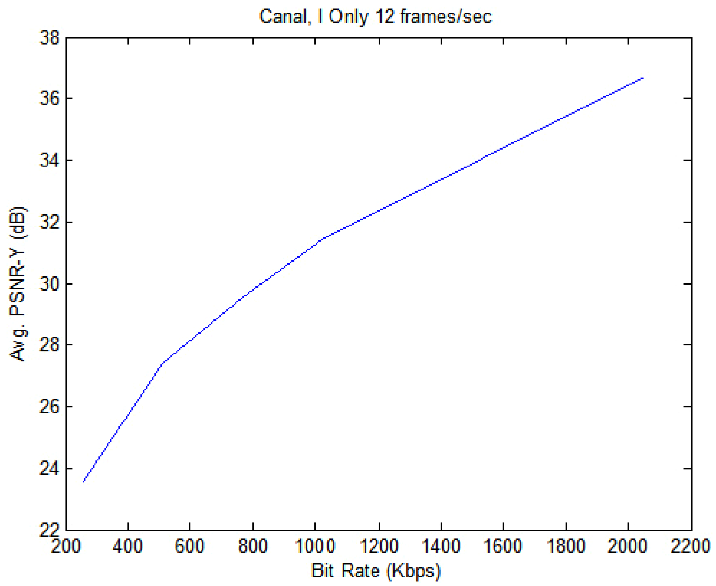


Fig. 6. Average PSNR performance of the Canal sequence with different bitrates for *I* frame only coding.

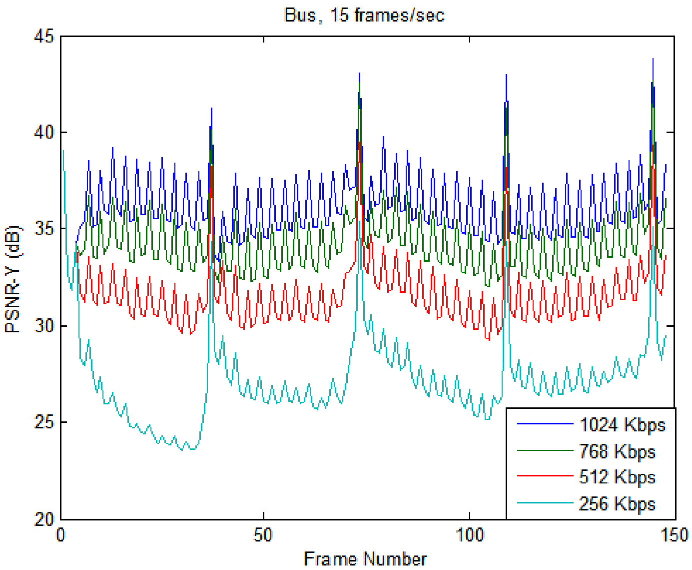


Fig. 7. PSNR performance comparison of the Bus sequence with different bitrates for normal coding.

#### 4.1.2. *Normal ($IL_2L_2L_1$ frame) coding*

Figure 7 shows the PSNR performance of the bus sequence in CIF format with normal coding for different bitrates. From the figure, we can clearly see that the stability of the algorithm is achieved after encoding the first GOP in all target bitrates.

Figure 8 shows the average PSNR results of bus sequence in CIF format with different bitrates for both Dirac and H.264 codecs. Even though Dirac has lower average PSNR performance, both codecs provide quite similar PSNR response to different bit rates. Moreover, there is no loss in terms of PSNR performance upon employing the proposed rate control algorithm. As shown in Fig. 8, the two curves of Dirac with and without using rate control give the PSNR performances which are almost the same. The average PSNR curve without rate control is generated by encoding with the constant $QF$ for the whole sequence. The value of $QF$ used here are 5, 6, 7, 8 and 9 which correspond to the rate 218.95, 333.32, 516.23, 779.88 and 1173.03 kbps, respectively.

