Proceedings of 2009 APSIPA Annual Summit and Conference, Sapporo, Japan, October 4-7, 2009

An Efficient Mode Decision Scheme by using RD Cost Correlation Coefficients in Scalable Video Coding

Po-Chun Wang, Gwo-Long Li, Shih-Chien Lin and Mei-Juan Chen
Department of Electrical Engineering
National Dong-Hwa University, Taiwan
m9623035@em96.ndhu.edu.tw, m9323004@em93.ndhu.edu.tw, u9523040@ems.ndhu.edu.tw, cmj@email.ndhu.edu.tw

***Abstract*—Scalable video coding (SVC) achieves layered coding structure with adaptive inter-layer prediction which causes high computational complexity. To decrease the computational complexity of SVC, we present a fast algorithm to speed up the inter-mode decision process for SVC. The proposed algorithm early terminates inter-mode decision in the enhancement layer by estimating the rate-distortion (RD) cost from the macroblocks of the base layer and the enhancement layer in temporal, spatial and inter-layer directions. Moreover, a search range decision algorithm is also proposed in this paper to further increase the motion estimation speed by using the motion vector information from temporal, spatial or inter-layer domain. The simulation results show that the proposed algorithm can determine the best mode and provide up to 75% of total coding time saving with 2.2% of bit rate increasing and 0.08dB PSNR degradation in average compared to JSVM 9.10.**

I. INTRODUCTION

With the development of Internet technology, the topic of multimedia communication receives more and more attention. Currently, the advancement of software and hardware technologies brings the multimedia applications such as video telephony, digital television, Video on Demand (VOD), Internet Protocol Television (IPTV) and so on into our daily life. The development of personal mobile communication systems introduces the application heterogeneities in video coding. Therefore, the video coding system must code the video sequence in different frame sizes, frame rates and bit rates to supply such heterogeneous demands.

To meet the requirements of application heterogeneities, the newest video coding standard call Scalable Video Coding [1][2] or H.264 Scalable Extension was recently standardized by Joint Video Team of ITU-T Video Coding Group and ISO/IEC Moving Picture Experts Group. Compared with the previous video coding standards, SVC supports three scalabilities in terms of temporal, spatial and quality. In SVC, the video sources would be coded into one base layer and several enhancement layers. The scalable video coding structure called pyramid coding scheme is revealed in Fig. 1. In this figure, the base layer is responsible for coding the smallest size of video sequence and it is H.264/MPEG4 Part 10 (AVC) compatible. To improve the coding performance of SVC, it prefers to remove the redundancies between different frame resolution layers when encoding enhancement layers. Although SVC can improve coding performance, it consumes significant computational complexity when compared to the original H.264 due to the inclusion of three inter-layer prediction tools.

