Improved Rate Control Algorithm for Scalable Video Coding*

Xin Lu and Graham R. Martin

Department of Computer Science
University of Warwick, Coventry
CV4 7AL, United Kingdom
{xin, grm}@dcs.warwick.ac.uk

Abstract

In the Scalable Video Coding (SVC) standard, a multi-layer based structure is utilised to support scalability. However in the latest Joint Scalable Video Model (JSVM) reference software, the rate control algorithm is implemented only in the base layer, and the enhancement layers are not equipped with a rate control scheme. In this work, a novel rate control algorithm is proposed for when inter-layer prediction is employed. Firstly, a Rate-Quantisation (R-Q) model, which considers the coding properties of different prediction modes, is described. Secondly, an improved Mean Absolute Difference (MAD) prediction model for the spatial enhancement layers is proposed, in which the encoding results from the base layer are used to assist the linear MAD prediction in the spatial/CGS enhancement layers. Simulation results show that, on average, rate control accuracy is maintained to within 0.07%. Compared with the default JVT-G012 rate control scheme employed in SVC, the proposed rate control algorithm achieves higher coding efficiency, namely an improvement of up to 0.26dB in PSNR and a saving of 4.66% in bitrate.

1998 ACM Subject Classification E.4 Coding and Information Theory

Keywords and phrases Inter-layer prediction, MAD prediction, Rate control, Scalable Video Coding (SVC), SVC extension of H.264/AVC.

Digital Object Identifier 10.4230/OASIcs.ICCSW.2013.73

1 Introduction

Scalable Video Coding (SVC), the scalable extension of the H.264/AVC standard, provides solutions for video applications with different network bandwidths, device capabilities and user demands. Bandwidth is a valuable resource, and often there are situations in which the bandwidth is insufficient or the network is unstable. Sometimes even the bitstream comprising the basic layer cannot be transmitted completely, resulting in frame skipping. In these cases, effective rate control is essential. In real-time applications, rate control enables the output bit rate to adjust quickly depending on the available channel bandwidth. With a proper control scheme, "overflow" and "underflow" of the buffer are prevented, which means that frame skipping and wastage of channel resources can be avoided. Thus, rate control extends the scalability of SVC. Furthermore, rate control appropriately allocates the available bits according to the complexity of the image content, so that the quality of the video is maximised.

Several rate control algorithms have been proposed for non-scalable video coders, such as the Test Model 5 (TM5) [10] for MPEG-2, Test Model Near-term 8 (TMN8) [5] for

* A longer version of this paper appeared in the proceedings of IEEE MMSP'13 [9].

2013 Imperial College Computing Student Workshop (ICCSW'13).
Editors: Andrew V. Jones, Nicholas Ng; pp. 73–80
OpenAccess Series in Informatics
OASICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany
ICCSW

H.263, Verification Model 8 (VM8) [11] for MPEG4 and JVT-G012 [6] for H.264/AVC. The JVT-G012 rate control algorithm is only implemented in the base layer of the latest JSVM reference software, and it does not support the enhancement layers which provide the scalability functions. Rate control algorithms that address the properties of the enhancement layers in SVC need to be developed. Several rate control algorithms have been suggested for SVC [16, 8, 4, 7]. They consider either precise target bit allocation or the optimisation of the rate quantisation model. Xu *et al.* proposed a rate control algorithm for spatial and Coarse Gain SNR (CGS) scalable coding in SVC [16]. This method employs the improved TMN8 model for quantisation parameter (Qp) estimation based on the mode analysis of I, P, and B frames. In [8], Liu *et al.* proposed that the MAD can be predicted from either the previous frame in the same layer or the corresponding frame in the base layer through a switching law. Hu *et al.* [4] proposed a frame level rate control algorithm for temporal scalability of scalable video coding by developing a set of weighting factors for bit allocation. In [7], Liu *et al.* proposed a bit allocation algorithm for SVC when the inter-layer dependency is taken into consideration.

