Tao ZHU[1], Xiong-wei ZHANG[2]

Institute of Communications Engineering, PLA Univ. of Sci. & Tech. (1), Institute of Command Automation, PLA Univ. of Sci. & Tech. (2)

Rate Control Scheme for Spatial Scalability of H.264/SVC

Abstract. *This paper presents a rate control scheme for spatial scalability of H.264/SVC. At first, an adaptive method is proposed to decide the initial quantization parameter. Then, based on the transform coefficients distribution of spatial scalability, a new natural logarithm domain rate-distortion model is derived from the inspiration of the Cauchy-Density-Based model. Next, with above method and model, a rate control scheme is developed to match the characteristic of spatial scalability. The excellent performance of the proposed scheme is demonstrated by experimental results.*

Streszczenie. *Artykuł przedstawia metodę kontroli szybkości system GH.264/SVC. Przy wykorzystaniu metody adaptacyjnej ustlany jest współczynnik kwantyzacji. Następnie bazując na rozkładzie przestrzennej skalowalności stosuje się nowy algorytm dopasowujący szybkość transmisji. (**Kontrola szybkości transmisji dla przestrzennej skalowalności systemu kodowania H.264/SVC**)*

Keywords: rate control; spatial scalability; rate-distortion (R-D) model; H.264/SVC
Słowa kluczowe: kodowanie H.264, szybkość transmisji.

1 Introduction

Modern video application system imposes many kinds of constraints for video coding, such as the constraints of channel bandwidth and transmission delay in video communication. As an interface between codec and channel, to achieve better reconstructed video quality, rate control adjusts the bit rate of the bit stream to match the channel bandwidth by deciding the Quantization Parameter (QP) at different control levels (such as frame layer and micro-block layer), while optimizing the overall coding performance, and prevents the buffer from overflow and underflow. Recent years, many kinds of rate control schemes have appeared, such as analytical methods-based schemes [1] and regressive search-based schemes [2], single-pass schemes [3] and two-pass [4] schemes and so on. Some of them are quickly adopted as the recommended rate control scheme by video coding specifications, such as TM5 [5] for MPEG-2, TMN8 [6] for H.263, VM8 [7] for MPEG-4, JVT-G012 [8] for H.264/AVC and so forth. Nevertheless, these schemes only focus on the non-scalable video coding or single layer video coding.

Scalable video coding has become the important point of the standardization process and the newest scalable video coding specification is the Scalable Video Coding (SVC) extension of H.264 Advance Video Coding (AVC) standard, which supports three kinds of scalabilities, i.e. temporal, spatial and quality scalability [9]. In coding performance comparison with video simulcast [10], spatial scalability of H.264/SVC can provide higher coding efficiency with the concept of inter-layer prediction, which consists of three kinds of inter-layer prediction mechanisms, i.e. Inter-Layer Motion Prediction, Inter-Layer Residual Prediction and Inter-Layer Intra Prediction. As depicted in Fig. 1, spatial scalability of H.264/SVC is implemented based on the multilayer coding structure. Each spatial layer is attached to a specified spatial resolution and indexed by a so-called dependency identifier *DId*. In each layer, inter-coding and intra-coding are performed as the usual way in single layer coding. However, this kind of multilayer coding structure would decrease the overall coding performance. Accordingly, the inter-layer prediction mechanisms are employed and the decoding of a spatial layer is dependent on its reference layers data. Obviously, spatial layers with smaller *DId* are more important than other layers and the number of bits of reference layers need to be counted into the bit-rate of a specified layer. We have to take these features into consideration during the development of rate control scheme for spatial scalability of H.264/SVC.

Although the reference software Joint Scalable Video Model (JSVM) [11] only adopts a JVT-G012-like rate control scheme for its base layer, some initial research has been carried out for rate control of spatial scalability. Y. Liu [12] proposed a rate control scheme for spatial scalability and a switched model to predict the mean absolute difference (MAD) with the multilayer structure. Recently, the famous ρ-domain model has been incorporated into the rate control scheme for spatial scalability [13].