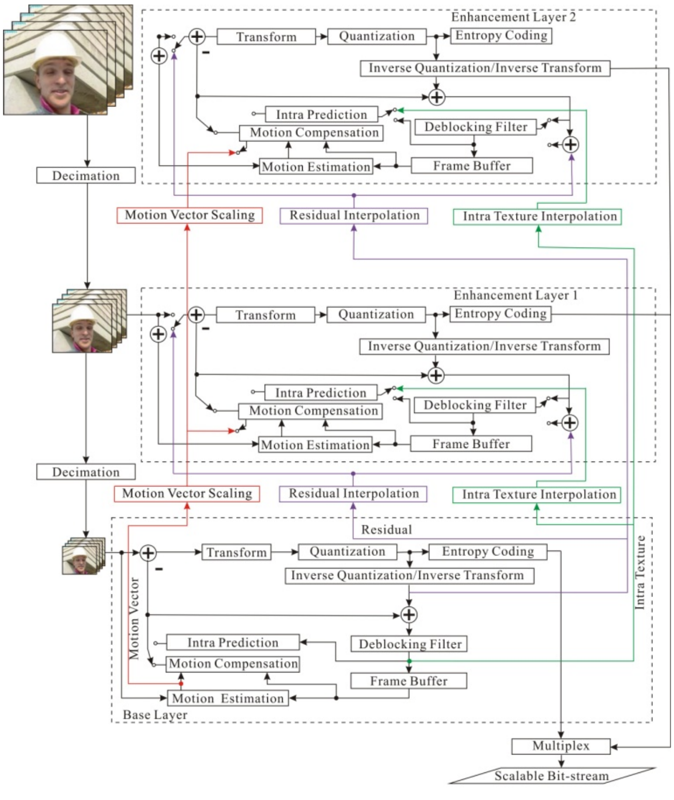

Fig. 1 Spatial scalability architecture of SVC

Several fast mode decision algorithms have been proposed for speeding up the SVC encoding process [3]-[12] . In these literatures, they attempted to decrease the motion estimation time in the enhancement layer for spatial scalability. In Xiong *et al*. [3] and Libo *et al*. [4], the authors reduced the intra prediction mode and prediction direction to decrease the coding time in the enhancement layer. Han *et al*. [5] proposed an intra prediction scheme for inter-layer intra prediction in the enhancement layer. Ye *et al*. [6] proposed a residual up-sampling scheme for the inter-layer residual prediction. Lange *et al*. [7][8] proposed an adaptive motion vectors selection method from base layer for the inter-layer motion prediction. Wang *et al*. [9] proposed the concept of different intra block sizes to improve the performance of inter-layer intra mode prediction. Li *et al*. [10] observed the intra mode relationship between the base layer and enhancement layer and the mode decision rule was introduced to reduce the computational complexity in the enhancement layer by adopting the observations. Lin *et al*. [11] analyzed the mode relationship between the base layer and enhancement layer for different quantization parameters and thus proposed a mode decision table to decide best mode in enhancement layer. In addition to predicting the best mode manners, the other way is to decide initial search point and reference frame from base layer. Lee *et al*. [12] used the bi-predictive zero motion, uni-predictive zero motion and zero coefficient block of base layer to decide what prediction modes that should be checked in enhancement layer. Ri *et al*.[13] used the correlation coefficient of rate distortion in spatial and temporal directions to decide the early termination threshold in mode decision.

To reduce the computational complexity of SVC, we propose a new fast inter-mode selection scheme by considering the rate-distortion cost correlation coefficients in the base layer and the enhancement layer to decide the macroblock mode in enhancement layer. In addition, according to the maximum correlation coefficient, an adaptive search range decision algorithm is also proposed to further increase the coding speed.

The rest of this paper is organized as follows. Section II gives an overview of inter-layer prediction in SVC. A new inter-mode decision scheme by using correlation coefficient is introduced in Section III. The experimental results and conclusion are presented in Section IV and Section V, respectively.

II. OVERVIEW OF INTER-LAYER PREDICTION IN SCALABLE VIDEO CODING

To receive better coding performance in H.264/AVC, seven different block sizes and shapes are supported for the inter mode prediction such as MODE_16x16, MODE_16x8, MODE_8x16, MODE_8x8, MODE_8x4, MODE_4x8 and MODE_4x4. Furthermore, for intra mode prediction, there are nine prediction directions for INTRA_4x4 and four prediction directions for INTRA_16x16. However, in addition to the inherent prediction modes supported in H.264/AVC, three more macroblock prediction modes called inter-layer motion prediction, inter-layer residual prediction and inter-layer intra prediction are additional supported to encode the macroblock of enhancement layers in SVC. In these inter-layer prediction modes, the base layer information is used as reference to further increase the coding performance. Since our proposed algorithm is designed to increase the mode decision speed in enhancement layers of SVC, it is necessary to understand these inter-layer prediction modes. The three inter-layer prediction modes adopted in SVC are described as follows.

A. Inter-layer motion prediction

In this prediction mode, when the enhancement layer as well as the base layer is inter prediction mode, the motion information of base layer can be used as reference for prediction in enhancement layer as shown in Fig. 2. In this manner, the enhancement layer macroblock partition is acquired from corresponding 8x8 block of the base layer associated with a scaling operation. For example, if the prediction mode of corresponding 8x8 block size is 4x8, the block size of 4x8 is scaled to 8x16 block size if the frame resolution ratio between the enhancement layer and the base layer is two. In addition to the block size, the motion vectors of the enhancement layer are obtained by multiplying the motion vectors of corresponding 8x8 block size in base layer by 2. Furthermore, the up-sampled motion information are used to refine the search results.

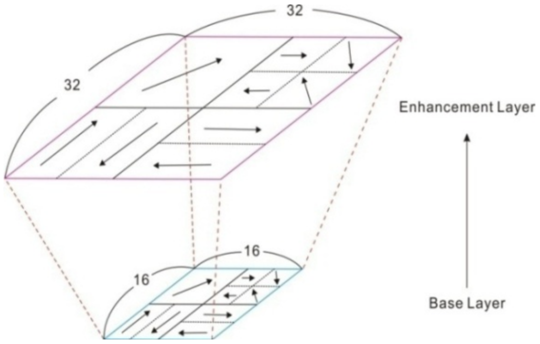

Fig. 2 Illustration of inter-layer motion prediction

B. Inter-layer residual prediction

Fig. 3 shows the concept of inter-layer residual perdition mode. When inter-layer residual prediction is performed, the residual data is up-sampled from corresponding 8x8 block of the base layer by bilinear interpolation. Afterward, the up-sampled residuals are used for predicting the current macroblock in the enhancement layer.

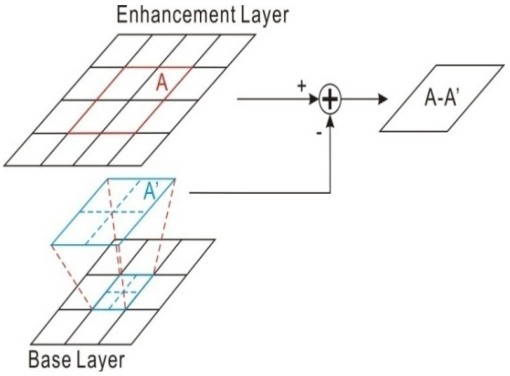

Fig. 3 Illustration of inter-layer residual prediction

C. *Inter-layer intra prediction*

Inter-layer intra prediction can be employed for macroblock in the enhancement layer if the corresponding block in the base layer is intra mode block. That is, the enhancement layer macroblock can be predicted by up-sampling the reconstructed macroblock of the base layer and the concept of inter-layer intra prediction is shown in Fig. 4. For up-sampling the reconstructed macroblock in the base layer, one-dimensional four-tape and bilinear filer are used for up-sampling the luminance and chrominance components, respectively.