SVC employs so called inter-layer prediction to reduce the redundancies between layers. However, the effects of inter-layer prediction are not taken into consideration in the rate control scheme of the JSVM. The rate control strategies assume that the statistical property of a video source is fixed [15], and then they derive a precise rate distortion model [12]. From observation and analysis, macroblocks coded using inter-layer prediction and those coded by intra-layer prediction (inter-frame prediction and intra-frame prediction) have dissimilar statistical properties. Furthermore, some encoding results of the base layer can be used to inform the encoding of the enhancement layers, thus benefiting from the bottom-up coding structure of SVC. These observations motivate us to propose a rate control scheme with a precise Rate-Quantisation (R-Q) model and optimised MAD prediction for the spatial enhancement layers.

The remainder of this paper is organised as follows. Section 2 discusses the formulation of the proposed rate distortion model, and the proposed MAD prediction scheme is presented in Section 3. Extensive experimental results are presented in Section 4 and conclusions are given in Section 5.

2 Rate-Distortion Model for Spatial Enhancement Layer

With the hypothesis that the residual coefficients obey a Laplacian distribution, the classic quadratic rate-distortion (R-D) model is described as follows [3],

$$\mathrm{R_{txt}} = \frac{\mathrm{X_1} \times \mathrm{MAD_{pred}}}{\mathrm{Q_{step}^2}} + \frac{\mathrm{X_2} \times \mathrm{MAD_{pred}}}{\mathrm{Q_{step}}} \tag{1}$$

where $\mathrm{R_{txt}}$ is the target number of bits assigned to code the texture information of a basic unit; $\mathrm{MAD_{pred}}$ indicates the mean absolute difference of the residual component; $\mathrm{Q_{step}}$ is the quantisation step size to be calculated and $\mathrm{X_1}$ and $\mathrm{X_2}$ are model coefficients. $\mathrm{X_1}$ and $\mathrm{X_2}$ are updated using a linear regression method after the coding of each basic unit [3].

In SVC, the inter-layer prediction tools are employed to reduce the redundancies between layers. However, inter-layer prediction is not efficient for coding sequences containing homogeneous texture or slow motion, since the high frequency components in the enhancement layer cannot be reconstructed well by upsampling the information of the base layer. Macroblocks with little detail and slow motion are more likely to be best matched with a block by inter-frame prediction in the same layer. Consequently, temporal prediction (inter-frame

Table 1 Relationship between average number of texture bits per coded unit and Qp for both inter-layer predicted macroblocks and intra-layer predicted macroblocks.

Sequence	Prediction mode	Qp				
		28	32	36	40	44
Bus	Intra-layer	98.98	47.91	18.53	6.26	1.66
	Inter-layer	98.68	52.22	28.05	15.28	7.94
Football	Intra-layer	77.89	43.92	22.07	9.41	3.79
	Inter-layer	100.3	56.63	32.12	20.31	11.76
Foreman	Intra-layer	18.61	6.52	2.40	0.86	0.39
	Inter-layer	28.81	14.27	7.59	4.72	3.20
Mobile	Intra-layer	185.2	77.89	22.57	6.76	2.23
	Inter-layer	162.4	96.88	46.41	22.33	12.79

prediction) is more efficient than inter-layer prediction, especially for sequences with slow motion. In contrast, inter-layer prediction performs well for fast moving sequences. However, macroblocks with fast movement usually need more bits to encode. So we observe that the average number of texture bits for inter-layer predicted macroblocks in the enhancement layers is significantly greater than for macroblocks which have been coded using temporal and spatial prediction within the same enhancement layer.

The significant difference in the number of texture bits generated for inter-layer predicted macroblocks and intra-layer predicted macroblocks leads to serious prediction errors in the rate-distortion model coefficients. To illustrate the mathematical relationship between Qp and the texture bits for both inter-layer predicted macroblocks and intra-layer predicted macroblocks, and to justify our proposed algorithm, we analysed the simulation results from processing four sequences with different degrees of activity and detail. The statistics were collected from the first 150 frames of each video sequence. QCIF sequences were used for the base layer and CIF were used for the enhancement layer.

Table (1) shows the average number of texture bits for inter-layer predicted macroblocks in the enhancement layer and those for intra-layer predicted macroblocks in the same enhancement layer, under a variety of Qp values. It is observed that the average number of bits used for encoding inter-layer predicted macroblocks is significantly different from that required for intra-layer predicted macroblocks. In general, as Qp increases, significantly more bits are consumed by inter-layer prediction coding than by intra-layer prediction.

