Fast Mode Decision Algorithm for Scalable Video Coding Based on Probabilistic Models

Yu-Che Wey[1], Mei-Juan Chen[1], Chia-Hung Yeh[2] and Chia-Yen Chen[3*]
[1]Dept. of Electrical Engineering, National Dong Hwa University, Hualien, Taiwan
E-mail: 610023037@ems.ndhu.edu.tw, cmj@mail.ndhu.edu.tw
[2]Dept. of Electrical Engineering, National Sun Yat-sen University, Kaohsiung, Taiwan
E-mail: yeh@mail.ee.nsysu.edu.tw
[3,*]Dept. of Computer Science and Information Engineering, National University of Kaohsiung, Kaohsiung, Taiwan
E-mail: ayen@nuk.edu.tw

***Abstract*—To reduce the computational complexity of the encoding process in Scalable Video Coding, we utilize the information of motion vector predictor (MVP) and the number of non-zero coefficients (NZC) to propose a fast mode decision algorithm. The probability models of MVP and the number of NZC are built to predict the partition mode in the enhancement layer. In addition, the search range of motion estimation is adaptively adjusted to further reduce computational complexity. Experiment results show that the proposed algorithm can reduce coding time by up to 76% in average and provide higher time saving and better performance than previous work.**

1. INTRODUCTION

Scalable Video Coding (SVC) [1,2] is an extended version of H.264/AVC, providing multi-layer encoding for various devices. SVC supports spatial, temporal and quality scalabilities which provide various picture sizes, frame rates and qualities for different layers, respectively. Due to the high correlation between base layer (BL) and enhancement layer (EL), SVC improves the rate-distortion efficiency of the ELs using three kinds of inter-layer predictions: inter-layer motion prediction, inter-layer residual prediction and inter-layer intra prediction.

In H.264/SVC, there are eight types of inter prediction modes following H.264/AVC for macroblock (MB) modes; these modes are Mode SKIP, Mode 16×16, Mode 16×8, Mode 8×16, Mode 8×8, Mode 8×4, Mode 4×8 and Mode 4×4. In addition, there are two kinds of intra prediction modes, INTRA 16×16 and INTRA 4×4. To select the best mode for the current MB, the rate-distortion cost (R-D cost) of each mode should be calculated, and the mode that has the minimum R-D cost is chosen as the best mode. Many fast mode decision algorithms have been proposed to reduce the computational complexity of mode decision. Since Mode SKIP has the lowest complexity in all inter modes, [3] uses the coding information of the co-located MB in BL and the neighboring MBs of the EL to predict the SKIP mode. Wang et al.'s method [4] is based on the priority-based mode decision (PBMD) and uses the R-D cost correlation between BL and EL to decide the mode priority in EL. In [5], the algorithm employs the correlation between BL and EL to predict the best mode of the EL and uses motion vector difference (MVD) of the BL to decide the search range of the EL to accelerate the encoding time. Kuo and Chan [6] propose a method based on the motion field distribution to efficiently determine the block mode for complexity reduction by likelihood.

The search range has a significant impact on the coding time of motion estimation. Therefore, the definition of search range is also an important research issue in video coding. In [7], the authors propose an efficient motion re-estimation scheme for H.264 B-frame and P-frame transcoding by utilizing the maximum likelihood to evaluate motion vector candidates and predict the best candidate. Lee et al. [8] propose a fast motion estimation scheme by applying the adaptive search range (SR). This method is based on the heuristic rule observed from experiments of motion vector characteristics. The results show that it can reduce significant coding time while retaining good coding performance. Kim et al. [9] propose an efficient learning method to control the EL's search range according to the block modes of BL.

Fast mode decision and motion estimation algorithms for SVC have been popular issues for a long time. Compared to H.264 / AVC, the coding complexity of SVC is relatively much higher due to its ability to fulfill a wider variety of user requirements. Many previous works use the inter-layer correlation to reduce the computational complexity. Most previous methods utilize just one parameter to determine the early termination of the mode decision; as such, the accuracy can be further improved. In this paper, in addition to utilizing the inter-layer correlation to decide the best mode at an early stage, we propose an algorithm which utilizes the probability models generated by motion vector predictor (MVP) and the number of non-zero coefficients (NZC) of Mode 16×16 in EL to find the most probable modes. Furthermore, the proposed method also shrinks the search range. Overall, the proposed method reduces the number of modes required to be tested and finds the suitable mode rapidly so that the encoding process of SVC can be accelerated significantly.