Figure 9 shows the PSNR performance comparison of two codecs with the Bus sequence in CIF format for the target bitrate 1024 kbps. Average PSNR-Y of the Dirac is 36.19 dB which is 1.47 dB lower than that of H.264. Even though Dirac suffers 1.47 dB loss in average, the PSNR value of $I$ frame in Dirac is even higher than that of H.264 in Fig. 9. However, the PSNR difference between $I$ frames and $L_1$ frames is much larger than that of H.264 and the same problem applies to the frames between $L_1$ and $L_2$, which gives the lower average PSNR value in Dirac.

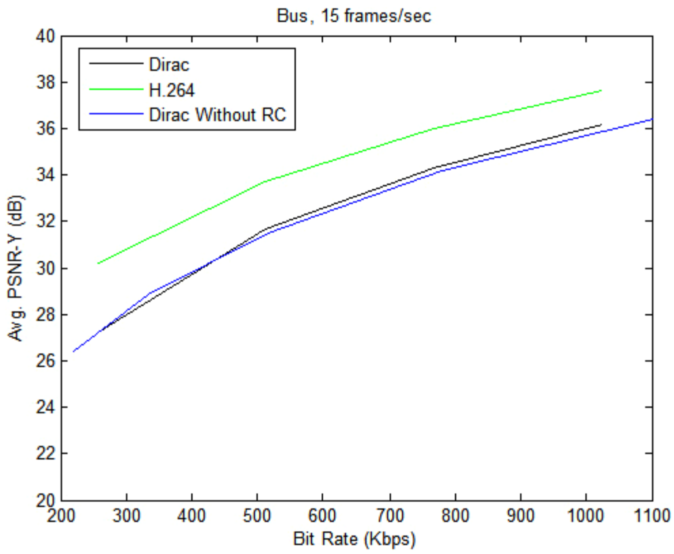


Fig. 8. Avg. PSNR performance of the Bus sequence with different bitrates for normal coding.

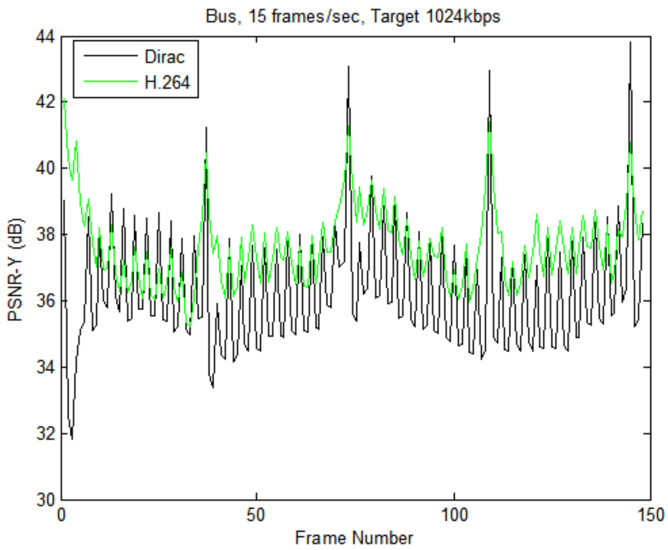


Fig. 9. PSNR performance comparison of H.264 and Dirac, the Bus sequence with target bitrates 1024 kbps.

From Table 2, we can clearly see that the maximum deviation error of the rate control technique[15] in JM11 reference software of H.264 is higher than that of the proposed technique with Dirac in all cases. Having better bitrate regulation is one of the important factors in real-time transmission since it can help to prevent buffer

Table 2. Comparison of rate control results for Dirac and H.264.