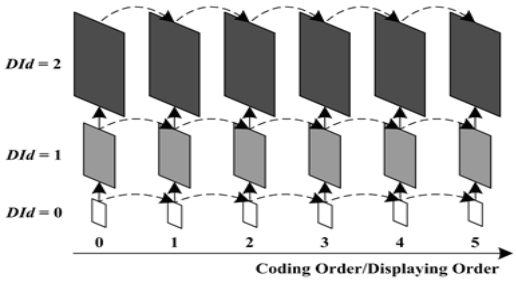

Fig.1. The multilayer structure of spatial scalability without encoding delay (i.e. B frame)

The rest of this paper is organized as follows. In section 2, an adaptive initial QP calculation method is proposed. In section 3, a quadratic logarithmic rate-distortion model is introduced based on the analysis of transform coefficient distribution characteristic. In section 4, based on above methods and models, a rate control scheme is proposed for spatial scalability. In section 5, the experimental results are presented to verify the performance of the proposed rate control scheme. At last, a conclusion is carried out.

2 Adaptive Initial QP Decision

The initial QP (denoted as "QP^{I_n}") is the QP for the first frame encoding of a sequence. Its value is a key factor that affects the performance of a rate control scheme. Generally speaking, under the restrained bit-rate of the rate control scenario, a sequence with high motions or complicated contents needs a large enough QP^I to save bits for other pictures in sequence; whereas, a small QP^I is suitable for sequences with low motions or simple contents to ensure the quality of reconstructed video as well as possible. Conventionally, QP^I is only decided by the bits per pixel (BPP) as JVT-G012 does. However, according to above analysis a better QP^I decision method should consider BPP and the video content at the same time.

This paper presents a gradient-based method to decide QP^I. The gradient is an important characteristic of an image. At first, the complexity of video content is measured by the gradient per pixel (GPP) of the video frame as follows

(1) $$\mathrm{GPP}_I = \frac{1}{W \cdot H}\left[\sum_{i=1}^{H-1}\sum_{j=1}^{W-1}\left(\left|I_{i,j} - I_{i+1,j}\right| + \left|I_{i,j} - I_{i,j+1}\right|\right)\right],$$

where $I_{i,j}$ is the intensity value at the location of (i, j), W and H are the image width and height respectively. For a frame with Y, Cb and Cr, the weighting GPP is calculated as

(2) $$\mathrm{GPP} = \left(4 \cdot \mathrm{GPP}_Y + \mathrm{GPP}_{Cb} + \mathrm{GPP}_{Cr}\right)/6.$$

To illustrate the relationship between the GPP and the bits number of I frame, at a specified QP the scatter plots of I frame bits versus GPP for spatial enhancement layers of nine sequences are shown in Fig. 2. Obviously, we can find that the frame bits number is proportional to its GPP and the relationship between them can be assumed to be linear.

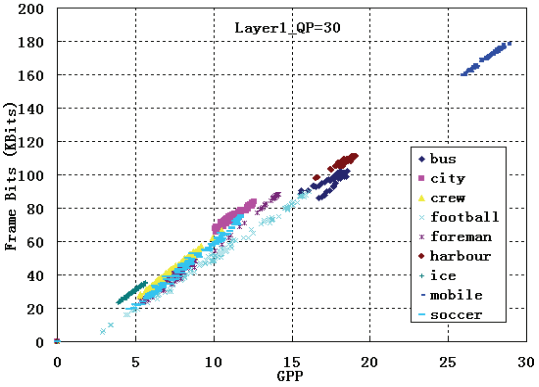

Fig.2. Scatter plots of frame bits versus GPP

Then, the concept of BPP has to be considered for the spatial scalability. Originally, according to [11], the BPP is

(3) $$\mathrm{BPP} = \frac{R}{F \times N_{pixel}},$$

where R is the target rate, F is the frame rate and N_{pixel} is the number of frame pixels. Because of the dependency relation between different spatial layers resulted from the inter-layer prediction mechanisms, the reference layers bits number of an enhancement layer should be taken into account when calculating the bit rate of this enhancement layer. Hence, the BPP concept should be redefined as

(4) $$\mathrm{BPP} = \frac{R_e - \sum_{i \in \Phi} R_r(i)}{F \times N_{pixel}},$$

where R_e is the target rate for the enhancement layer, Φ denotes the set of reference layers and R_r is the target rate for the reference layer, F is the frame rate and N_{pixel} is the number of picture pixels at this spatial layer.