The model relationship between Qp and $\mathrm{Q_{step}}$ is

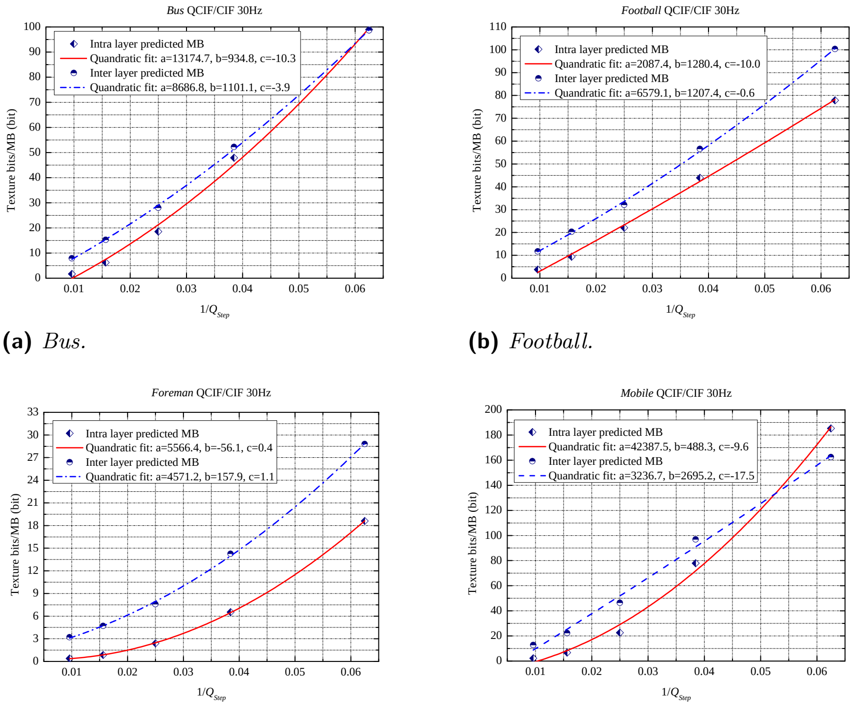
$$\mathrm{Qp} = 2^{\frac{\mathrm{Q_{step}}}{6}} \zeta(\mathrm{Q_{step}}\%6) \tag{2}$$

where $\zeta(0)$=0.675; $\zeta(1)$=0.6875; $\zeta(2)$=0.8125; $\zeta(3)$=0.875; $\zeta(4)$=1.0; $\zeta(5)$=1.125 [13].

Figure (1) illustrates the relationship between $\mathrm{Q_{step}}$ values and the obtained number of texture bits $\mathrm{R_{txt}}$ for both inter-layer predicted macroblocks and intra-layer predicted macroblocks. The quadratic curves fitting the measured data are also presented. It can be seen that the measured data can be represented by quadratic functions very well.

$$\mathrm{R_{txt}} = \frac{\mathrm{a}}{\mathrm{Q_{step}^2}} + \frac{\mathrm{b}}{\mathrm{Q_{step}}} + \mathrm{c} \tag{3}$$

where $\mathrm{c} \approx 0$. The coefficients of the quadratic model are obtained by finding the minimal fitting error. Although the observed $\mathrm{R_{txt}}$ - $\mathrm{Q_{step}}$ relationship can be represented by quadratic models, the model coefficients are significantly different. From these observations, it is concluded that for optimised rate control within the enhancement layers, separate, and



(a) *Bus.* **(b)** *Football.* **(c)** *Foreman.* **(d)** *Mobile.*

Figure 1 Relationship between average number of texture bits and Q_{step} for both inter-layer coding and non-inter-layer coding. Points are actual data; curves are fitted to the data.

