Proceedings of APCC2008 copyright © 2008 IEICE 08 SB 0083

Adaptive Rate Control in Frame-layer for Real-time H.264/AVC

Myoung-Jin Kim and Min-Cheol Hong, *Member, IEEE*
School of Electronic Engineering
Soongsil University
Sangdo5-dong Dongjak-gu, Seoul, 156-743, Korea
webzealer@ssu.ac.kr, mhong@ssu.ac.kr

Abstract— **Rate control has been treated as a critical issue of H.264/AVC including previous video coding standards. The purpose of this paper is to improve allocation of the number of bits without skipping the frame by accurately estimating the target bits in H.264/AVC rate control. In our scheme, we propose an enhancement method of the target frame rate based H.264/AVC bit allocation. The enhancement uses a frame complexity estimation to improve the existing mean absolute difference (MAD) complexity measurement. Bit allocation to each frame is not just computed by target frame rate but also adjusted by a combined frame complexity measure. To prevent an undesirable buffer overflow or underflow in short of channel bandwidth, the computed quantization parameter (QP) for the current frame is adjusted based on actual encoding results at that point. The objective of QP and adjustment is to produce bits as close to the target frame as possible, which is especially important for low bandwidth based real-time applications. Simulation results show that the proposed rate control scheme achieves a similar or smaller PSNR deviation, and achieves time saving of more than 99 % over JM 12.1 rate control algorithm.**

Keywords : AVC, Peak Signal-tonoise Ratio (PSNR), Quantization Parameter (QP), Rate Control

I. INTRODUCTION

H.264/AVC is the latest international video coding standard developed by Joint Video Team (JVT) of ISO Motion Picture Expert Group (MPEG) and ITU-T Video Coding Expert Group (VCEG), in order to provide an enhanced video coding standard [1–5]. This is mainly intended for video transmission in all areas where bandwidth or storage capacity is limited (e.g. video telephony or video conferencing over mobile channels and devices) by supplying an enhanced coding efficiency and an improved network adaptation [6]. Since many target applications concern video transmission over time-varying bandwidth channels, we need to control bit rate algorithms that allow modifying coding parameters according to channel's variations. Therefore, a rate control scheme is highly desirable for H.264 which must be both accurate and computationally efficient. One fundamental problem in the encoder design is the selection of quantization parameter (QP) to maximize visual quality under constraints imposed by the computational complexity and bandwidth.

H.264 encoder employs more complicated approaches in the coding procedure. One of the important approaches is the utilization of rate distortion optimization (RDO), but this imposes a big problem for rate control in H.264, which is the well-known chicken-and-egg dilemma in the RDO process [7-9]. To perform RDO, QP should be first determined by using the mean absolute difference (MAD) of the current frame and/or MB. However, in order to perform rate control, QP can only be obtained according to the coding complexity and number of target bits that are calculated by motion compensated residues after the determination of RDO mode. To resolve this dilemma, Li *et al*. [10] presented a linear model to predict MAD and adopted a fluid flow traffic model to allocate target bit rate for current frame or MB. To meet the hypothetical reference decoder (HRD) requirements, the target bits are further bounded Ref. [9]. However, to estimate the target bits for each frame, a common straightforward way is used, namely, an equal number of bits is allocated to each frame regardless of its complexity. Moreover, the linear MAD model is weak in predicting picture characteristics.

Many rate control schemes have been proposed in previous works [11-13]. However, they are difficult to be applied directly to H.264 rate control since they need the information after actually encoding the current frame to decide the appropriate QP. It does not comply with the H.264 RDO procedure. Kamaci *et al* in [14] have proposed a Cauchy-Density-Based rate model. Although this model is accurate, it needs additional computations to decide the Cauchy-based rate and distortion model parameters.