| | Avg. Rate (kbps) | | Max. Dev. Error (%) | | Avg. Dev. Error (%) | | PSNR-Y (dB) | |
|---|---|---|---|---|---|---|---|---|
| Sequence | Dirac | H.264 | Dirac | H.264 | Dirac | H.264 | Dirac | H.264 |
| QCIF, 10 Hz, 32 kbps | | | | | | | | |
| Carphone | 32.28 | 32.02 | 5.70 | 5.75 | 0.87 | 0.07 | 29.27 | 34.54 |
| Highway | 32.10 | 32.03 | 5.60 | 7.28 | 0.31 | 0.10 | 33.37 | 37.89 |
| QCIF, 10 Hz, 64 kbps | | | | | | | | |
| Carphone | 64.22 | 64.03 | 4.04 | 5.59 | 0.34 | 0.04 | 34.46 | 37.97 |
| Highway | 64.10 | 64.08 | 2.06 | 5.34 | 0.16 | 0.13 | 36.98 | 39.95 |
| QCIF, 10 Hz, 128 kbps | | | | | | | | |
| Carphone | 127.94 | 128.13 | 3.00 | 4.09 | 0.044 | 0.10 | 38.56 | 41.57 |
| Highway | 128.17 | 128.12 | 1.20 | 5.38 | 0.13 | 0.09 | 39.42 | 41.75 |
| CIF, 15 Hz, 256 kbps | | | | | | | | |
| Bus | 255.64 | 256.17 | 1.24 | 3.00 | 0.14 | 0.06 | 27.19 | 30.20 |
| Foreman | 257.66 | 257.40 | 2.56 | 6.70 | 0.65 | 0.55 | 34.83 | 36.89 |
| CIF, 15 Hz, 512 kbps | | | | | | | | |
| Bus | 511.32 | 511.95 | 1.00 | 3.32 | 0.133 | 0.01 | 31.64 | 33.73 |
| Foreman | 515.59 | 515.18 | 1.87 | 5.18 | 0.70 | 0.62 | 38.23 | 39.89 |
| CIF, 15 Hz, 1024 kbps | | | | | | | | |
| Bus | 1023.14 | 1023.90 | 1.24 | 3.76 | 0.084 | 0.01 | 36.19 | 37.66 |
| Foreman | 1029.38 | 1031.50 | 2.30 | 5.06 | 0.525 | 0.73 | 41.21 | 42.90 |
| HD720P, 24 Hz, 2 Mbps | | | | | | | | |
| Knight Shields | 1999.87 | 2002.17 | 0.79 | 4.30 | 0.0065 | 0.1086 | 35.58 | 36.25 |
| HD1080P, 24 Hz, 2 Mbps | | | | | | | | |
| Pedestrian Area | 2001.55 | 2002.2 | 1.44 | 3.62 | 0.077 | 0.11 | 35.998 | 36.5 |

overflow and underflow. Even though average deviation error from the target bitrate of the proposed technique with Dirac is higher than H.264 in most of the cases, especially for lower target bitrates, the percentage of the average deviation error is always within 1%. However, the average deviation error becomes comparable or even lower than that of H.264 when the target bitrate is higher especially at 128 kbps for QCIF format and 1024 kbps for CIF format coding. In terms of PSNR, Dirac suffers lower PSNR values for all types of components because of the higher PSNR difference between $I$, $L_1$ and $L_2$ frames. The encoder still needs to be developed further in order to achieve better or at least comparable PSNR performance with H.264. However, more importantly, Dirac was designed to maximize the subjective quality[16] instead of PSNR and so it is less likely to obtain better PSNR than H.264 in future versions of the Dirac.

### 4.2. *Deviation error from the target bitrate*

#### 4.2.1. *I frame only coding*

Figure 10 shows the percentage of deviation error from the target bitrate for Canal sequences in CIF format with the target bitrates 256 kbps. The proposed algorithm