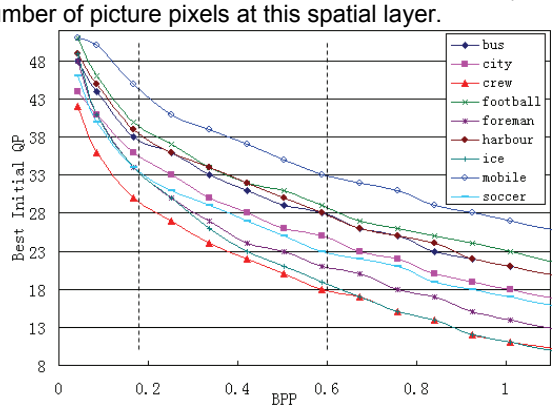

Fig.3. Relationship between best initial QP and BPP.

To achieve the best initial QPs (denoted as "QP^I_{Best}") for various BPPs, all possible values of QP^I are tested and the one that produces the best R-D performance is considered as QP^I_{Best}. The experimental results of nine sequences are depicted in Fig. 3, where the BPP axis can be partitioned into three regions: region-1 (BPP < 0.18), region-2 (0.18 ≤ BPP < 0.6), and region-3 (0.6 ≤ BPP). Within each of these regions, for most cases the linear decrease model can be used to approximate the relation between QP^I_{Best} and BPP.

Further more, comparing Fig. 2 with Fig. 3, it can be observed that the sequences with larger GPP in Fig. 2 correspond to upper curves in Fig. 3. This phenomenon means that GPP is an important factor to describe the relation between QP^I_{Best} and BPP. Taking the middle bit-rate (0.18 ≤ BPP < 0.6) as an example, we can obtain scatter plots of QP^I_{Best} versus GPP for three specified BPPs as depicted in Fig. 4, from which we can conclude that the relation of QP^I_{Best} versus GPP is approximately linear. After careful investigation, the three segments linear relation of QP^I_{Best}, GPP and BPP can be approximated by

(5) $$\mathrm{QP}^I_{Best} = \begin{cases} 43.49 + 0.59 \times \mathrm{GPP} - 106.45 \times \mathrm{BPP} & \mathrm{BPP} \le 0.18 \\ 25.12 + 0.69 \times \mathrm{GPP} - 29.23 \times (\mathrm{BPP} - 0.18) & 0.18 < \mathrm{BPP} \le 0.6 \\ 13.93 + 0.74 \times \mathrm{GPP} - 18.40 \times (\mathrm{BPP} - 0.6) & 0.6 < \mathrm{BPP} \end{cases}.$$

For the first I frame in the sequences, after its gradient information is calculated, the best QP for this I frame can be decided under certain BPP value.

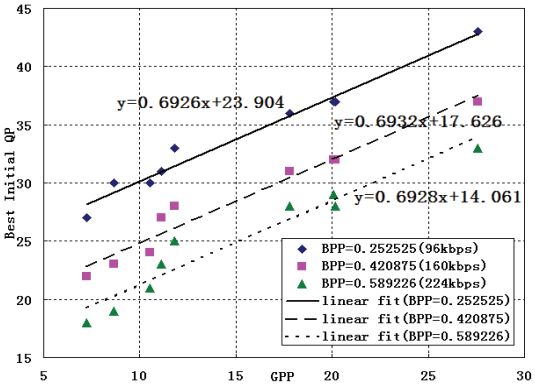

Fig.4. Linear relation between best initial QP and GPP

3 A Natural Logarithm Domain R-D Model

The R-D behaviour of video encoder is characterized by the rate-quantization (R-Q) model and the distortion-quantization (D-Q) model, which are jointly called as R-D model. In the past years, several R-D models have been proposed from various points of view. The most classical quadratic model was deduced from the Taylor expanding of the R-D function based on the Laplacian distribution [14]. According to the property of the DCT zeros coefficients percentage ρ, He et al. designed a ρ-domain model [15] to describe the relationship between rate and ρ. Due to the closer DCT coefficient statistical distribution to the Cauchy distribution than to the Laplacian distribution, a Cauchy-Density-Based R-D model has been proposed by Kamaci [16]. Since the header bits may occupy a larger portion of the total bits, rate model can be developed for the texture bits and its related header bits respectively [17]. Recently, an exponential rate model [18] was proposed and can be used to elegantly resolve the "chicken and egg dilemma" in H.264/AVC rate control. However, scarcely any study about R-D model of scalable video coding has been carried out.