To resolve this problem, we propose an enhancement frame estimation model to select the appropriate QP for inter-frames. Using a frame complexity estimation, the performance of the MAD based complexity measurement can be improved. The proposed model utilizes complexity measurement for inter-frames, which can be obtained without pre-encoding of the target frame. Bit allocation to each frame is not only just computed by target frame rate but also adjusted by a combined frame complexity measurement. Simulation results show that our proposed method achieves better rate control for inter-coded frames without the degradation of coding performance.

This paper is organized as follows. In Section 2, we describe the development of our proposed frame estimation scheme. Section 3 demonstrates the experimental results. Finally, we present a conclusion in Section 4.

II. PROPOSED SCHEME

Similar to earlier standards, H.264/AVC exploits the spatial, temporal and statistical redundancies in the sequence. Since the level of redundancy changes from frame to frame, the number of bits generated per frame is variable. In general, rate control scheme has been treated in frame layer level and/or in the MB layer level. Frame layer rate control allocates a target number of bits to each frame. For a given frame, rate control determines a QP to achieve the frame target bits.

In this paper, we propose a new frame-layer rate control algorithm for target bit allocation. Considering the target bandwidth and the spatial complexity of the frames, we determine the optimal number of target bits for the current frame. In our scheme, the total number of basic unit is frame.

Figure 1 shows the hierarchical diagram of the proposed rate control algorithm. Specifically, the proposed rate control is consisted of four main operations:

1) Initialize parameters
2) Control GOP-level and Determination of QP
3) Estimate the target bit rate
4) Update modeling parameters

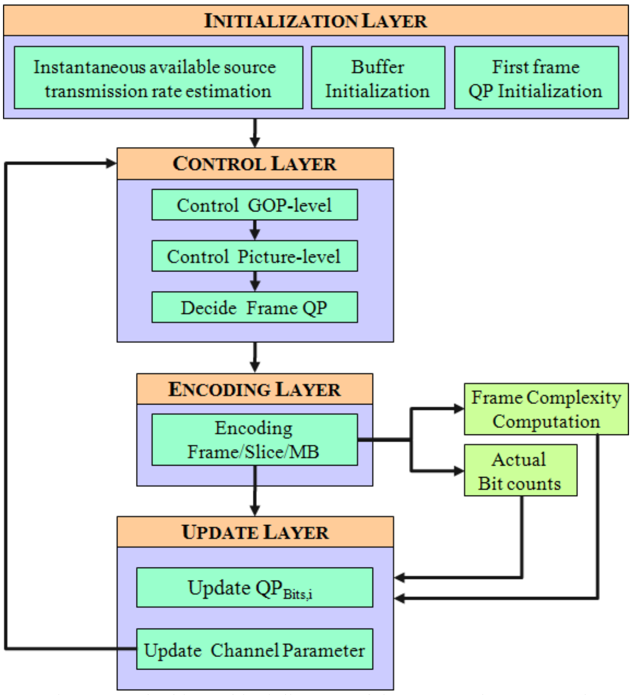

Figure 1. The hierarchical diagram of the proposed rate control algorithm.

In followings, we explain each operation in more detail.

A. First Stage

We initialize several parameters, such as frame counter, produced bits after encoding a frame, the rest of the available bits from the frame rate, and quantization parameter for first I-frame of the sequence. Using these parameters, Figure 2 shows the procedure for the choice of initial QP, that is described in JVT reference software. While I-frame is encoded, the desired bit rate is computed as

$$bpp = \frac{TargetFrameBits}{FrameRate \times (width \times height)}, \quad (1)$$

where *TargetFrameBits*, *FrameRate*, *width*, and *height* represent target bits per frame, frame rate per second, horizontal size of a frame, and vertical size of a frame.