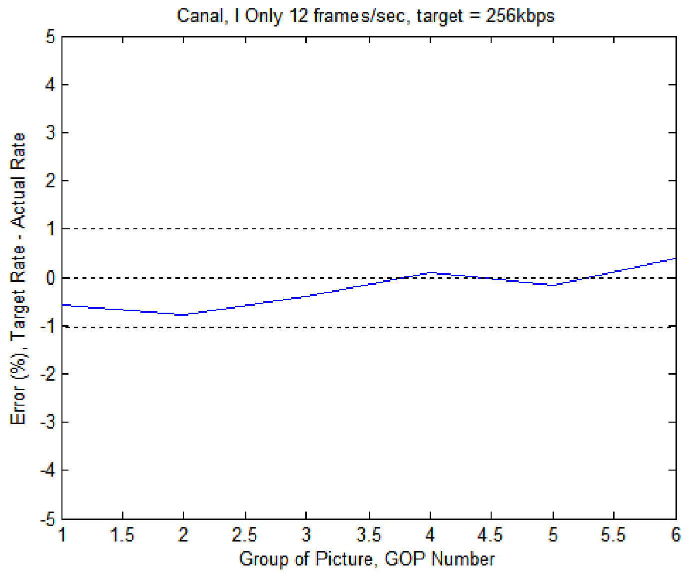


Fig. 10. Percentage of bitrate error from the target bitrate, 256 kbps, the Canal sequence in CIF format.

for $I$ frame only coding performs very well and the precision is within 1% of the target bitrates.

#### 4.2.2. *Normal ($IL_2L_2L_1$ frame) coding*

Figure 11 shows the percentage of deviation error from the target bitrate for Bus sequence in CIF format with the target bitrates 1024 kbps. The proposed algorithm performs very well and the precision is around 1% of target bitrates. Moreover, Dirac's rate control algorithm offers better bitrate regulation over each GOP than H.264. According to the Table 2, the maximum deviation error from the target bitrate of Dirac is only 1.24% which is much smaller compared with 3.76% for H.264.

## 5. Conclusion

This paper has presented the rate control algorithm which is efficient and simple to implement. Even though the algorithm is implemented and tested in Dirac, it can also be used in other types of video codec, e.g. H.264 by incorporating a parameter which controls the quality of the encoded video sequence. Experimental results have shown that the proposed algorithm can control the bitrate within 1% of the target bitrate in average and it has better bitrate regulation over each GOPs than rate control algorithm of H.264. It is an advantage which is crucial in real-time multimedia data streaming in preventing buffer overflow or underflow. Moreover, the proposed

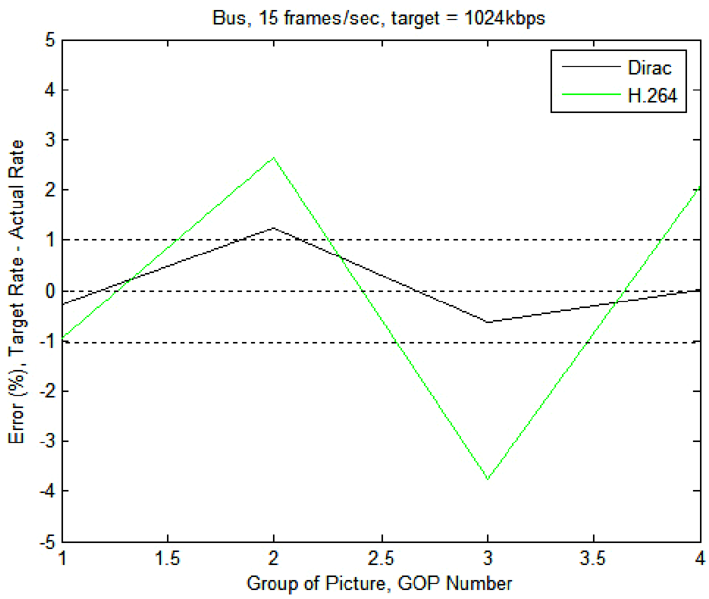


Fig. 11. Percentage of bitrate error for the target bitrate 1024 kbps, the Bus sequence.

algorithm is a complete one pass process and it is not required to iterate the calculation for finding the optimum $QF$ value. The calculation of $QF$ is based upon the simple mathematical equation and it does not even need to calculate the $QF$ for each frame in normal coding mode. The algorithm is also capable of controlling large range of bitrates from a few to several thousand kbps and so it is practically applicable for all types of video frame sizes. It is obvious that the rate control process in wavelet based video encoder becomes a lot easier with the proposed idea of $QF$ optimization since the requirement of bit allocation for each subband can be scrapped completely. In term of application, it is interesting to find that the proposed rate control algorithm can also be combined with the idea of progressive image transmission for wavelet based image coder,[17,18] in order to achieve secure and high quality video transmission through bandwidth limited wireless channel. More importantly, the proposed method can also be applied on any type of video encoders; either wavelet or non-wavelet based (e.g. H.264) as long as the encoder employs *RDO* encoding.