Knowledge of transform coefficient distribution can be employed to investigate the R-D characteristic of transform coding based video encoder. Thereby, it is necessary to examine the transform coefficient distribution of spatial scalability. With the Maximum Likelihood Estimation method and the luminance components transform coefficient data of spatial scalability on enhancement layers, two goodness-of-fit tests, the Kolmogorov-Smirnov (KS) test and the χ^2 test, are performed to decide the best distribution among three famous distributions, the Generalized Gaussian distribution

(GGD), the Laplacian distribution (LAP) and the Cauchy distribution (CCH), for the transform coefficients of video sequence encoded by H.264/SVC encoder. For nine video test sequences, the KS and χ^2 statistic are computed for each location of Zig-Zag scan order under above three distribution models. The distribution that gives the minimum χ^2 statistic is chosen as the best fit model. The number of certain model chosen as the best fit is sum up for every frame type, and then the selected data of coefficients with Zig-Zag index 0, 5, 15 (C0, C5, C15) are listed in Table 1. These results demonstrate that the CCH can be considered as the best distribution to model the coefficients for spatial scalability. That is to say that the Cauchy distribution can still be considered as the best description for the transform coefficients of spatial scalability in H.264/SVC.

Table 1. The number of models chosen as the best fit

Data		GGD	LAP	CCH
I	C0	0	0	9
	C5	2	0	7
	C15	2	0	7
	Total	5	0	23
P	C0	4	0	5
	C5	5	0	4
	C15	4	1	4
	Total	13	1	13
B	C0	1	0	8
	C5	3	0	6
	C15	0	0	9
	Total	4	0	23
Total		22	1	59

From [16], we know that the R-D model based on the Cauchy distribution can be expressed as follows:

$$\begin{cases} \ln R(\mathrm{Q}) = a - \alpha \ln(\mathrm{Q}) \\ \ln D(\mathrm{Q}) = b + \beta \ln(\mathrm{Q}) \end{cases}. \tag{6}$$

However, because of the only consideration for source, this model is not accurate enough in actual encoder, especially for H.264/SVC. In H.264/SVC, many kinds of coding modes are employed to reduce redundancy. This feature leads to the increase of the header bits relative to the previous specifications and the difficulty of designing a model for it. Consequently, the relationship between total bits and QS doesn't follow the model of (6) for most cases. Nevertheless, motivated by the Cauchy-Density-Based R-D model, we can investigate the relationships between the frame total bits, the frame distortion (represented by mean-square-error) and QS in natural logarithm domain. At first, the H.264/SVC encoder of JSVM [11] is exhaustively run on sequences with different QSs and the numbers of actual total bits used for encoding frames and their Mean Square Error (MSE) are recorded respectively. Then, through performing curve fitting on extensive testing data of I, P and B frame, it is find that these R-D relationships can be modelled by quadratic function in natural logarithm domain as illustrated in Fig. 5. Hence the R-D model in natural logarithm domain is expressed by

$$\begin{cases} \ln R(Q) = a\ln^2(Q) + b\ln(Q) + c \\ \ln D(Q) = d\ln^2(Q) + e\ln(Q) + f \end{cases}, \tag{7}$$

where a, b, c, d, e, f are the model parameters, R is the total bits for a frame, Q is the quantization step size. From Fig. 5, this R-D model can accommodate to very broad quantization range, from low rate end to high rate end. And if let Q = 1, namely ln$Q$ = 0, we can obtain c = ln$R$ and f = ln$D$, which are invariant with Q and describe the inherent coding complexity of certain video encoder from rate and distortion point of view respectively.

