A FAST INTRA PREDICTION MODE SELECTION METHOD IN H.264/AVC SCALABLE VIDEO CODING

Sung-Jae Park, Yeo-Song Lee*, Chae-Bong Sohn*, S. Y. Jeong***, Kwang-sue Chung**, Ho-Chong Park*, Chang-Bum Ahn*, and Seoung-Jun Oh*,***

{*VIA-Multimedia Center, **BnC Convergence Platform Center} Kwangwoon University,
*** ETRI, Korea
E-mail: {sjpark,kunix, cbsohn}@media.kw.ac.kr, jsy@etri.re.kr,
{kchung, hcpark, chahn, sjoh}@kw.ac.kr

ABSTRACT

In this paper, we propose a fast intra prediction mode selection method in Scalable Video Coding(SVC) which is an emerging video coding standard as an extension of H.264/Advanced Video Coding(H.264/AVC). The proposed method decides a candidate intra prediction mode based on the characteristic of macroblock smoothness. Statistical analysis is applied to computing that smoothness in spatial enhancement layer. We also propose an early termination scheme for Intra_BL mode decision where the RD cost value of Intra_BL is utilized. Compared with JSVM software, our scheme can reduce about 55% of the computation complexity of intra prediction on average, while the performance degradation is negligible; For low QP values, the average PSNR loss is very negligible, equivalently the bit rate increases by 0.01%. For high QP values, the average PSNR loss is less than 0.01dB, which equals to 0.25% increase in bitrate on average.

Keywords: H.264, AVC, SVC, intra prediction mode

1. INTRODUCTION

Scalable Video Coding (SVC) has been developed as an extension of H.264/MPEG-4 Part10 Advanced Video Coding (H.264/AVC) by the joint video team (JVT) of ISO/IEC/ JTC 1/SC 29/WG 11(MPEG) and ITU-T SG16 WP3 Q.60 (VCEG)[1]. SVC is based on H.264/AVC and provides the scalability feature. It encodes a video in multiple layers with different temporal, spatial and quality scalability and decodes from partial streams depending on the specific rate and resolution required by a certain application [2]. Since its scalability capability is attained great interest, SVC can be applied in many application fields. Very high computational complexity prevents SVC from being applied in real time applications. Therefore, it becomes an interesting research work to design fast algorithms for core SVC modules which require very high computation.

High coding efficiency in H.264/AVC can be achieved by using rate distortion optimization (RDO) technique. RDO technique gets the best coding result in terms of maximizing coding quality and minimizing resulting data bits. The basic process of RDO is as following. The encoder encodes a macroblock (MB) using all the possible modes. The coding mode which produces the least rate distortion (RD) cost will be used in the final coding. However, as the RD cost is obtained only after all encoding process, it results in extremely high computational complexity in the encoder. Therefore, intra prediction is one of the most computational complexity modules in H.264/AVC. Furthermore, SVC uses the three intra modes that are applied to H.264/AVC and intra texture prediction is applied additionally.

There is a number of fast intra prediction algorithms proposed for H.264/AVC video coding [3-6] recently. However, these methods are fast scheme about most Intra4x4 mode. These methods are not suitable to spatial enhancement layer of SVC because are being based on H.264/AVC. Consequently, the intra prediction method in spatial enhancement layer of SVC should be researched for the fast methods. In this paper, we propose an efficient intra prediction mode decision in SVC. We proposed method decides a candidate intra prediction mode based on the characteristic of macroblock smoothness. Statistical analysis is applied to computing that smoothness in spatial enhancement layer. We also propose an early termination scheme for Intra_BL mode decision where the RD cost value of Intra_BL is utilized.

This paper is organized as follows. In Section 2, we will first analyze the intra prediction mode in spatial enhancement layer. Section 3 presents our intra prediction mode decision method through analysis of intra prediction mode in spatial enhancement layer. Then, Section 4 shows the simulation results. Finally, we conclude the paper in Section 5

2. STATISTICAL ANALYSIS OF INTRA PREDICTION MODE IN THE SPATIAL ENHANEMENT LAYER

Three modes such as Intra4x4, Intra8x8 and Intra16x16 are adopted in the intra coding of base layer in H.264/AVC. And in spatial enhancement layer, three intra modes and intra texture prediction are adopted in intra coding. Intra texture prediction using information from the previous spatial layer is provided with the Intra_BL macroblock mode, where the enhancement layer residue (the difference between the current macroblock and upsampled base layer) is transformed and quantized. As a key feature of SVC, Intra_BL improves the coding efficiency greatly [7].