2. Observation and Analysis

From experimental observations, we find that there are high correlations between the best mode with MVP and NZC in EL in the SVC. The observation motivated us to use various statistical analyses to find the presented probability model. In this section, we first observe the MVP and the number of NZC of Mode 16×16 in EL and record MVP and NZC from eight CIF sequences. We then analyze the features using Laplacian distributions. Finally, we combine this method with mode priority decision to make an early decision on the best mode.

2.1 Analysis of Motion Vector Predictor

The starting search point for motion estimation in the current MB is determined by the median of motion vectors from the neighboring left, top and top-right MB. We use the MVP to represent the motion characteristics of the current block. Usually, static blocks with less motion use larger size modes, like Mode 16×16. In contrast, high-motion blocks use smaller size modes, such as Mode 8×8. The MVP of Mode 16×16 in EL is analyzed to discover the correlation between the MVP of Mode 16×16 and the best mode of the MB. Eight sets of sequences with different complexities are used to analyze the MVP. The sequences are Akiyo, Foreman, Football, Soccer, Stefan, Coastguard, Table and News. These sequences contain differing amounts of motion. The base layer resolution is QCIF/CIF, with QP of 38. The enhancement layer resolution is CIF/4CIF, with QP of 32. The search range is ±32, applied on 3 reference frames, the group of pictures size is 16, and the total number of frames is 150. Fig. 1 shows the probability of the horizontal component of the MVP in each final selected mode, where the probabilities in each subfigure sum up to 1. The subscripts Mode_1, Mode_2, Mode_3 and Mode_4 represent Mode 16×16, Mode 16×8, Mode 8×16 and Mode 8×8, respectively. The peak values of the probabilities in Figs. 1(a) to (d) are 0.3913, 0.3581, 0.3901, and 0.3471, respectively. According to Fig. 1, when the associated mode in the EL is Mode 16×16, the probability distribution of the MVP tends to be zero or close to zero. When further subdivided into model with smaller blocks, the probability distribution model is more dispersed, as shown in Fig. 1(d) when the best mode is Mode 8×8.

2.2 Analysis of number of non-zero coefficients

The number of non-zero quantized coefficients can represent the complexity of block. From this viewpoint, we also analyze the correlation between the best mode and the number of NZC of Mode 16×16 in EL. The same sequences used in Section 2.1 are used in this section as well. The base layer resolution is QCIF/CIF, with QP of 28, 32, 36 and 40. The enhancement layer resolution is CIF/4CIF, with QP of 28, 32, 36 and 40. The search range is ±32, applied on 3 reference frames, the group of pictures size is 16, and the total number of frames is 150. Fig. 2 shows the probabilities of the number of NZC in each mode, where the probabilities in each subfigure sum up to 1, and the subscripts Mode_1, Mode_2, Mode_3 and Mode_4 represent Mode 16×16, Mode 16×8, Mode 8×16 and Mode 8×8, respectively. As seen from Fig. 2(a), the numbers of NZC in Mode 16×16 are less than other modes. The QP_{BL}/QP_{EL} values used in Fig. 2 are 28/28, and the observations are similar for all four QP pairs, namely 28/28, 32/32, 36/36, and 40/40. When the selected best mode has a smaller size, its number of NZC will have higher probabilities of being larger values, as seen in Fig. 2 (b) and (c) where the selected best modes are smaller than in Fig. 2(a) and the NZC are more spread out. The phenomenon is even more pronounced in Fig. 2 (d), which has the smallest selected best mode.