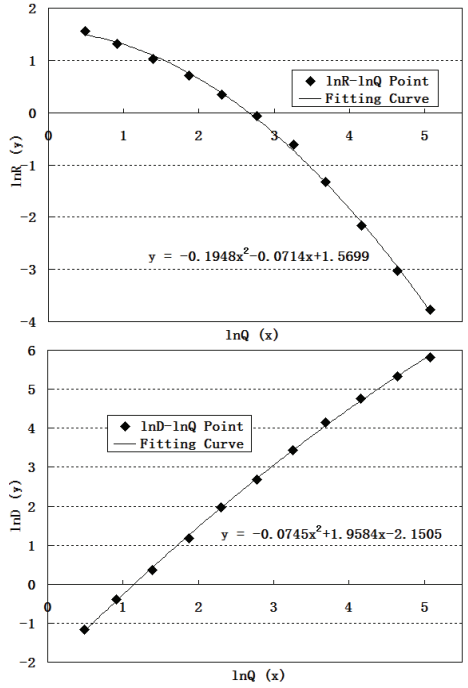

Fig.5. The curve fitting of lnR-lnQ and lnD-lnQ points for spatial enhancement layer

Our proposed rate control scheme regulates the bit-stream at frame level in every spatial layer. Namely, the proposed scheme can be considered as the single layer rate control applied in every spatial layer with some small changes. The main flowchart for each spatial layer of the proposed scheme is depicted in Fig. 6. QP^I calculation is presented in section 2; the used R-Q model is introduced in section 3 and its update is dependent on the method of linear regression as JVT-G012; hence, buffer fullness pre-estimation, frame base QP calculation, frame bit allocation and QP calculation, and final QP decision are our mainly focus in this section. Other details are similar to JVT-G012.

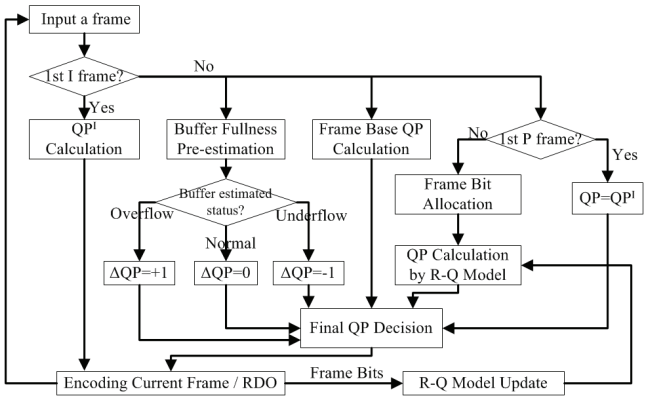

Fig.6. The flowchart of the proposed rate control scheme

4.1 Buffer Fullness Pre-estimation

To avoid undesirable buffer overflow and underflow, a buffer fullness pre-estimation mechanism is developed. Its key is the bits number estimation of frame to be encoded, which is the average frame bits number in a slide window as

$$B_{est}(i) = \sum_{j=\max(1,i-WinSize)}^{i-1} B_{act}(j) \Big/ \min(WinSize, i-1), \tag{8}$$

where $B_{est}(i)$ and $B_{act}(i)$ are the estimated and actual bits number of i^{th} frame on certain layer respectively, *WinSize* is the size of slide window and set as 5 usually. And then, the buffer fullness can be estimated as

$$Buf_{est}(i) = Buf_{act}(i-1) + B_{est}(i) + B_{ref}(i) - R/F, \quad (9)$$

where $Buf_{est}(i)$ and $Buf_{act}(i)$ are the estimated and actual buffer fullness after i^{th} frame encoded respectively, R is the channel bit rate of current layer, F is the frame rate, $B_{ref}(i)$ is the bits sum of reference layers. Finally, the QP increment ΔQP can be attained as

$$\Delta QP = \begin{cases} +1 & Buf_{est} > 0.85 BufSize \\ 0 & \text{other} \\ -1 & Buf_{est} < 0.15 BufSize \end{cases}. \quad (10)$$

From (10), if $Buf_{est} > 0.85 BufSize$, overflow may occur and we increase the frame QP; if $Buf_{est} < 0.15 BufSize$, underflow may happen and we decrease the frame QP.

4.2 Rate Control on Each Layer

This part of rate control is to determine the QP for frame encoding with frame budget bits, buffer status and QPs of past frames. This procedure is designed as follows:

Step 1: Frame base QP (QP_{base}) calculation. QP_{base} is an estimation of current frame QP from the QPs of past several frames. Similar to (8), slide window is also used as follows

$$QP_{base}(i) = \sum_{j=\max(1,i-WinSize)}^{i-1} QP(j) \Big/ \min(WinSize, i-1), \quad (11)$$

where $QP(i)$ is the actual QP used to encode frame i.