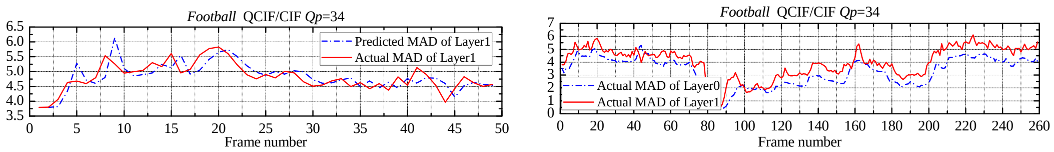
sufficiently accurate, models must be used for inter-layer prediction and intra-layer prediction. Consequently, $Q^{\text{inter}}_{\text{step}}$ and $Q^{\text{intra}}_{\text{step}}$ can be modelled accurately by quadratic functions with their respective model coefficients.

$$
\begin{aligned}
R_{\text{txt}} &= \frac{X_1^{\text{inter}} \times \text{MAD}_{\text{pred}}}{\left(Q^{\text{inter}}_{\text{step}}\right)^2} + \frac{X_2^{\text{inter}} \times \text{MAD}_{\text{pred}}}{Q^{\text{inter}}_{\text{step}}} \\
R_{\text{txt}} &= \frac{X_1^{\text{intra}} \times \text{MAD}_{\text{pred}}}{\left(Q^{\text{intra}}_{\text{step}}\right)^2} + \frac{X_2^{\text{intra}} \times \text{MAD}_{\text{pred}}}{Q^{\text{intra}}_{\text{step}}}
\end{aligned}
\qquad (4)
$$

where R_{txt} is the target number of texture bits for the current basic unit; MAD_{pred} is the MAD predicted from the previous coding results, $Q^{\text{inter}}_{\text{step}}$ and $Q^{\text{intra}}_{\text{step}}$ are the desired quantisation step sizes for inter-layer prediction and intra-layer prediction respectively, X_1^{inter} and X_2^{inter}, X_1^{intra} and X_2^{intra}, are the model coefficients for inter-layer prediction and intra-layer prediction, each updated after the coding of a basic unit using inter-layer prediction and intra-layer prediction respectively.

3 Optimisation of MAD Prediction for Spatial Enhancement Layer

From equations (4) it can be seen that the quantisation step sizes $Q^{\text{inter}}_{\text{step}}$ and $Q^{\text{intra}}_{\text{step}}$ depend on the model coefficients, the target number of bits R_{txt} for the current basic unit, and the predicted MAD value of the current basic unit. However, the MAD value is unknown before



(a) **(b)**

Figure 2 The MAD relationships (part of *Football* sequence) (a) Predicted and actual MAD values; (b) Actual MAD values of base layer and spatial enhancement layer.

rate-distortion optimisation (RDO). The MAD value can only be obtained after coding the current basic unit using the quantisation step size, but the model needs the MAD value to calculate the quantisation step size. This is a "chicken and egg situation". The JVT-G012 rate control algorithm overcomes this problem by using the MAD value of the basic unit in the same position of the previous frame to predict the MAD value of current basic unit, thus permitting the quantisation step size to be calculated. A linear MAD model is adopted here as [6]:

$$\mathrm{MAD_j} = \mathrm{a_1} \times \mathrm{MAD_{j-1}} + \mathrm{a_2} \tag{5}$$

where $\mathrm{a_1}$ and $\mathrm{a_2}$ are model coefficients, updated after the coding of each frame. $\mathrm{MAD_j}$ denotes the predicted MAD of the current basic unit and $\mathrm{MAD_{j-1}}$ denotes the actual MAD of the basic unit in the corresponding position of the previous frame.

In the quadratic R-D model, MAD prediction is very important as it directly affects the allocation of bits. As the prediction is not always accurate, there is always some small error in bit allocation, and this cannot be avoided in the JVT-G012 algorithm. As shown in Figure (2)(a), if the MAD fluctuates due to fast motion or scene changes in the video sequence, the linear model performs poorly, and there is always a delay. In the example, this phenomenon is particularly obvious at frames 9, 41 and 44.