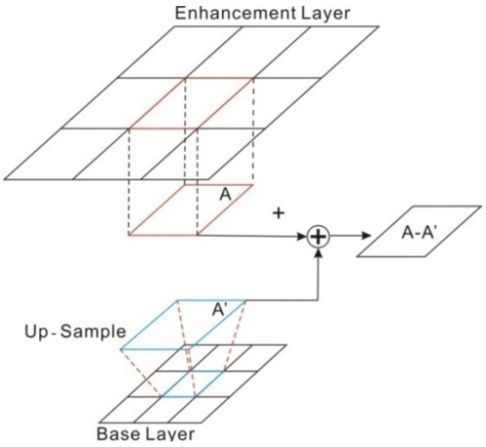

Fig. 4 Illustration of inter-layer intra prediction

III. PROPOSED FAST INTER-MODE DECISION SCHEME IN SVC

In this section, a fast inter-mode decision algorithm is proposed to enhance the coding performance of SVC encoder. The proposed algorithm includes fast mode decision based on the correlation coefficient of rate-distortion costs. To decide the best mode of the current macroblock in the enhancement layer, the proposed fast mode decision algorithm is described as follows.

A. *Rate-Distortion Cost Prediction*

In H.264, the best mode is decided by the rate-distortion cost and it can be expressed as Eq. (1).

$$J(s,c,Mode|QP) = SSD(s,c,Mode|QP) + \lambda_{Mode} \cdot R(s,c,Mode|QP) \quad (1)$$

where s is the original block, c refers to the reconstructed block, QP means the quantization parameter and λ_{Mode} denotes the Lagrangian multiplier. In Eq. (1), the SSD is the sum of the square difference that is used to measure the distortion. $Mode$ is the encoding mode and $R(s,c,Mode|QP)$ represents the number of bits associated with the *Mode* and motion vectors.

In SVC, different video resolutions are supported by spatial scalability to satisfy the diversities of user requirement. In the encoding procedure, the video sequence in the base layer is the downscaled version of the enhancement layer. This mechanism results in the high correlation between the base layer and the enhancement layer. Hence, we can use the correlation coefficient of rate-distortion costs to decide a best mode of a macroblock in the enhancement layer.

Before describing our proposed algorithm, the definition of correlation coefficient should be identified. The correlation coefficient is a well-known rule in statistics and probability theory. It is usually used to indicate the strength and direction between two variables X and Y as Eq. (2)

$$corr(X,Y) = \rho_{X,Y} \quad (2)$$

$$\rho_{X,Y} = \frac{cov(X,Y)}{\sigma_X \sigma_Y} \quad (3)$$

where *corr* is the correlation coefficient and *cov* refers to the covariance and the σ being the standard deviation. The correlation coefficient can be estimated by the sample Pearson product-moment correlation coefficient. As a result, the correlation coefficient can be computed as Eq. (4).

$$r_{XY} = \frac{\sum x_i y_i - (\sum x_i \sum y_i)/n}{\sqrt{\sum {x_i}^2 - (\sum x_i)^2/n}\sqrt{\sum {y_i}^2 - (\sum y_i)^2/n}} \quad (4)$$

where $x_i \in X$ and $y_i \in Y$. However, since the strong relationship existing in video data itself, the relationships in spatial, temporal, and inter-layer domains are taken into account for constructing our fast mode decision algorithm. In this paper, the rate-distortion costs are used as the samples for computing the correlation coefficients. Fig. 5 shows the corresponding macroblocks (samples) that have been used in our proposed algorithm for deriving the correlation coefficients. The definitions of various samples for spatial, temporal and inter-layer domains are shown in Eqs. (5), (6) and (7), respectively.

Spatial samples

$$\begin{aligned} X_S &= \{RDcost_B, RDcost_{B_{-1}}, RDcost_{B_{-2}}\} \\ Y_{S\alpha} &= \{RDcost_{B_\alpha}, RDcost_{B_{\alpha-1}}, RDcost_{B_{\alpha-2}} | \\ &\qquad \alpha \in \{L,U,UL,UR\}\} \end{aligned} \quad (5)$$

Temporal samples

$$\begin{aligned} X_T &= \{RDcost_B, RDcost_{B_L}, RDcost_{B_{UL}}, RDcost_{B_U}\} \\ Y_{T\beta} &= \left\{ \begin{matrix} RDcost_{B_{-\beta}}, RDcost_{B_{L-\beta}}, \\ RDcost_{B_{UL-\beta}}, RDcost_{B_{U-\beta}} \end{matrix} \right| \beta \in \{1,2\}\} \end{aligned} \quad (6)$$