Step 2: Frame bit allocation and QP calculation. Frame bit allocation is as follows

$$Bit(i) = \eta \left[\frac{R}{F} - B_{ref}(i) \right] + (1-\eta) B_{est}(i), \quad (12)$$

where η is a weighting factor and empirically set to 0.75. QP is transformed from QS that is the solution of R-Q model in (7), according to the relation between QP and QS in H.264.

Step 3: Final QP decision. Final QP is computed as

$$QP = (1-\omega) QP_{base} + \omega QP_{Calc} + \Delta QP, \quad (13)$$

where ω is determined before i^{th} frame encoded as

$$\omega = \min\{0.75, \max\{0.25, (i-1)/WinSize\}\}. \quad (14)$$

Finally, QP should be restricted in the usually used range

$$QP = \min\{50, \max\{12, QP\}\}. \quad (15)$$

After the frame QP is derived, frame encoding can be performed. And the R-Q model is updated by new frame bits to prepare for the QP calculation of next frames.

5 Experimental Results

The experiments are performed on JSVM 9.18, which provides a software tool called "FixedQPEncoder" (denoted as "FixQP"). FixQP adopts a logarithmic search algorithm to find a QP for meeting a bit rate and its performance can be used as a benchmark to compare with other rate control schemes. Many sequences are tested by FixQP and our proposed scheme (denoted as "ProRC").

To evaluate the performance of the proposed scheme, several video sequences are tested. At first, the anchor bit rates for each spatial layer were decided by encoding test sequences at fixed QPs, i.e. 24, 30, 36, 42 for layer 0 and layer 1, using scalable encoder of JSVM 9.18. Then FixQP and ProRC set these anchor bit rates as target bit rates for corresponding spatial layer. Finally, the performance difference between FixQP and ProRC can be evaluated by delta PSNR of Rate-Distortion (R-D) curves. The search range for motion estimation is 32 pixels; the context-based adaptive binary arithmetic coding is the entropy coding. For FixQP, the encoding iterations (Itr.) number is not allowed to exceed 15 and the acceptable mismatch range is ±1%. The buffer size is set as 0.5×bit rate. All encoding structures of sequences are IPPP.

Table 2. The final rate control result for spatial scalability

Seq.	Layer No.	Tgt. Bitr. (kb/s)	Meth.	Act. Bitr. (kb/s)	PSNR (dB)	Itr.
bus	0	86.872	FixQP	87.378	31.948	4
			ProRC	86.528	31.903	1
	1	138.86	FixQP	139.82	29.558	3
			ProRC	139.78	29.562	1
city	0	21.062	FixQP	21.818	33.201	13
			ProRC	21.352	33.186	1
	1	95.424	FixQP	95.832	33.563	3
			ProRC	95.517	33.328	1
crew	0	67.093	FixQP	67.398	32.611	1
			ProRC	67.251	32.532	1
	1	232.94	FixQP	234.08	34.138	1
			ProRC	233.05	34.116	1
foot -ball	0	133.80	FixQP	132.49	32.37	5
			ProRC	132.89	32.384	1
	1	891.80	FixQP	893.46	37.297	1
			ProRC	892.26	37.302	1

As in Table 2, "Seq." represents names of sequences, "Tgt. Bitr." and "Act. Bitr." are the target rate and actual rate for rate control respectively and their unit is "kbps". "PSNR" is the weighting average Peak Signal-to-Noise Ratio of three components, namely Y, Cb and Cr, in video sequence, namely defined as

$$PSNR = \frac{4 \times YPSNR + CbPSNR + CrPSNR}{6}, \quad (16)$$

and its unit is "dB".