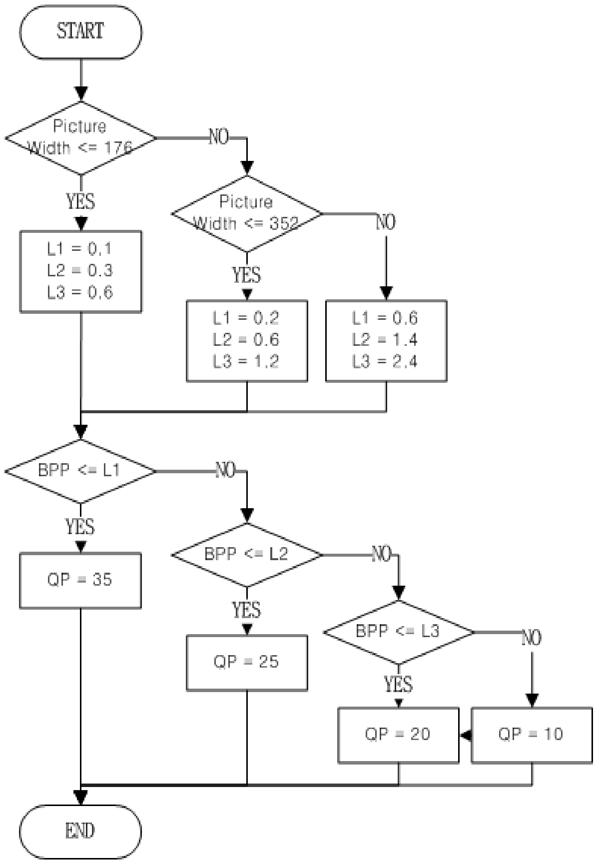

Figure 2. Initial QP algorithm

B. Second Stage

In hybrid video coding, the structure of GOP influences the whole coding efficiency. GOP is consisted of one I picture and the different picture mode such as P picture and B picture. I picture is the most important frame since it used as a reference picture for P picture.

In our scheme, we assume that the first frame in a GOP is an intra-coded I-frame, and the remaining frames are all predicted as P-type frames. We decide the frame QP by clipping the computed QP (denoted as QP_c). To maintain the smoothness of visual quality among successive frames, the computed QP is limited to change within a range. In our scheme, the QP for encoding the current frame, which is denoted as *nQP*, is decided by

$$\begin{aligned} &QP_c = choice\_QP(UBit_{QPn}), \quad (1 \le n \le 51) \\ &nQP = \min\{QP_p + \Delta QP, \max\{\mathrm{QP}_p - \Delta QP, QP_c\}\} \end{aligned} \quad (2)$$

where QP_p is the QP of previous frame, *choice_QP()* is a function to decide on QP within $UBit_{QPn}$, $UBit_{QPn}$ is the estimated bits by QP of the previous frames, that is computed by equation (5) in the final stage, where the increment or decrement of Δ QP is limited with ±2.

C. Third Stage

We estimate the target bit rate for the current frame by three steps. At first, we compute the total number of coding bits used in the previous frame. Next, complexity of the current frame is computed. Finally, the target bits for the current frame are estimated. Over the past few years, several studies have been made on rate control by frame complexity measurement, but what seems to be lacking is consideration of the encoding time of each frame for the real transmission application. It may be helpful to consider some important factors of achieving accurate frame complexity here. In order to estimate the number of bits to each frame, it is necessary to find out the average bits of five various sequences. Computed average bits of five CIF sequences (slow and smooth sequence “Container”, “News”, normal sequence “Foreman”, fast and detail sequence “Mobile”, “Stefan”) are reported in Table I.

TABLE I
AVERAGE BITS OF P-FRAMES BY QP

QP	container	foreman	mobile	news	stefan	QP Range
…						
19	36361	59772	155991	21704	140248	77201
20	28845	49106	137400	18580	124006	67587
21	21139	42086	123580	16430	111703	59867
22	20027	36069	110859	14542	99234	52376
23	16023	30299	96701	12610	87396	45247
24	12896	25432	84335	10993	75787	39617
25	10634	22152	75887	9831	68227	34368
26	8371	18285	64033	8398	57865	29263
…						

Average bits of the P-frame by experiment (Table I) have to do with the QP and the required bits, it can be derived from this table as

$$QP_{bits,n} = \alpha \times e^{(\beta \times (QP_{n-1}+1))}, \quad (1 \le n \le 51) \tag{3}$$

where QP_{bits} was just once calculated in the first stage, and updated after encoding a frame. Using the table, the parameters of equation (3) can be approximately determined. In our work, α and β are used as constant values.