Inter-layer samples

$$\begin{aligned} X_L &= \{RDcost_{B_L}, RDcost_{B_{UL}}, RDcost_{B_U}\} \\ Y_L &= \{RDcost_{E_L}, RDcost_{E_{UL}}, RDcost_{E_U}\} \end{aligned} \quad (7)$$

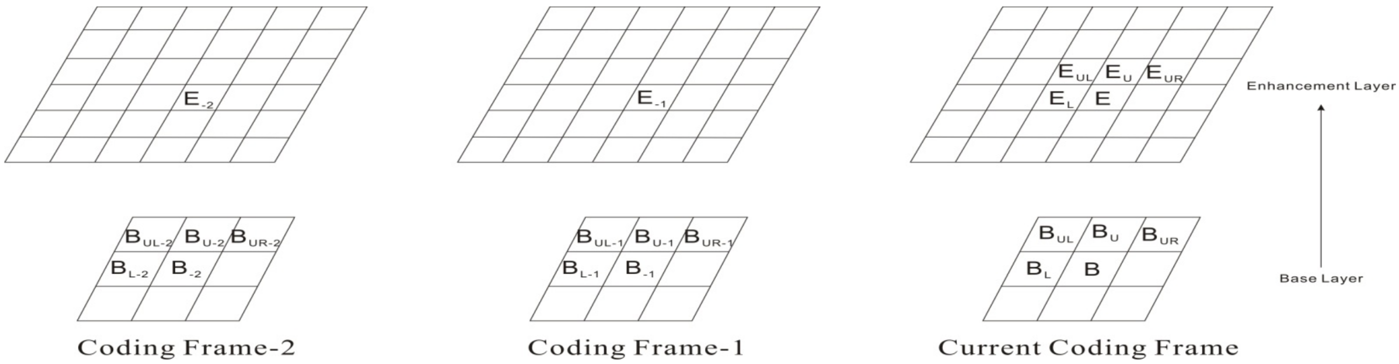

Fig. 5 Illustration of the used macroblocks in our preposed

where *RDcost* refers to the rate distortion cost of certain macroblock. The X_S, X_T and X_I are the samples in spatial, temporal and inter-layer domain, respectively. They will be used to substitute for the *X* component in Eq. (4) to calculate the correlation coefficients. Similarly, the $Y_{S\alpha}$, $Y_{T\beta}$, and Y_I are the samples which will be used to substitute for the *Y* component in Eq. (4) to obtain the correlation coefficients. Once the samples of spatial, temporal and inter-layer domains are completely defined, the correlation coefficients of these domains are calculated by Eq.(4) separately. Consequently, the maximum correlation coefficient is selected by Eqs. (8)-(11) since the most related macroblock should be extracted to be further used as prediction reference.

$$C_S = max\{\gamma_{X_S Y_{S\alpha}}\} \tag{8}$$

$$C_T = max\{\gamma_{X_T Y_{T\beta}}\} \tag{9}$$

$$C_L = \gamma_{X_L Y_L} \tag{10}$$

$$C_{Pred} = max\{C_S, C_T, C_L\} \tag{11}$$

By the assistance of Eqs.(8)-(11), the direction which has the maximum correlation coefficient can be obtained and consequently to be used as the prediction reference for the current encoding macroblock in the enhancement layer. Hence, the prediction reference derived from the previous step is further adopted to decide an early termination threshold (*TH*) to avoid exhaustive mode test. The *TH* is based on the maximum correlation coefficient, C_{Pred}, and it can be adaptively adjusted in encoding of the enhancement layer because of the high relationship between the base and the enhancement layers. As a result, the *TH* can be derived by the following rules.

$$TH = \begin{cases} RDcost_{E_{-1}} & : C_{Pred} = \gamma_{X_T Y_{T1}} \\ RDcost_{E_{-2}} & : C_{Pred} = \gamma_{X_T Y_{T2}} \\ RDcost_{E_L} & : C_{Pred} = \gamma_{X_S Y_{SL}} \\ RDcost_{E_U} & : C_{Pred} = \gamma_{X_S Y_{SU}} \\ RDcost_{E_{UL}} & : C_{Pred} = \gamma_{X_S Y_{SUL}} \\ RDcost_{E_{UR}} & : C_{Pred} = \gamma_{X_S Y_{SUR}} \\ RDcost_B & : C_{Pred} = \gamma_{X_L Y_L} \end{cases} \tag{12}$$

After obtaining the *TH*, it will be used during mode decision process to check whether the modes waiting to be examined should be further tested or not. If the rate distortion cost of the current testing mode is smaller than *TH*, the mode decision process will be terminated immediately.