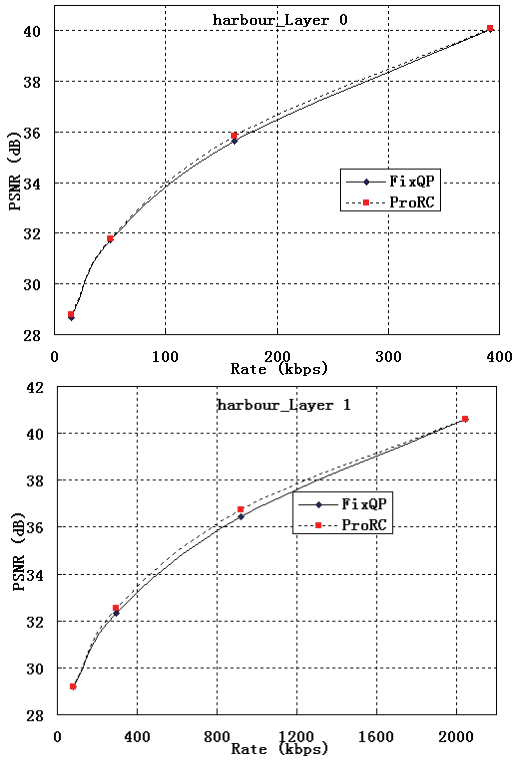

Fig.7. The R-D curves of ProRC and FixQP at different layers for test sequence "harbour"

The R-D curves of ProRC and FixQP at different layers are depicted in Fig. 7, where the difference between two R-D curves, i.e. ΔPSNR, denotes the performance increment of ProRC relative to FixQP. ΔPSNR can be calculated as

the contribution of Bjontegaard [19]. Some of the calculation results are list in Table 3.

Table 3. The performance increment of ProRC relative to FixQP

Sequence		foreman	harbour	bus	mobile	soccer
ΔPSNR (dB)	L0	0.156	0.118	-0.02	0.073	0.038
	L1	0.133	0.176	-0.04	0.080	0.028

An example of buffer status for spatial scalable coding is illustrated in Fig. 8. With the method of ProRC, the buffer is well controlled at different spatial layers.

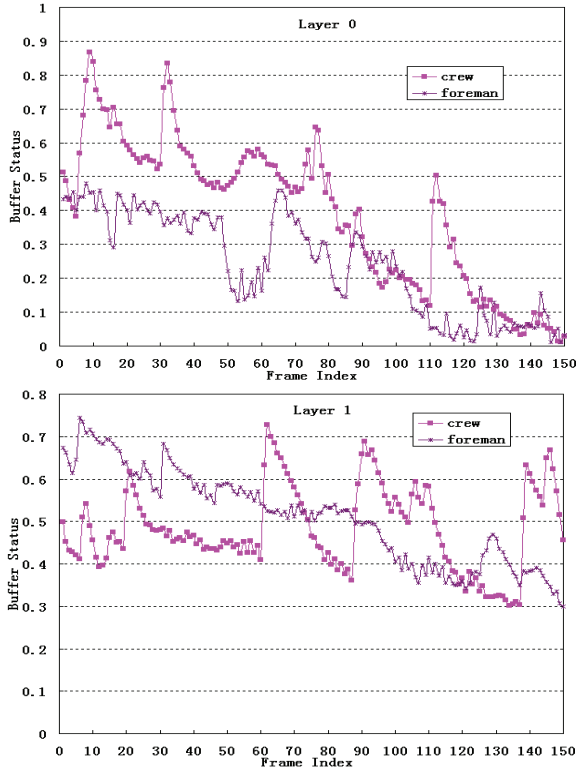

Fig.8. The buffer status of the proposed rate control scheme

From above experimental results, compared with FixQP, ProRC has performance characteristics as follows. At first, ProRC can achieve more accurate bit-rate control with only one encoding iteration instead of the several iterations that FixQP needs as demonstrated in Table 2. Then, ProRC has comparable or a little better performance with FixQP as explained in Fig. 7 and Table 3. Thirdly, reasonable buffer control is achieved by ProRC with no buffer overflow and underflow. Hence, it can be concluded that the experimental results prove the good performance of ProRC at several specified target rates.

6 Conclusions

This paper proposes a rate control scheme for spatial scalability of H.264/SVC. Based on the frame gradient, the best initial QP can be decided by an adaptive method. From the research on R-D characteristic of transform coefficient for frame in spatial multilayer structure, the relation between rate and distortion can be described by a quadratic R-D model in natural logarithm domain. Then, using above method and model, the proposed rate control scheme for spatial scalability can match the target rates at each spatial layer. On the basis of experiment results, we can conclude that the developed rate control scheme can provide good performance and can be expected to be used in practical video application.