However, the additional computations it introduces significantly increase the complexity of encoder at the same time.

The concept behind the method is that intra prediction mode is highly dependent on the characteristic of macroblock smoothness. Through a hypothesis testing [8], we analyze its intra prediction mode by the characteristic of macroblock smoothness in spatial enhancement layer. Hypothesis experiment results are classified by “Smoothness”, “No Decision”, and “No Smoothness”, according to the characteristic of macroblock smoothness. The hypothesis test is summarized as following steps:

Step 1: Calculate the mean and the variance of each subblock:

$$\mu_1, \mu_s, \cdots, \mu_s, \quad S_1^2, S_s^2, \cdots, S_s^2$$

where s is the number of subblocks. There are 16 8x2 and 16 2x8 subblocks in a macroblock.

Step 2: Compute the test statistics SC.

$$SC = MAX(\mu_1, \mu_s, \cdots, \mu_s) - MIN(\mu_1, \mu_s, \cdots, \mu_s)$$

Step 3: Calculate the critical values $k_{\alpha:s:s(p-1)}$

$$k = k^* \sqrt{SS^*_{wit} / p}$$

$$SS^*_{wit} = SS_{wit} / s(p-1)$$

$$SS_{wit} = p \sum_{i=1}^{s} S_i^2$$

k^* is obtained from the an ANOVA table.

Step 4: Decide the characteristic of macroblock smoothness

$$\begin{cases} SC < k_{\alpha_1:s:s(p-1)} : smothness \\ SC > k_{\alpha_2:s:s(p-1)} : no\, smothness \\ else : no\, decision \end{cases}$$

where $\alpha_1 = 0.05, \alpha_2 = 0.01$

Table 1. Statistical analysis of intra prediction modes in enhancement layer

Decision	Mode	QP 24	QP 28	QP 32	QP36
Smooth-ness	Intra BL	27.3%	29.3%	31.0%	32.5%
	Intra 4x4	13.0%	10.8%	8.6%	6.2%
	Intra16x16, 8x8	0.48%	0.67%	1.16%	2.08%
No Decision	Intra BL	25.3%	26.4%	27.6%	28.6%
	Intra 4x4	6.8%	5.7%	4.3%	2.8%
	Intra16x16, 8x8	0.08%	0.18%	0384%	0.82%
No smooth-ness	Intra BL	19.6%	20.6%	22.3%	23.5%
	Intra 4x4	7.4%	6.3%	4.4%	3.0%
	Intra16x16, 8x8	0.05%	0.13%	0.26%	0.58%

For the statistical analysis, we test 8 sequences such as Bus, Football, Foreman, Mobile, City, Crew, Harbour and Soccer with SVC reference software JSVM8.9 [9]. Hypothesis experiment results are shown in Table 1. From Table 1, we can find that the intra prediction mode is highly dependent on the characteristic of macroblock smoothness. While the percentage of both “No Decision” as well as “No Smoothness” in Intra8x8 togather with Intra16x16 is very small, that of “Smoothness” is relatively large. As QP value becomes larger, the number of both Intra16x16 and Intra8x8 modes becomes larger and that of Intra4x4 becomes smaller. Therefore, all modes are testes if the macroblock is decided as “Smoothness” while only both Intra_BL and Intra4x4 modes are tested if it is decised as either “No Decision” or “No Smoothness”.

3. A NEW INTRA PREDICTION MODE DECISION METHOD

3.1 Detection of no smoothness regions in enhancement layer

As shown in chapter 2, intra prediction mode in enhancement layer is highly dependent on the characteristic of macroblock smoothness. A complex block tends to be coded as either Intra4x4 or Intra_BL while a smooth block is coded as one among Intra16x16, Intra8x8, Intra4x4, or Intra_BL. To simplified the decision of smoothness, we categorize macroblock smoothness into two classes using the value of SCM(Statistical Character Mean) defined by Eq.(1): “Smoothness” and “No smoothness”.

$$SCM = MAX(m_1, m_2, \cdots, m_s) - MIN(m_1, m_2, \cdots, m_s) \qquad (1)$$

where m_i is the mean of the i–th subblock. If the texture is complex, SCM is large. Otherwise, that is small.