B. Adaptive Search Range

As mentioned above that the motion estimation of SVC consumes lots of computational complexities. To further decrease the coding time, the exhaustive search point checking is highly expected to be avoided during motion search in the enhancement layer. To aim at this goal, the relationship between the motion vectors of the current encoding macroblock and the macroblock with maximum correlation coefficient is analyzed. Table I shows the probability that the motion vector magnitude of the current encoding macroblock is smaller than or equal to the motion vector magnitude of the macroblock with maximum correlation coefficient. From Table I, it can be observed that the motion vector magnitude between current encoding macroblock and the macroblock with maximum correlation coefficient are very similar. This situation shows us that the best motion vector of the current encoding macroblock in the enhancement layer can be obtained within a restricted search area which can be decided by the motion vectors of the macroblock with maximum correlation coefficient. As a result, the search range can be adaptively adjusted with the change of the macroblock which has maximum correlation coefficient. The search range for the current encoding macroblock can be calculated as follows.

$$SR = \begin{cases} MV_{E_{-1}} & : C_{Pred} = \gamma_{X_T Y_{T1}} \\ MV_{E_{-2}} & : C_{Pred} = \gamma_{X_T Y_{T2}} \\ MV_{E_L} & : C_{Pred} = \gamma_{X_S Y_{SL}} \\ MV_{E_U} & : C_{Pred} = \gamma_{X_S Y_{SU}} \\ MV_{E_{UL}} & : C_{Pred} = \gamma_{X_S Y_{SUL}} \\ MV_{E_{UR}} & : C_{Pred} = \gamma_{X_S Y_{SUR}} \\ MV_B \times m & : C_{Pred} = \gamma_{X_L Y_L} \end{cases} \tag{13}$$

where *MV* is the motion vector magnitude of the corresponding macroblock. The *m* refers to the ratio of the frame resolution between layers. If the maximum correlation coefficient belongs to inter-layer prediction, the search range should be multiplied by a factor of *m* to adjust search range size since the motion vectors of the base layer are upsampled for prediction of the enhancement layer in SVC.

TABLE I PROBABILITY THAT THE MOTION VECTOR MAGNITUDE OF THE CURRENT MACROBLOCK IS SMALLER THAN OR EQUAL TO THE MOTION VECTOR MAGNITUDE OF THE MACROBLOCK WITH MAXIMUM CORRELATION COEFFICIENT

Sequence	Probability
Akiyo	98.52 %
Foreman	95.22 %
Flower	96.73 %
News	96.48 %

C. Proposed Fast Inter-Mode Decision Algorithm

The flowchart of overall proposed algorithm is shown in Fig. 6 and the coding procedure is described as follows. First, the base layer is encoded by the H.264 compatible encoder. Afterward, for encoding the enhancement layers, we calculate seven correlation coefficients from spatial, temporal and inter-layer domains by Eqs. (4)-(7) and the maximum correlation coefficients are obtained by Eqs. (8)-(11). After obtaining the maximum correlation coefficient direction, the threshold, *TH*, based on the C_{Pred} is derived by Eq. (12) to define an early termination criterion for increasing the mode decision speed. Meanwhile, the search range of the current encoding macroblock is dynamically decided by C_{Pred} in Eq. (13). For an upcoming test mode, the position with the minimum rate distortion cost is found out within the decided search range first. If the minimum rate distortion cost of the current testing mode is smaller than *TH*, the test for the unchecked modes is terminated immediately and the best mode is selected from the previous checked modes. Otherwise, the other modes are checked in turns.

IV. EXPERIMENTAL RESULTS

The proposed fast mode decision algorithm is implemented in JSVM 9.10 encoder and the PSNR, bit rate and total encoding time are employed to measure the performance of the proposed algorithm. The computer used for simulation has 3.0GHz Intel Pentium 4 CPU, 512M bytes RAM with Windows XP professional operating system. The simulation tests ten sequences including *Akiyo, City, Coastguard, Crew, Foreman, Flower, Harbour, Ice, Mobile and News* for CIF format. The frame format of the base layer is QCIF and the enhancement layer is CIF so that the *m* is 2. The group of picture is set to 16 for temporal scalable at each layer. The maximum search range is set to +16/-16 pixels and the number of reference frame is set to 3 for motion estimation. The coding parameters are shown in Table II. The quantization parameters are set to 38 in the base layer. Each test sequence contains total of 100 frames, with one I frame followed by 99 B or P frames. For the performance comparison, three methods of Lin's [11], Lee's [12] and the original JSVM 9.10 reference software [14] are compared with our algorithm. We implemented Lin's algorithm without MODE_SR in JSVM 9.10 since the MODE_SR has been removed in reference software JSVM 9.10.