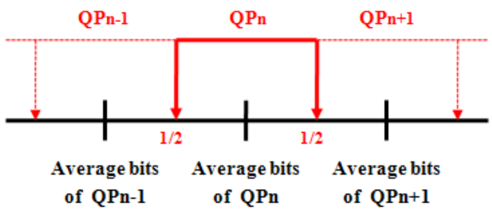

Figure 3. QP of the current frame is decided in the QR.

In this case, QP range for encoding the current frame is updated by actual bits produced from the previous frames. Since the bits as a function of quantization index take Gaussian distribution, QP range can be written as

$$QR_n = QP_{bits,n} - \frac{\left|QP_{bits,n} - QP_{bits,n+1}\right|}{2}, \quad (0 \le n < 51) \tag{4}$$

According to the QR, the number of bits in the $QP_{bits,n(0\sim51)}$ is estimated, and QP of the current frame can be computed.

D. Third Stage

We update modeling parameters for the current frame QP. These parameters are consisted of weighted combination of two values: 1) the number of bits produced from the previous frame; 2) the number of bits by scaling the average bits from the reference twenty frames. For frame-level rate control, the target bits for each frame are first determined adaptively according to the frame complexity. To estimate the current frame complexity, we use the modeling parameters above. In order to show improvement of the mismatch ratio between the target bits and actual achieved bits per frame, we define as

$$UBit_{QPn} = \lambda \times RBit_{t-1} + (1-\lambda) \times SBit \tag{5}$$

where $UBit_{QPn}$ is estimated bits by QP from the previous frames, which is updated after encoding a frame, and is reflected in the $QP_{bits,n(0\sim51)}$. $RBit_{t-1}$ is the number of bits produced from the previous frame, $SBit$ is the average bits from the previous frames with the same QP value in the reference frames, λ is weighting factor from among the previous frames because current frame is close correlation between the adjacent frames, is set to 0.67 as experimental value.

The method to allocate the number of bits of the current frame is decided by

$$TB_c = \frac{TB_n + (TB_p - Bits_p)}{nF}, (0 < nF \le FrameRate) \tag{6}$$

where TB_c is the number of estimated target bits per frame in the GOP for the current frame, TB_n is available bits in the rest frame rate, TB_p is estimated target bits of the previous frame, $Bits_p$ is produced actual bits from the previous frame, nF is the rest of the frame in the frame rate.

III. EXPERIMENTAL RESULT

The proposed rate control algorithm is tested for various video sequences. Every test sequence is encoded with only one I-frame of the first frame followed by P-frames at 30 frames/s. As a reference for comparisons, the H.264/AVC rate control algorithm was selected (as is implemented on reference software JM12.1). We employed ten test sequences of the QCIF 4:2:0 and size (176×144 pixels), a total of 300 frames were coded without skipping the frames. The H.264 encoder was configured to have one reference frame for inter motion search, (1/4)-pel motion vector resolution UVLC for symbol coding, rate-distortion optimized mode decisions, and search range of 16. More results are reported in Table II and III. These tables compare the average PSNR values and average encoding time with the proposed and the JM12.1 algorithms. For “Glasgow”, “Paris”, and “Table” sequence, we have improved the average PSNR values by up to 0.16 dB, 0.13 dB, and 0.06 dB respectively. Other sequences are similar or smaller than JM12.1. As shown in Table II and III, in encoding time, our method is more stable and very fast to the various target bit rates.

In experimental results, we achieved a similar PSNR over the most test set. Table II and Table III also show that the time saving of 99.999 % in the all sequences when compared to the JM12.1 rate control algorithm. These results are all very desirable for various target bit rates or frame rates in real time applications.