In the SVC encoding process, for each frame, the base layer is encoded first, prior to the enhancement layers. Furthermore, the content of the base layer and enhancement layers are highly correlated. As shown in Figure (2)(b), even though the MAD values of the two layers are not the same, they have a similar tendency in the presence of abrupt changes. This leads to the idea that some encoding results of the base layer can be used to inform the coding of the enhancement layer(s), thus benefitting from the bottom-up coding structure of the standard. Therefore, a new MAD prediction model for the spatial enhancement layer using the encoding results from the base layer as a factor in the MAD prediction procedure is proposed. The new prediction model is defined as:

$$\mathrm{E\_MAD_j} = \mathrm{a_1} \times \mathrm{E\_MAD_{j-1}} + \mathrm{a_2} + \mathrm{a_3} \times \Gamma_\mathrm{j} \tag{6}$$

As for equation (5), $\mathrm{a_1}$ and $\mathrm{a_2}$ are model coefficients updated after the coding of each frame. The prefix E_ indicates the enhancement layer. Note that the model is a general model and can be used at the basic coding unit (e.g. macroblock) level or at frame level. Therefore, $\mathrm{E\_MAD_j}$ may refer to the predicted MAD value of the $\mathrm{j^{th}}$ frame or that of one basic unit in the $\mathrm{j^{th}}$ frame. Similarly, $\mathrm{E\_MAD_{j-1}}$ may refer to the actual MAD value of the previous frame or that of one basic unit in the same position of the previous frame. Γ_j refers to the difference between the actual and predicted MAD of the base layer of the $\mathrm{j^{th}}$ frame in the base layer, and is defined as

$$\Gamma_\mathrm{j} = \mathrm{B\_MAD_{actual,j}} - \mathrm{B\_MAD_{predicted,j}} \tag{7}$$

Table 2 Comparison of rate control accuracy.

Sequence	Target bitrate (kbps)	FixedQp		JVT-G012		Proposed	
		BR (kbps)	Mism. (%)	BR (kbps)	Mism. (%)	BR (kbps)	Mism. (%)
Bus	384	391.5	1.94	385.0	0.26	384.4	0.10
	512	522.1	1.96	513.3	0.25	512.0	0.00
	768	745.6	2.91	769.1	0.14	768.6	0.08
	1280	1303.9	1.87	1282.1	0.16	1280.6	0.05
Football	768	743.2	3.23	768.5	0.07	768.0	0.00
	1024	1042.3	1.79	1024.5	0.05	1024.6	0.06
	1536	1519.7	1.06	1538.3	0.15	1535.5	0.03
	2560	2513.3	1.82	2560.1	0.00	2559.2	0.03
Foreman	192	191.8	0.12	192.9	0.47	192.3	0.16
	256	257.1	0.42	256.7	0.27	256.3	0.12
	384	389.2	1.35	385.0	0.26	384.3	0.08
	640	637.7	0.36	641.3	0.20	639.8	0.03
Mobile	256	251.3	1.83	256.1	0.04	256.6	0.23
	384	370.3	3.57	384.7	0.18	384.4	0.10
	512	522.0	1.96	512.8	0.16	512.1	0.02
	768	777.4	1.22	769.0	0.13	767.6	0.05
Average			**1.71**		**0.17**		**0.07**

where the prefix B_ indicates the base layer. a_3 is the Γ_j weighting factor and typically is assigned a value of 0.1. The value of a_3 is determined from consideration of a very large number of training samples and many different types of picture content. Consequently it is reliable and widely-applicable. It may be possible to define an adaptive threshold that is even more accurate, and this may be pursued in the future.

With the above model, the prediction errors from the base layer are used to assist estimation of the MAD in the enhancement layers. When encoding the enhancement layers, the encoder is made aware of the abrupt changes of MAD in advance and promptly adjusts the MAD prediction to reduce the prediction errors. In this way, the bits are allocated more appropriately and not only is there an improvement in the rate control accuracy, but also an increase in the quality of the reconstructed video.

4 Experimental Results

The proposed algorithm was incorporated in the SVC reference software JSVM9.19.14 [1]. In order to validate the effectiveness of the proposed algorithm, video sequences with different degrees of motion activity and picture detail were coded, and the results compared with the JVT-G012 algorithm. The first 150 frames of each video sequence were coded to generate a reliable result. Two spatial layers are evaluated and the proposed algorithm is applied to the enhancement layer. Adaptive inter-layer prediction is enabled for the enhancement layer. As our algorithm attempts to optimise the R-Q model and MAD prediction model, which are only involved in P frames, the GOP structure is set to IPPP. In this work, a macroblock is chosen as the basic unit for rate control. To compare the results with the JVT-G012 algorithm, the initial Qp value is set to 32 for both schemes. Other parameters are set to the default values of the reference software.