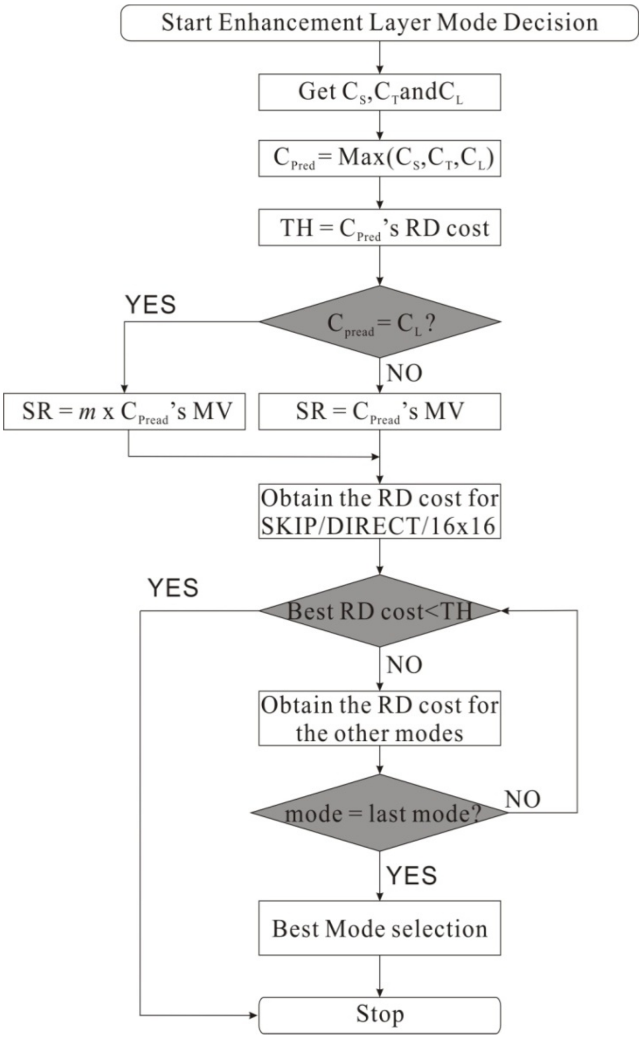

Fig. 6 Flowchart of the proposed scheme

Table III, Table IV and Table V show the comparison of PSNR, bitrate and the coding time saving (see Eq. (14)) between the proposed algorithm and the other fast mode decision methods with base layer QP=38 and enhancement layer QP=30. From these simulation results, the proposed algorithm provides up to 75% of total coding time saving. The maximum coding time saving is up to 77% for the *City* sequence. Furthermore, the average decrease of PSNR is only 0.08dB and the bitrate increase is 2.2%.

$$Time\ Save = \frac{T_{JSVM} - T_{Reference}}{T_{JSVM}} \times 100\% \qquad (14)$$

Table VI, Table VII and Table VIII show the performance comparisons with base layer QP=38 and enhancement layer QP=10. As shown in these tables, the proposed algorithm provides up 75% of total coding time saving in average. In addition, the PSNR degradation and bitrate increase is only 0.09dB and 0.7% when compared to JSVM9.10, respectively. Fig. 7 and Fig. 8 present the rate distortion (RD) curve for the *Akiyo* and *Foreman* sequence, respectively. In these figures, we can observe that our proposed algorithm does not deflate the coding efficiency and frame quality noticeably.

V. CONCLUSIONS

In this paper, a fast inter-mode decision algorithm is proposed to speedup the encoder process of SVC. The proposed fast inter-mode decision is based on the rate-distortion cost correlation coefficient of the base layer to determine the mode of macroblocks in the enhancement layer. In addition, according to the maximum correlation coefficient direction, the search range for the current encoding macroblock can be dynamically decided. Experimental results demonstrate that the proposed algorithm can provide up to 75% time saving with slight PSNR degradation and bit rate increase when compared to the JSVM 9.10.

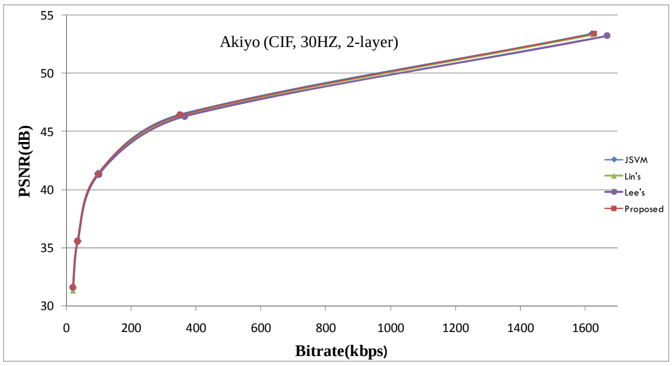

Fig. 7 RD curve for Akiyo sequence

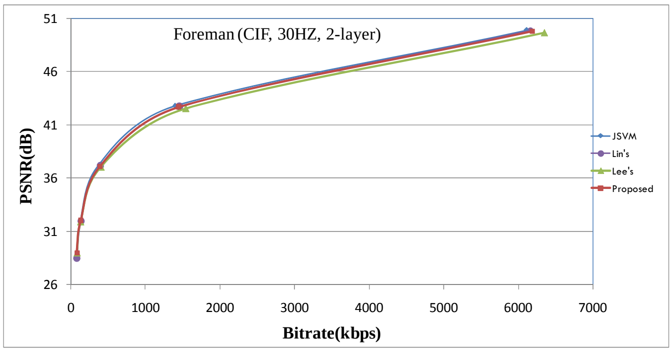

Fig. 8 RD curve for Foreman sequence

TABLE II CODING PARAMETERS

Parameter	value
GOP size	16
Search range	+16/-16
Number of reference frame	3
Intra period	-1
Motion vector accuracy	1/4
Frame rate	30HZ