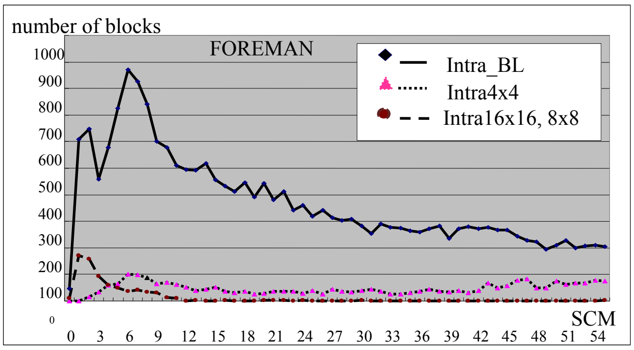

Fig. 1: The histogram of *SCM* for intra prediction mode in Foreman sequence

Fig. 1 shows *SCM* histogram for each intra prediction mode in Foreman sequence. We can insist as following; If a macroblock has a large value of SCM, then all intra modes are tested for this macroblock. Otherwise, two modes such as Intra4x4 and Intra_BL are tested.

In our decision method, the *SCM* value of the current macroblock is compared with a given threshold value T_k defined by Eq.(2).

$$T_h = c + \frac{QP}{6} \qquad (2)$$

where c = 9 determined empirically and QP is a quantization parameter in H.264/AVC. If $SCM > T_k$, we can say that this macroblock is complex and is coded as either

Intra4x4 or Intra_BL. Otherwise, it is considered as in a smoothness region and is coded as one among Intra16x16, Intra8x8, Intra4x4, and Intra_BL. The threshold value is linearly related to the quantization parameter of QP in H.264/AVC since block smoothness intrinsically depends on quantization step size.

3.2 Early decision of Intra_BL

According to [10], most macroblocks are coded as Intra_BL in enhancement layer. Therefore, we can reduce the computational complexity load of intra coding if we can decide as early as possible whether or not the macroblock is coded as Intra_BL in enhancement layer. For Intra_BL decision, the proposed method uses RD Cost of Intra_BL in the current macroblock.

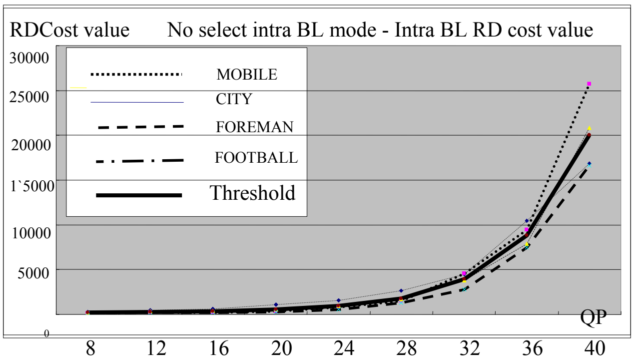

Fig. 2: The smallest RD Cost of unselected Intra_BL in City, Mobile, Football, and Foreman sequences

From Fig. 2, if the RD Cost of Intra_BL is smaller than this curved line in each sequence, the block is always coded as Intra_BL. However, those lines are not same, a representative line is chosen empirically after testing several well known sequences, which is shown as a thick solid line in Fig. 2. This solid line becomes a threshold for each QP. The value of TBL in Eq.(3) is the corresponding point in the line to each QP.

$$RDCost_{BL} < T_{BL} \tag{3}$$

This approximated threshold value may occur visual quality degradation. Thus, we need to provide a threshold update mechanism to reduce this effect shown in Eq. (4).

$$T_{BL\_new} = \begin{cases} T_{BL\_current} & \text{Best mode : Intra\_BL} \\ \min(T_{BL\_current},\ best\ RD\ \cos t) & \\ & \text{Best mode : No Intra\_BL} \end{cases} \tag{4}$$

where $T_{BL\_current}$ is threshold of current macroblock, *best RD cost* is the RD cost of the best mode among Intra16x16, Intra8x8, and Intra4x4. T_{BL_new} is the threshold of next macroblock. If $RDCost_{BL}$ satisfies Eq. (3), the best mode of macroblock is decided as Intra_BL. Otherwise, calculate *RD cost*'s for Intra16x16, Intra8x8, and Intra4x4 modes and decide the best mode.

3.3 Overall Algorithm Description

Based on the above observations and analysis, we propose an efficient intra prediction mode decision in spatial enhancement layer, where we decide the candidate intra modes using block smoothness in spatial enhancement layer. Time consumption can greatly be reduced by simply checking whether the early termination condition of Intra_BL is satisfied or not. The threshold value for the condition is decided using the RD Cost value shown at section 3.2. The entire procedure is outlined as follows:

Step 1: Fine the best prediction mode of INTRA_BL.
Step 2: Examine the inequality Eq.(3).
　　If $RDCost_{BL} < T_{BL}$ goto step 7;
　　　Otherwise goto step 3.
Step 3: Calculate *SCM* using Eq.(1).
　　If *SCM* < T goto step 4;
　　　Otherwise goto step 6.
Step 4: Fine the best prediction mode of the macroblock from Intra16x16 mode.
Step 5: Fine the best prediction mode of the macroblock from Intra8x8 mode.
Step 6: Fine the best prediction mode of the macroblock from Intra4x4 mode.
Step 7: Decide the best mode with the smallest cost.
Step 8: Update the threshold using Eq.(4).