TABLE II
AVERAGE PSNR-Y OF FRAME AND ENCODING TIME OF THE 5 SEQUENCE

Sequence	Target Bit Rate	PSNR(Y)			Encoding Time(μs)		
		JM	Prop.	Gain	JM	Prop.	%
Container	32K	34.74	34.70	-0.04	905	1	0.001
	64K	37.42	37.06	-0.36	923	1	0.001
	128K	40.24	40.05	-0.19	945	1	0.001
	192K	41.99	41.88	-0.11	959	1	0.001
	256K	43.37	43.33	-0.04	966	1	0.001
	384K	45.45	45.13	-0.32	980	1	0.001
	512K	47.15	46.76	-0.39	971	1	0.001
Foreman	**32K**	**28.39**	**28.78**	**0.39**	**881**	**1**	**0.001**
	64K	32.15	32.05	-0.10	922	1	0.001
	128K	35.75	35.60	-0.15	964	1	0.001
	192K	37.83	37.34	-0.49	982	1	0.001
	256K	39.33	39.15	-0.18	995	1	0.001
	384K	41.56	41.28	-0.28	1007	1	0.001
	512K	43.22	42.88	-0.34	1014	1	0.001
Glasgow	**32K**	**24.80**	**25.71**	**0.91**	**874**	**1**	**0.001**
	64K	**27.01**	**27.70**	**0.69**	**906**	**1**	**0.001**
	128K	**29.73**	**30.15**	**0.46**	**936**	**1**	**0.001**
	192K	**31.55**	**32.15**	**0.60**	**956**	**1**	**0.001**
	256K	**33.15**	**33.32**	**0.17**	**966**	**1**	**0.001**
	384K	35.18	34.80	-0.38	981	1	0.001
	512K	36.56	35.97	-0.59	993	1	0.001
Paris	**32K**	**26.57**	**26.70**	**0.13**	**892**	**1**	**0.001**
	64K	**29.05**	**29.35**	**0.30**	**920**	**1**	**0.001**
	128K	**33.65**	**33.78**	**0.13**	**939**	**1**	**0.001**
	192K	**35.90**	**36.03**	**0.13**	**953**	**1**	**0.001**
	256K	38.12	37.99	-0.13	957	1	0.001
	384K	**41.09**	**41.24**	**0.15**	**970**	**1**	**0.001**
	512K	**43.80**	**43.90**	**0.10**	**966**	**1**	**0.001**
Table	**32K**	**29.27**	**29.87**	**0.60**	**890**	**1**	**0.001**
	64K	**33.11**	**33.12**	**0.01**	**926**	**1**	**0.001**
	128K	**36.56**	**36.75**	**0.19**	**951**	**1**	**0.001**
	192K	38.67	38.66	-0.01	966	1	0.001
	256K	40.52	40.36	-0.16	973	1	0.001
	384K	**43.11**	**43.18**	**0.07**	**977**	**1**	**0.001**
	512K	**44.88**	**44.89**	**0.01**	**978**	**1**	**0.001**

Encoding Time(μs) is only measured time unit for the rate control algorithm, especially at timer, which is the current value of the high-resolution performance counter.

TABLE III
AVERAGE PSNR-Y OF FRAME AND ENCODING TIME OF THE SEQUENCES

Sequence	PSNR(Y)			Encoding Time(μs)		
	JM	Prop.	Gain	JM	Prop.	%
Container	42.87	42.67	-0.20	959.770	1	0.001
News	43.93	43.73	-0.20	934.091	1	0.001
Foreman	38.56	38.29	-0.27	982.418	1	0.001
Mobile	29.62	29.49	-0.13	983.391	1	0.001
Glasgow	**32.52**	**32.69**	**0.16**	**957.924**	**1**	**0.001**
Akiyo	48.71	48.53	-0.18	897.822	1	0.001
Coastguard	34.70	34.56	-0.15	982.142	1	0.001
Hall Monitor	42.06	41.87	-0.19	951.651	1	0.001
Paris	**37.40**	**37.53**	**0.13**	**951.994**	**1**	**0.001**
Table	**39.79**	**39.85**	**0.06**	**962.149**	**1**	**0.001**

Sequences are tested for the various bit rates (e.g., 32 k bits/s to 512 k bits/s).