**Acknowledgments**

The authors gratefully acknowledge the financial support from the BBC R and D, UK and Brunel University, UK. The authors would also like to thank Dr Thomas Davies and Dr Tim Borer from the BBC R & D for their insightful and constructive comments on this work.

## Appendix. The Flow Chart of Overall Rate Control Algorithm

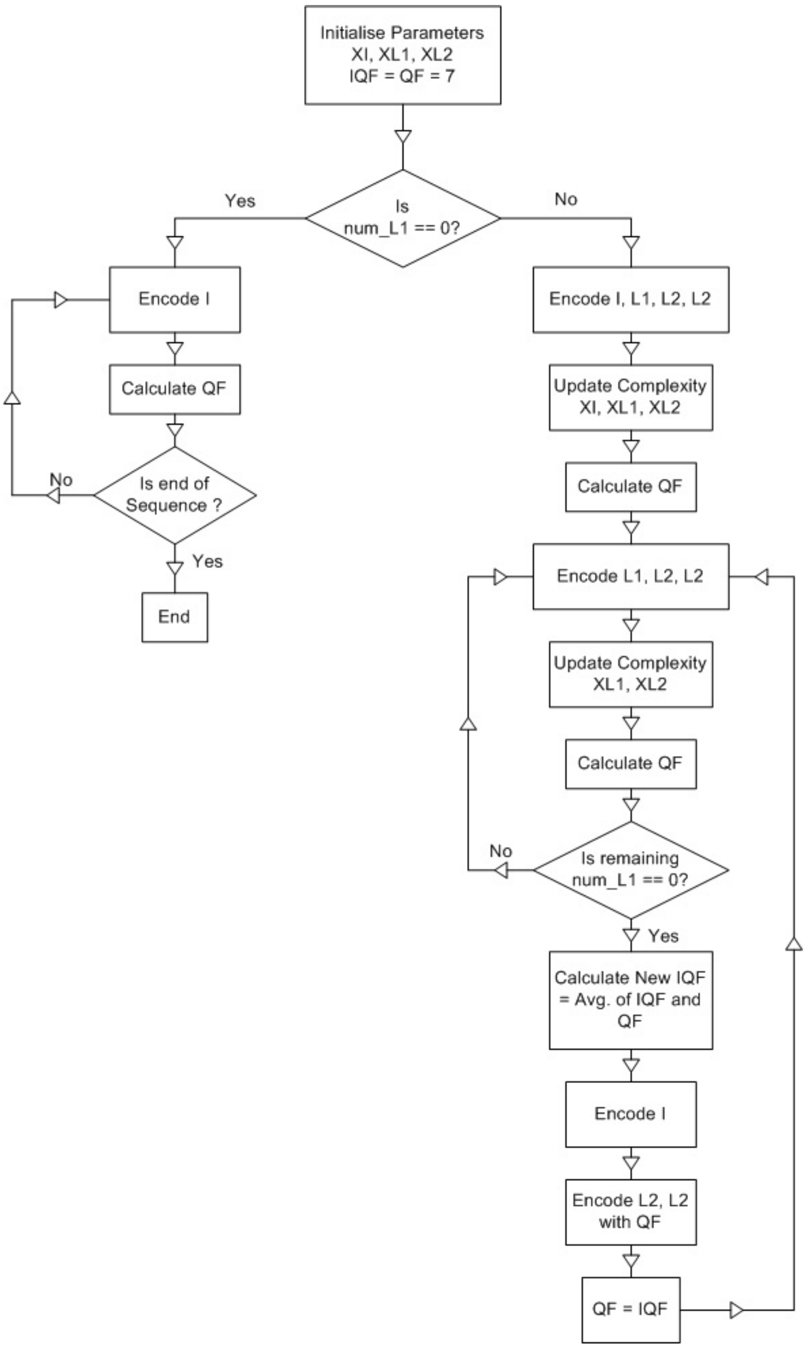


Fig. A. Complete operation of the proposed rate control algorithm.