The bit-rate mismatch (%Mism.) and rate distortion performance in terms of , BDBR (%), BDPSNR (dB), ΔBR (%), and ΔPSNR (dB) [2] were measured against the JVT-G012 scheme to evaluate the coding performance of the proposed rate control algorithm. ΔPSNR(dB) is computed according to $\Delta PSNR = PSNR_{Proposed} - PSNR_{G012}$, where $PSNR_{Proposed}$ and $PSNR_{G012}$ denote the PSNR resulting from the proposed algorithm and JVT-G012, and

Table 3 Comparison of rate distortion performance.

Sequence	JVT-G012		Proposed		ΔBR	ΔPSNR	BDBR	BDPSNR
	BR (kbps)	PSNR (dB)	BR (kbps)	PSNR (dB)	(%)	(dB)	(%)	(dB)
Bus	385.0	28.00	384.4	28.13	-0.16	0.13	-3.55	0.17
	513.3	29.28	512.0	29.43	-0.25	0.15		
	769.1	31.10	768.6	31.28	-0.07	0.18		
	1282.1	33.67	1280.6	33.82	-0.12	0.15		
Football	768.5	32.94	768.0	33.11	-0.07	0.17	-4.66	0.26
	1024.5	34.29	1024.6	34.59	0.01	0.30		
	1538.3	36.53	1535.5	36.78	-0.18	0.25		
	2560.1	39.39	2559.2	39.66	-0.04	0.27		
Foreman	192.9	32.85	192.3	32.97	-0.31	0.12	-2.68	0.11
	256.7	34.06	256.3	34.17	-0.16	0.11		
	385.0	35.69	384.3	35.78	-0.18	0.09		
	641.3	37.64	639.8	37.77	-0.23	0.13		
Mobile	256.1	23.98	256.6	24.07	0.20	0.09	-4.01	0.16
	384.7	25.61	384.4	25.71	-0.08	0.10		
	512.8	26.64	512.1	26.83	-0.14	0.19		
	769.0	28.08	767.6	28.38	-0.18	0.30		
Average					**-0.12**	**0.17**	**-3.73**	**0.18**

ΔBR(%) is computed as $\Delta BR = (BR_{Proposed} - BR_{G012})/BR_{G012} \times 100\%$, where $BR_{Proposed}$ and BR_{G012} denote the bit-rate resulting from the proposed algorithm and JVT-G012, respectively.

The rate control accuracy for four target bit rates is summarised in Table (2). All the test sequences and bitrates used in the experiments are those recommended by the JVT in document JVT-Q205 [14]. It can be seen that both the proposed algorithm and the JVT-G012 scheme work well at various target bit rates. Although both methods produce the target bit rates, the accuracy of the proposed algorithm is better than JVT-G012 in most cases. This is because the proposed optimised MAD prediction model results in a smaller prediction error when fast motion occurs. Most of the mismatch errors are less than 0.1% and the maximum error is 0.23%. The overall average absolute mismatch error is 0.07%. Consequently, it can be considered that bit rate is precisely controlled using the proposed algorithm.

The proposed rate control mechanism also achieves better rate-distortion performance for the enhancement layers than the JVT-G012 scheme. The comparative performance results are shown in Table (3). The results show that 1) given the same bit rate, the proposed algorithm increases the average PSNR by up to 0.26dB, and 2) given the same video quality (PSNR), the proposed algorithm produces a saving in average bit rate of up to 4.66%, compared to the JVT-G012 algorithm. In general, the PSNR of each of the four sequences is increased at all ranges of target bit rate. The maximum coding gain is 0.30dB. Therefore, the proposed algorithm improves the coding efficiency compared with the JVT-G012 rate control algorithm, and this is true regardless of target bit rate.