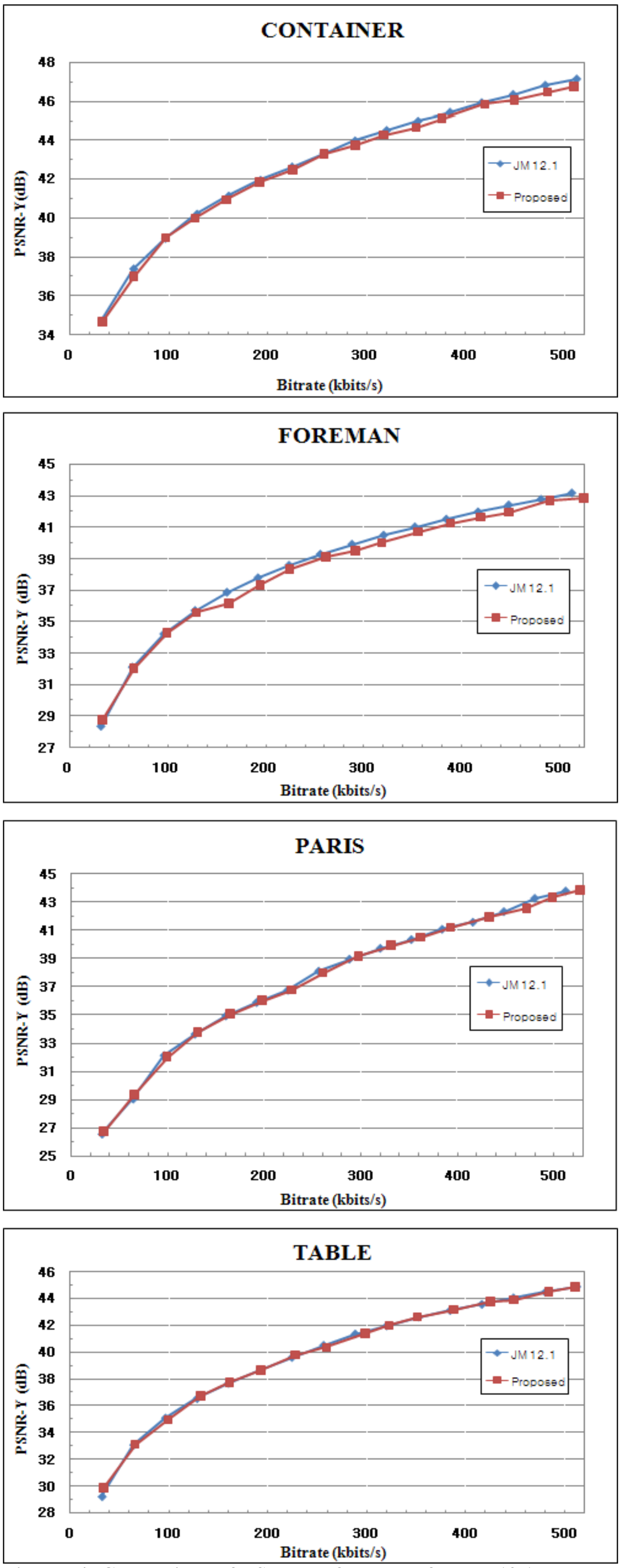

Figure 4. Comparison of PSNR performance for JM 12.1 and our proposed scheme.

IV. CONCLUSION

In this paper, we have presented an efficient real-time rate control scheme without skipping the frame to effectively allocate the number of bits for H.264/AVC video encoding. Our new and simple frame complexity measurement is developed to enhance the existing MAD-based method and is applied to our bit allocation for real-time rate control. QP accuracy is very important to prevent the overflow or underflow to target channel of the low bandwidth. Therefore, we have presented a QP control scheme to adjust the computed QP mainly based on actual encoding results of previous coded frames.