4. EXPERIMENTAL RESULTS

We apply the proposed method to JSVM8.9. Two layers are adopted in our experiments. Each sequence is coded as a spatially scalable bit-stream with the base layer resolution equals to half of the enhancement layer resolution. The CIF/QCIF sequences we test are City, Foreman, Football, Soccer, and Crew at 15fps. All the sequences are coded with only I-pictures at fixed QP (8, 12, 16, 20, 24, 28, 32, 36). CAVLC is used as entropy coding.

Table 2 shows the simulation results. In the table, Full Search is the algorithm implemented in JSVM8.9 software, Fast Intra is our proposed method. *TS* represents the average time saving in encoding process, which is the same meaning of the complexity reduction of intra prediction. *TS* is computed as Eq.(5):

$$TS = \frac{T_{ref} - T_{proposed}}{T_{ref}} \times 100\ [\%] \tag{5}$$

where T_{ref} and $T_{proposed}$ are, respectively, the encoding times of Full Search software and our Fast Intra software. In the table, ENC_TS(ENCoding Time Saving) is the average time saving percentage for encoding the entire sequence and IP_TS(Intra Prediction Time Saving) is the average time saving percentage for intra prediction only.

From the simulation results, we can see that our proposed method has a good performance. Compared with Full Search, our scheme reduces about 55% of the computation complexity of intra prediction on average, while the performance degradation is negligible; For low QP values, the average PSNR loss is very negligible, equivalently the bit rate increases by 0.01%. For high QP values, the average PSNR loss is less than 0.01dB, which equals to 0.25% increase in bitrate on average. It shows that our proposed method is an effective intra coding approach for scalable video coding.

5. CONCLUSION

In this paper, we propose an efficient intra prediction mode decision for Scalable Video Coding. It reduces the candidate intra modes by removing Intra16x16 and Intra8x8 from intra coding in enhancement layer based on the characteristic of macroblock smoothness, proposing an early termination condition, together with updating the threshold value for minimizing visual quality degradation caused by early termination. The RD cost value of Intra_BL is utilized for Intra_BL mode decision.

Experimental results show that compared with full search adopted in SVC, our method can reduce the computational complexity of intra prediction by about 55% on average, while the performance degradation is so negligible.

ACKNOWLEDGEMENT

This work was supported by grant from the Ministry of Knowledge Economy and Korea Industrial Technology Foundation through the Human Resource Training Project for Strategic Technology in 2008, and Electronics and Telecommunications Research Institute, grant No.9122-2008-0392 in 2008.

REFERENCES

[1] ISO/IEC JTC 1/SC 29/WG 11: ‘Text of ISO/IEC 14496-10:2006/PDA M3 Scalable Video Coding’, MPEG 75/N7797, Thailand, January 2006.

[2] J.-R Ohm, “Advances in scalable video coding,” Proc. IEEE, vol. 93, n0. 1, pp. 42-56, Jan. 2005.

[3] B. Meng, C. Oscar, C. W. Wong, and H. Lan, “Efficient intra-prediction algorithm in H.264,” Proceedings of the 2003 IEEE International Conference on Image Processing, pp.837-840, Barcelona, Spain, Sep. 2003.

[4] B. Meng, C. Oscar, C. W. Wong, and H. Lan, “Efficient intra-prediction mode selection for 4x4 blocks in H.264,” Proceedings of the 2003 International Conference on Multimedia and Expo, pp.521-524, Baltimore, Maryland, July 2003.

[5] F. Pan, L. S. Rahardja, K. P. Lim, L. D. Wu, W. S. Wu, C. Zhu, W. Ye, Z. Liang, “Fast mode decision algorithm for H.264/AVC video coding,” in Proceedings of International Conference on Image Processing, vol. 2, pp. 781-784, Oct. 2004.

[6] C. Kim, H. Shin and C. J. Kuo, “Multistage mode decision for intra prediction in H.264 codec”, Proceedings of Visual Communications and Image Processing 2004, pp.355-363, California, USA, 2004.