TABLE III AVERAGE PSNR COMPARISON OF DIFFERENT METHODS
(BL QP = 38, EL QP = 30, 100 FRAMES)

Sequence	JSVM	Lin's [11]		Lee's [12]		Proposed	
	PSNR	PSNR	△PSNR	PSNR	△PSNR	PSNR	△PSNR
Akiyo	41.39	41.36	-0.03	41.31	-0.08	41.36	-0.03
City	36.50	36.42	-0.08	36.32	-0.18	36.41	-0.09
Coastguard	34.14	34.09	-0.05	34.04	-0.1	34.04	-0.1
Crew	36.86	36.79	-0.07	36.71	-0.15	36.77	-0.09
Foreman	37.27	37.17	-0.1	37.03	-0.24	37.13	-0.14
Flower	34.37	34.33	-0.04	34.22	-0.15	34.28	-0.09
Harbour	33.74	33.72	-0.02	33.68	-0.06	33.70	-0.04
Ice	39.85	39.69	-0.16	39.50	-0.35	39.69	-0.16
Mobile	33.64	33.60	-0.04	33.53	-0.11	33.58	-0.06
News	39.70	39.65	-0.05	39.56	-0.14	39.64	-0.06
Average	**36.74**	**36.68**	**-0.06**	**36.59**	**-0.15**	**36.66**	**-0.08**

TABLE IV AVERAGE BITRATE COMPARISON OF DIFFERENT METHODS
(BL QP = 38, EL QP = 30, 100 FRAMES)

Sequence	JSVM	Lin's [11]		Lee's [12]		Proposed	
	Bitrate	Bitrate	△Bitrate	Bitrate	△Bitrate	Bitrate	△Bitrate
Akiyo	99.73	99.27	0.08%	101.48	1.75%	97.79	0.06%
City	394.54	407.16	3.20%	426.32	8.05%	403.69	2.32%
Coastguard	923.76	930.79	0.76%	950.30	2.87%	933.30	1.03%
Crew	573.81	603.96	5.25%	604.25	5.30%	590.11	2.84%
Foreman	378.01	396.66	4.93%	404.08	6.90%	388.29	2.72%
Flower	1088.97	1116.55	2.53%	1168.67	7.32%	1128.00	3.58%
Harbour	1039.9	1055.49	1.50%	1073.92	3.27%	1053.06	1.27%
Ice	364.36	386.36	6.04%	391.76	7.52%	380.85	4.53%
Mobile	1053.34	1081.52	2.68%	1111.25	5.50%	1071.29	1.7%
News	225.37	232.77	3.28%	236.23	4.82%	229.18	1.69%
Average	**613.93**	**631.05**	**2.79%**	**646.82**	**5.36%**	**627.55**	**2.22%**

TABLE V AVERAGE TIME-SAVING COMPARISON OF DIFFERENT METHODS
(BL QP = 38, EL QP = 30, 100 FRAMES)

Sequence	Lin's [11]	Lee's [12]	Proposed
Akiyo	62.66	74.41	77.01
City	39.41	71.53	77.25
Coastguard	43.9	70.63	74.93
Crew	52.62	66.07	76.31
Foreman	47.03	68.94	75.05
Flower	47.58	68.61	73.81
Harbour	42.16	63.92	72.55
Ice	54.18	67.34	76.36
Mobile	38.53	69.52	76.65
News	61.32	70.47	75.32
Average	**48.93**	**69.14**	**75.52**

TABLE VI AVERAGE PSNR COMPARISON OF DIFFERENT METHODS
(BL QP = 38, EL QP = 10, 100 FRAMES)

Sequence	JSVM	Lin's [11]		Lee's [12]		Proposed	
	PSNR	PSNR	△PSNR	PSNR	△PSNR	PSNR	△PSNR
Akiyo	53.42	53.36	-0.06	53.25	-0.17	53.41	-0.01
City	49.70	49.67	-0.03	49.57	-0.13	49.67	-0.03
Coastguard	49.79	49.75	-0.04	49.46	-0.33	49.56	-0.23
Crew	50.13	50.12	-0.01	49.77	-0.36	49.87	-0.26
Foreman	49.90	49.86	-0.04	49.67	-0.23	49.80	-0.1
Flower	50.43	50.36	-0.07	50.14	-0.29	50.35	-0.08
Harbour	49.52	49.50	-0.02	49.38	-0.14	49.48	-0.04
Ice	51.44	51.38	-0.06	51.01	-0.43	51.33	-0.11
Mobile	49.57	49.53	-0.04	49.42	-0.15	49.55	-0.02
News	52.14	52.08	-0.06	51.75	-0.39	52.08	-0.06
Average	**50.60**	**50.56**	**-0.04**	**50.34**	**-0.26**	**50.51**	**-0.09**