As demonstrated in our experiments, in comparison to H.264/AVC rate control, our proposed algorithm achieves higher average PSNR with smoother visual quality. The actually produced bits by each frame are closer to the target bits. Furthermore, the proposed improvements can be extended to unit and MB level rate control.

ACKNOWLEDGMENT

This work was supported by Seoul R&BD Program (10544).

REFERENCES

[1] ISO-IEC/JTC1/SC29/WG11, *Information technology—coding of audio-visual objects—part 10: advanced video coding* Final Draft International Standard, ISO/IEC FDIS 14 496-10, December 2003.

[2] T.Wiegand, “Draft ITU-T recommendation and final draft international standard of joint video specification (ITU-T Rec. H.264—ISO/IEC 14496-10 AVC),” in *Joint Video Team (JVT) of ISO/ICE MPEG and ITU-T VCEG, VT-G050*, Pattaya, Thailand, March 2003.

[3] T. Sikora, “Trends and perspectives in image and video coding,” *Proceedings of the IEEE*, vol. 93, no. 1, pp. 6–17, 2005.

[4] T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra, “Overview of the H.264/AVC video coding standard,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 560–576, 2003.

[5] G. J. Sullivan and T. Wiegand, “Video compression—from concepts to the H.264/AVC standard,” *Proceedings of the IEEE*, vol. 93, no. 1, pp. 18–31, 2005.

[6] G. J. Sullivan, T. Wiegand, T. Stockhammer, *Using the draft H.26L video coding standard for mobile applications*, IEEE International Conference on Image Processing 2001, Thessaloniki, Greece, October 2001.

[7] S. W. Ma, W. Gao, F. Wu, and Y. Lu, “Rate control for JVT video coding scheme with HRD considerations,” in *Proceedings of the IEEE International Conference on Image Processing (ICIP ’03)*, vol. 3, pp. 793–796, Barcelona, Spain, September 2003.

[8] G. Sullivan, T. Wiegand and K. P. Lim, "Joint model reference encoding methods and decoding concealment methods," Joint Video Team of ISO/IEC MPEG and ITU-T VCEG, JVT-1049, September 2003.

[9] Z. Li, W. Gao and F. Pan *et al*, "Adaptive rate control with HRD consideation," Joint Video Team of ISO/IEC MPEG and ITU-VCEG 8th Meeting, JVT-H014.

[10] Z. Li, F. Pan *et al*., “Adaptive basic Unit Layer Rate Control for JVT,” Joint Video Team of ISO/IEC MPEG and ITU-VCEG 7th Meeting, JVT-G012.

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

[12] Z. He, Y. K. Kim and S. K. Mitra, "Low delay rate control for DCT video coding via p domain source modeling," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, no 8, pp 928-940, Aug 2001.

[13] F. Pan, Z. G. Li and K. P. Lim et al, "Adaptive intra-frame quantization for very low bit rate video coding," in *Proc. Int. Symp. Circuits and Systems, ISCAS'*04, pp 781-784.

[14] N. Kamaci, Y. Altunbasak and R. M. Mersereau, "Frame bit allocation for the H.264/AVC video coder via Cauchy-Density-Based rate and distortion models," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 8, pp 994-1006, Aug 2005.

[15] H.-J. Lee, T. Chiang, and Y.-Q. Zhang, “Scalable rate control for MPEG-4 video,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 10, no. 6, pp. 878–894, Sep. 2000.

[16] F. Pan, Z. Li, K. Lim, and G. Feng, “A study of MPEG-4 rate control scheme and its improvements,” *IEEE Trans. Circuit Syst. Video Technol.*, vol. 13, 440-446, 2003.

[17] S. Ma, Z. Li, and F. Wu, Proposed Draft of Adaptive Rate Control, JVT-H017. Geneva, Switzerland, May 2003.

[18] JM12.1, http://iphome.hhi.de/suehring/tml/download/