[7] P. Yin, H. Boyce and P. Pandit, “spatial intra prediction on enhancement layer residue”, JVT-O053, Busan, Korea, April 2005

[8] W. W. Hines, C. Montgomery, D. M. Goldsman and C. M. Borror, “Probability and Statistics in Engineering,”, Wiley 4th Edition, 2003.

[9] ISO/IEC JTC 1/SC 29/WG 11/N8457: ‘JSVM8.9 Software’, Hangzhou, China, October 2006.

[10] L. Yang, Y. Chen, J. Zhai, F. Zhang “Low Complexity Intra Prediction for enhancement layer”, JVT-Q084, Nice, France, Oct. 2005.

Table 2. Performance of the proposed method in the enhancement layer

Sequence	QP	Full Search		Fast Intra		ENC_TS [%]	IP_TS [%]
		PSNR (dB)	Bitrate (kbit/s)	PSNR (dB)	Bitrate (kbit/s)		
Foreman	8	53.2	7445.7	53.2	7445.1	28.4	49.3
	12	49.8	5467.5	49.8	5467.6	29.1	51.5
	16	46.5	3951.4	46.5	3952.3	28.5	51.6
	20	43.3	2685.3	43.3	2686.8	28.1	51.7
	24	40.5	1779.8	40.5	1781.5	27.5	52.1
	28	37.9	1163.9	37.9	1166.2	28.0	53.5
	32	35.3	741.8	35.3	745.4	27.0	56.5
	36	33.0	478.3	32.9	484.1	31.0	60.6
City	8	53.3	9077.2	53.3	9079.7	29.6	55.2
	12	49.6	7067.9	49.6	7067.9	31.0	54.8
	16	46.1	5400.8	46.1	5401.0	31.0	55.0
	20	42.4	3896.1	42.4	3896.5	30.0	54.4
	24	39.2	2743.1	39.2	2743.3	29.0	54.1
	28	36.1	1843.0	36.1	1844.1	29.3	54.5
	32	33.0	1136.2	33.0	1138.3	28.6	54.4
	36	30.4	671.6	30.4	674.5	28.8	55.5
Crew	8	53.0	7591.4	53.0	7592.8	27.2	49.0
	12	49.7	5596.7	49.7	5597.2	27.4	48.6
	16	46.7	4013.2	46.6	4013.6	26.8	48.8
	20	43.5	2715.5	43.5	2716.3	26.1	48.6
	24	40.8	1817.6	40.8	1818.8	25.9	48.7
	28	38.1	1177.0	38.1	1178.8	26.5	50.3
	32	35.5	732.0	35.4	735.2	26.4	52.4
	36	33.1	451.0	33.1	456.0	28.5	54.9
Harbour	8	53.1	10336.0	53.1	10337.2	31.3	57.7
	12	49.5	8289.7	49.5	8289.9	32.3	57.5
	16	46.0	6537.4	46.0	6537.4	32.1	57.7
	20	42.3	4927.0	42.3	4927.5	31.4	57.2
	24	38.9	3641.2	38.9	3641.3	31.8	57.3
	28	35.7	2597.2	35.7	2597.8	31.5	57.4
	32	32.4	1746.4	32.4	1748.2	31.0	57.6
	36	29.6	1133.3	29.6	1137.2	30.7	58.1
Mobile	8	53.1	13447.2	53.1	13447.4	31.1	56.7
	12	49.6	11162.8	49.6	11163.1	31.4	56.8
	16	46.1	9146.3	46.1	9146.5	31.5	59.1
	20	42.3	7218.0	42.3	7218.4	31.5	57.1
	24	38.7	5556.7	38.7	5557.0	31.4	57.4
	28	35.2	4127.4	35.2	4127.7	31.4	57.6
	32	31.6	2903.5	31.6	2903.9	31.3	57.7
	36	28.4	1954.9	28.4	1955.7	31.1	58.2

Table 3. Average PSNR difference and its corresponding bitrate increase as well as time saving

Sequence	QP (8,12,16,20)			QP (24,28,32,36)		
	Avg. PSNR Diff (dB)	Avg. Δ Bitrate (%)	Avg. IP_TS (%)	Avg. PSNR Diff (dB)	Avg. Δ Bitrate (%)	Avg. IP_TS (%)
Foreman	-0.004	0.02	51.0	-0.011	0.50	55.7
City	-0.003	0.01	54.8	-0.003	0.17	53.5
Crew	-0.001	0.02	48.8	-0.015	0.44	51.6
Harbour	-0.004	0.02	57.5	-0.008	0.12	57.6
Mobile	-0.001	0.00	57.4	-0.002	0.02	57.7
Average	-0.002	0.01	53.9	-0.008	0.25	55.2