TABLE VII AVERAGE BITRATE COMPARISON OF DIFFERENT METHODS (BL QP = 38, EL QP = 10, 100 FRAMES)

Sequence	JSVM	Lin's [11]		Lee's [12]		Proposed	
	Bitrate	Bitrate	△Bitrate	Bitrate	△Bitrate	Bitrate	△Bitrate
Akiyo	1621.91	1623.83	0.12%	1668.23	2.86%	1629.77	0.48%
City	6282.03	6285.58	0.06%	6425.51	2.28%	6343.41	0.98%
Coastguard	8745.84	8729.31	0.15%	8836.71	1.04%	8769.10	0.27%
Crew	7271.16	7294.22	0.32%	7410.02	1.91%	7293.20	0.3%
Foreman	6117.02	6160.34	0.71%	6348.74	3.79%	6185.70	1.12%
Flower	8027.64	8085.48	0.72%	8295.19	3.33%	8130.96	1.29%
Harbour	9698.65	9677.03	0.24%	9766.50	0.7%	9726.75	0.29%
Ice	2807.06	2885.45	2.79%	3035.38	8.13%	2883.30	2.72%
Mobile	10378.32	10417.8	0.38%	10617.5	2.3%	10444.2	0.63%
News	2768.17	2781.34	0.48%	2868.47	3.62%	2779.46	0.41%
Average	**6371.78**	**6394.03**	**0.35%**	**6527.22**	**2.44%**	**6418.59**	**0.73%**

TABLE VIII AVERAGE TIME-SAVING COMPARISON OF DIFFERENT METHODS (BL QP = 38, EL QP = 10, 100 FRAMES)

Sequence	Lin's [11]	Lee's [12]	Proposed
Akiyo	63.53	75.49	77.04
City	39.23	71.63	74.95
Coastguard	43.79	70.82	75.78
Crew	52.59	66.36	76.09
Foreman	47.10	69.52	73.95
Flower	48.07	69.02	75.80
Harbour	42.00	63.93	76.52
Ice	55.48	69.22	75.36
Mobile	38.40	69.61	76.13
News	62.01	73.55	75.12
Average	**49.08**	**69.88**	**75.67**

REFERENCES

[1] T. Wiegand, G. Sullivan, J. Reichel, H. Schwarz and M. Wien "Joint Draft ITU-T Rec. H.264 | ISO/IEC 14496-10 / Amd.3 Scalable video coding" *Joint Video Team (JVT) JVT-X201*, 2007.

[2] H. Schwarz, D. Marpe and T. Wiegand, "Overview of scalable video coding extension of H.264/AVC," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 17, No. 9, pp. 1103-1120, September 2007.

[3] L. Xiong, "Reducing enhancement layer directional intra prediction modes," *Joint Video Team (JVT) JVT-P041*, 2005.

[4] Y. Libo, "Low complexity intra prediction for enhancement layer," *Joint Video Team (JVT) JVT-Q084*, 2005.

[5] W. J. Han "Modified IntraBL design using smoothed reference," *Joint Video Team (JVT) JVT-R091*, 2006.

[6] Y. Ye and Y. Bao, "CE2: improved residual upsampling for ESS," *Joint Video Team (JVT) JVT-W117*, 2007.

[7] L. Lange, "Extended inter-layer motion vectors prediction in scalable video coding – case study and improvement proposal," *Joint Video Team (JVT) JVT-R094*, 2006.

[8] L. Lange, "Extended inter-layer motion vectors prediction in JSVM – case study and experimental results," *Joint Video Team (JVT) JVT-S063*, 2006.

[9] T. S. Wang, C. S. Park, J. H. Kim, M. S. Yoon and S. J. Ko, "Improved inter-layer intra prediction for scalable video coding," *in Proceedings of IEEE TENCON*, pp.1-4, 2007.

[10] H. Li, Z. G. Li, C. Y. Wen and L. P. Chau, "Fast mode diction for spatial scalable video coding," *in Proceedings of IEEE International Symposium on Circuits and Systems*, pp. 3305-3308, May 2006.

[11] H. C. Lin, W. H. Peng, H. M. Hang and W. J. Ho, "Layer-adaptive mode decision and motion search for scalable video coding with combined coarse granular scalability (CGS) and temporal scalability," *in proceedings of IEEE International Conference on Image Processing*, Vol. 2, pp. 289-929, October, 2007.

[12] B. S. Lee, M. C. Kim, S. J. Hahm, C. S. Park and K. S. Park, "A fast mode selection scheme in inter-layer prediction of H.264 scalable extension coding," *in proceedings of IEEE International Symposium on Broadband Multimedia Systems and Broadcasting*, pp. 1-5, March, 2008.

[13] S. H. Ri, Y. Vatis and J. Ostermann, "Fast Inter-Mode Decision in an H.264/AVC Encoder Using Mode and Lagrangian Cost Correlation," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 19, No. 2, pp. 302-306, February 2009.

[14] J. Reichel, H. Sclwarz and M. Wien, "Joint Scalable Video Model JSVM-9," *Joint video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG (ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6), JVT-V203,2007.*