## References


1. J. Xu and Y. He, A novel rate control for H.264, in *Proc. Int. Symposium on Circuits and Systems, ISCAS '04*, Vol. 3 (IEEE Press, 2004), pp. 809–812.
2. H. Xiong, J. Sun, S. Yu, J. Zhou and C. Chen, Rate control for real-time video network transmission on end-to-end rate-distortion and application-oriented QoS, *IEEE Trans. Broadcasting* **51**(1) (2005) 122–132.
3. S. Hong, S. Yoo, S. Lee, H. Kang and S. Hong, Rate control of MPEG video for consistent picture quality, *IEEE Trans. Broadcasting* **49**(1) (2003) 1–13.
4. S. Ma, W. Gao and Y. Lu, Rate distortion analysis for H.264/AVC video coding and its application to rate control, *IEEE Trans. Circuit Syst. Vid. Technol.* **15**(12) (2005) 1533–1544.
5. Z. He and S. K. Mitra, Optimum bit allocation and accurate rate control for video coding via $\rho$-domain source modeling, *IEEE Trans. Circuits Syst. Vid. Technol.* **12**(10) (2002) 840–849.
6. Z. He, Y. K. Kim and S. K. Mitra, $\rho$-domain source modeling and rate control for video coding and transmission, in *IEEE Int. Conf. Acoustics, Speech, and Signal Processing, 2001 (ICASSP '01)*, Vol. 3 (IEEE Press, 2001), pp. 1773–1776.
7. 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, in *IEEE Int. Conf. on Acoustics, Speech, and Signal Processing, 2001 (ICASSP '01)*, Vol. 3 (IEEE Press, 2001), pp. 1777–1780.
8. S. Kim and Y.-S. Ho, Rate control algorithm for H.264/AVC video coding standard based on rate-quantization model, in *IEEE International Conference on Multimedia and Expo, 2004. ICME '04*, Vol. 1 (IEEE Press, 2004), pp. 165–168.
9. S. Kim, S.-H. Kim and Y.-S. Ho, Adaptive model-based quantization for H.264 video rate control, in *IEEE Region 10 Conference TENCON 2004*, Vol. 1 (IEEE Press, 2004), pp. 331–334.
10. D. Lazar and A. Averbuch, Wavelet-based video coder via bit allocation, *IEEE Trans. Circuit Syst. Vid. Technol.* **11**(7) (2001) 815–832.
11. BBC Research and Development, Dirac video codec, http://dirac.sourceforge.net/.
12. M. Tun, K. K. Loo and J. Cosmas, Error-resilient performance of dirac video codec over packet-erasure channel, *IEEE Trans. Broadcasting* **53**(3) (2007) 649–659.
13. Test Model 5, http://www.mpeg.org/MPEG/MSSG/tm5.
14. H264/AVC JM11Reference Software, http://iphome.hhi.de/suehring/tml/.
15. Z. G. Li, F. Pan, K. P. Lim, G. N. Feng, X. Lin and S. Rahardaj, Adaptive basic unit layer rate control for JVT, in *Proc. of the Seventh Meeting*, Pattaya II, Thailand (7–14, March, 2003). JVT-G012.
16. T. Davies, A modified rate-distortion optimization strategy for hybrid wavelet video coding, in *IEEE International Conference on Acoustics, Speech, and Signal Processing, 2006 (ICASSP '06)*, Vol. 2 (IEEE Press, 2006), pp. 909–912.
17. C. P. Huang and C. C. Li, Secure and progressive image transmission through shadows generated by multiwavelet transform, *Int. J. Wavelets Multiresolut. Inf. Process.* **6**(6) (2008) 907–931.
18. T. H. Oh and K. G. Tan, Image tramsmission through MC-CDMA channel: An image quality evaluation, *Int. J. Wavelets Multiresolut. Inf. Process.* **6**(6) (2008) 827–850.