In order to test the robustness of the proposed algorithm, it was applied to video sequences of larger spatial resolution and with multiple enhancement layers. The experimental results show that the proposed algorithm consistently achieves a significant improvement in RD performance compared with that of JVT-G012. Due to limitations on space, the detailed results are presented in the longer version of this paper.

5 Conclusions

A rate control scheme for the spatial enhancement layer(s) in SVC has been described. The scheme introduces a separate rate-quantisation (R-Q) model for inter-layer prediction coding in the enhancement layer. An improved MAD prediction model is also proposed, where the MAD from previous temporal frames and previous spatial frames are considered together. In applying each of the above techniques, both the target bit rate mismatch is reduced and the coding efficiency is significantly improved. Simulation results show that the proposed method achieves better rate control accuracy than the JVT-G012 scheme, the average rate control mismatch error being 0.07%. Furthermore, the proposed algorithm attains higher coding efficiency than the JVT-G012 rate control algorithm. The improvement, averaged over the different types of video sequences coded, is an increase in PSNR of 0.18dB or a saving in bit rate of 3.73%.

References

1 JSVM (Joint Scalable Video Model) reference software for SVC. Online. Available: CVS server garcon.ient.rwth-aachen.de.

2 G. Bjøntegaard. Calculation of average PSNR differences between RD-curves. *VCEG-M33*, Apr. 2001.

3 T. Chiang and Y. Zhang. A new rate control scheme using quadratic rate distortion model. *IEEE Trans. Circuits Syst. Video Technol.*, 7(1):246–250, 1997.

4 S. Hu, H. Wang, S. Kwong, T. Zhao, and C.-C. J. Kuo. Rate control optimization for temporal-layer scalable video coding. *IEEE Trans. Circuits Syst. Video Technol.*, 21(8):1152–1162, 2011.

5 ITU-T Study Group 16. Video Codec Test Model, Near-Term, Version 8 (TMN8). *ITU-T/SG16/VCEG/Q15 A59*, Jun. 1997.

6 Z. G. Li, W. Gao, F. Pan, S. W. Ma, K. P. Lim, G. N. Feng, X. Lin, S. Rahardja, H. Q. Lu, and Y. Lu. Adaptive rate control for H.264. *J. Vis. Commun. Image Represent.*, 17(2):376–406, 2006.

7 J. Liu, Y. Cho, Z. Guo, and C.-C. J. Kuo. Bit allocation for spatial scalability coding of H.264/SVC with dependent rate-distortion analysis. *IEEE Trans. Circuits Syst. Video Technol.*, 20(7):967–981, 2010.

8 Y. Liu, Z. Li, and Y. Soh. Rate control of H.264/AVC scalable extension. *IEEE Trans. Circuits Syst. Video Technol.*, 18(1):116–121, 2008.

9 X. Lu and G. R. Martin. Rate control for scalable video coding with rate-distortion analysis of prediction modes. In *Proc. IEEE MMSP*, pages 289–294, 2013.

10 MPEG-2 Video Test Model Editing Committee. MPEG-2 Video Test Model 5 (TM5). *ISO/IEC JTC1/SC29/WG11 N0400*, Apr. 1993.

11 MPEG Video Group. MPEG-4 Video Verification Model Version 8 (VM8). *ISO/IEC JTC1/SC29/WG11 N1796*, Jul. 1997.

12 J. Ribas-Corbera and S. Lei. Rate control in DCT video coding for low-delay communications. *IEEE Trans. Circuits Syst. Video Technol.*, 9(1):172–185, 1999.

13 I. E. Richardson. *H.264 and MPEG-4 video compression: video coding for next-generation multimedia.* Wiley, 2003.

14 M. Wien and H. Schwarz. Testing conditions for SVC coding efficiency and JSVM performance evaluation. *JVT-Q205*, Jul. 2005.

15 J. Xie and L. Chia. Study on the distribution of DCT residues and its application to R-D analysis of video coding. *J. Vis. Commun. Image Represent.*, 19(7):411–425, Oct. 2008.

16 L. Xu, W. Gao, X. Ji, D. Zhao, and S. Ma. Rate control for spatial scalable coding in SVC. In *Proc. PCS*, 2007.