Acknowledgement

The authors wish to acknowledge the financial support of Natural Science Foundation of Jiangsu Province in 2009 (BK2009059).

REFERENCES

[1] Y. Liu, Z. G. Li, Y. C. Soh, A Novel Rate Control Scheme for Low Delay Video Communication of H.264/AVC Standard, *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 1, Jan. 2007, pp. 68-78.

[2] W. Yuan et al., Optimum bit allocation and rate control for H.264/AVC, *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, no. 6, Jun. 2006, pp. 705-715.

[3] Z. G. Li, W. Gao et al, Adaptive rate control for H.264, *J. Vis. Commun. Image R.*, vol. 2006, no. 17, 2006, pp. 376-406.

[4] D. D. Zhang, K. N. Ngan et al, A two pass rate control algorithm for H.264/AVC high definition video coding, *Sig. Proc.: Image Commun.*, vol. 2009, no. 24, 2009, pp. 357-367.

[5] MPEG-2, Test Model 5, *Doc. ISO/IEC JTC1/SC29 WG11/93-400*, 1993.4.

[6] ITU-T/SG15. Video Codec Test Model, Near-Term, TMN8 ITU Study Group 16, *Video Coding Experts Groups*, Portland, USA, 1997, Doc. Q15-A-59.

[7] MPEG-4 Video Verification Model v8.0, *ISO/IEC JTC1/SC29/WG11 Coding of Moving Pictures and Associated Audio MPEG97/N1796*, 1997.7.

[8] Z. G. Li, F. Pan et al, Adaptive basic unit layer rate control for JVT, *ISO/IEC JTC1/SC29/WG11 and ITU-T SG16/Q6*, Doc. JVT-G012, In the 7th Meeting, Pattaya II Thailand, 2003.3.

[9] H. Schwarz, D. Marpe, T. Wiegand, Overview of the scalable video coding extension of the H.264/AVC standard, *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 9, pp. 1103-1120, 2007.9.

[10] M. Wien, H. Schwarz, T. Oelbaum, Performance Analysis of SVC, *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 9, pp. 1194-1203, 2007.9.

[11] Julien Reichel, Heiko Schwarz, Mathias Wien, Joint Scalable Video Model JSVM-12 text, *ISO/IEC JTC1/SC29/WG11 and ITU-T SG16/Q6*, Doc. JVT-Y202, In the 25th Meeting, Shenzhen, 2007.10.

[12] Y. Liu, Z. G. Li, Y. C. Soh, Rate control of H.264/AVC scalable extension, *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, no. 1, pp. 116-121, 2008.1.

[13] Y. Pitrey, M. Babel, O. Déforges, J. Viéron, *ρ*-domain based rate control scheme for spatial, temporal, and quality scalable video coding, *In Visual Comm. and Image Proces. 2009, SPIE Electronic Imaging*, vol. 7257, 2009.1.

[14] T. Chiang, Y.-Q. Zhang, A new rate control scheme using quadratic rate distortion model, *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, no. 1, pp. 246-250, 1997.2.

[15] Z. He, S. K. Mitra, A unified rate-distortion analysis framework for transform coding, *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, no. 12, pp. 1221-1236, 2001.12.

[16] N. Kamaci, Y. Altunbasak, R. 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, 2005.8.

[17] Do-Kyoung Kwon, Mei-Yin Shen, C.-C. Jay 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, 2007.5.

[18] Yimin Zhou Yu Sun, Zhidan Feng, Shixin Sun, New rate-distortion modeling and efficient rate control for H.264/AVC video coding, *Signal Processing: Image Communication*, vol. 24, pp. 345-356, 2009.2.

[19] Gisle Bjontegaard. Calculation of average PSNR differences between RD-curves. *ITU-T SG16/Q6*, Doc. VCEG-M33, In the 13th Meeting, Austin, Texas, USA, 2001.4.

***Authors**:*
Tao Zhu, Postgraduate Team 2, Institute of Communications Engineering, Biaoyin 2, Yudao Street, Nanjing, China, 210007, E-mail: zhutao007@gmail.com;
Professor Xiong-wei Zhang, Institute of Command Automation, PLA Univ. of Sci. & Tech..