2.3 Mode priority decision by base layer

As mentioned in [4], Wang *et al.* analyze the R-D cost correlation between the BL and the EL, and discover that the R-D costs of each sub-MB_{BL} can be used to efficiently decide the mode priority of the EL. Eq. (1) shows the R-D costs of all sub-MB_{BL} including Mode 8×8, Mode 8×4, Mode 4×8 and Mode 4×4 after sorting. In Eq. (1), J_{BL} is a sorted set of R-D costs for sub-MB in BL, wherein the elements are sorted by increasing magnitude and given corresponding indices, such that $J_{BL\_1} < J_{BL\_2} < J_{BL\_3} < J_{BL\_4}$. The minimum J in J_{BL} is set to be the first priority and the maximum J is set to be the last one. We use the R-D cost order of the BL sub-MB as EL's execution order for their corresponding mode. Mode 8×8, Mode 8×4, Mode 4×8 and Mode 4×4 in BL correspond to Mode 16×16, Mode 16×8, Mode 8×16 and Mode 8×8 in EL, respectively. For example, if the first mode priority in BL is Mode 8×4, then we will execute Mode 16×8 first in EL. The mode priority of EL is represented by $\text{Mode}_{BL\text{-}p}$ (p = 1 to 4). $\text{Mode}_{BL\text{-}1}$ means the first priority mode and $\text{Mode}_{BL\text{-}4}$ represents the last one.

$$J_{BL} = \{ J_{BL\_p} \mid J_{BL\_1} < J_{BL\_2} < J_{BL\_3} < J_{BL\_4},\ p = 1 \text{ to } 4 \} \tag{1}$$

2.4 Adaptive search range

We also analyze the correlation between the search range and the modes by experiments. In the coding of the base layer, the R-D cost from Mode 8×8 to Mode 4×4 are sorted and used as the coding order of enhancement layer Mode 16×16 to Mode 8×8. In the coding of the MB in the enhancement layer, we first check the corresponding mode in the base layer, if it is Mode SKIP, then according to Table I which empirically evaluates the relationship between the bit rate and the search range, the search range for the enhancement layer is set to 10, and coding is only performed for Mode SKIP and Mode 16×16. The search range has been

determined through observations of previous experiments. According to Table II, the decrement in ΔBit rate begins to taper off for search range larger than 4. Therefore, we set search range 4 as the minimum search range. If the corresponding mode in the base layer is not Mode SKIP, then the search range for the enhancement layer is determined based on Eq. (2).

Table I. Relationship between bit rate and search range for Mode SKIP in the base layer.

$SR_{SKIP_{BL}}$	SR = 6	SR =8	SR = 10	SR = 12
Δ Bit rate(%)	0.114%	0.012%	-0.002%	-0.003%

Table II. Analysis of the minimum SR for EL

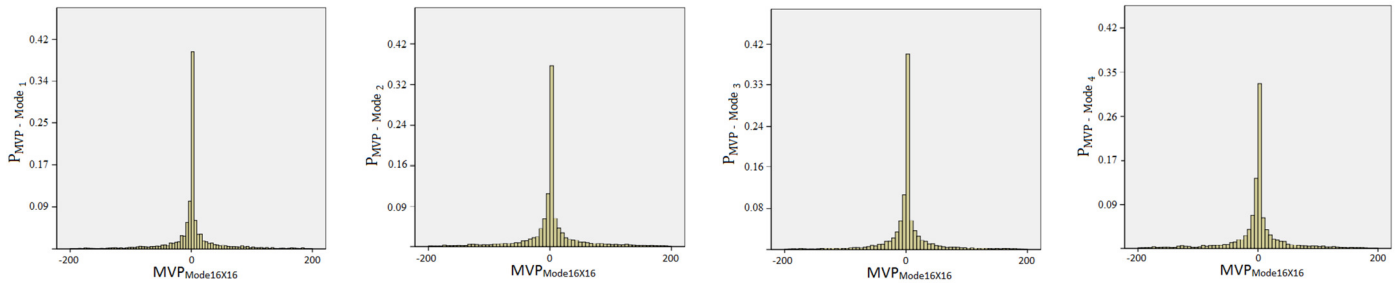
Min SR_{EL}	SR = 2	SR =3	SR = 4	SR = 5
Δ Bit rate(%)	0.124%	0.108%	0.085%	0.073%

Through the experiments, we propose an adaptive search range algorithm to shrink the search range by the following conditions:

(a) If an MB is only coded with Mode SKIP and Mode 16×16, the search range is set as 10.

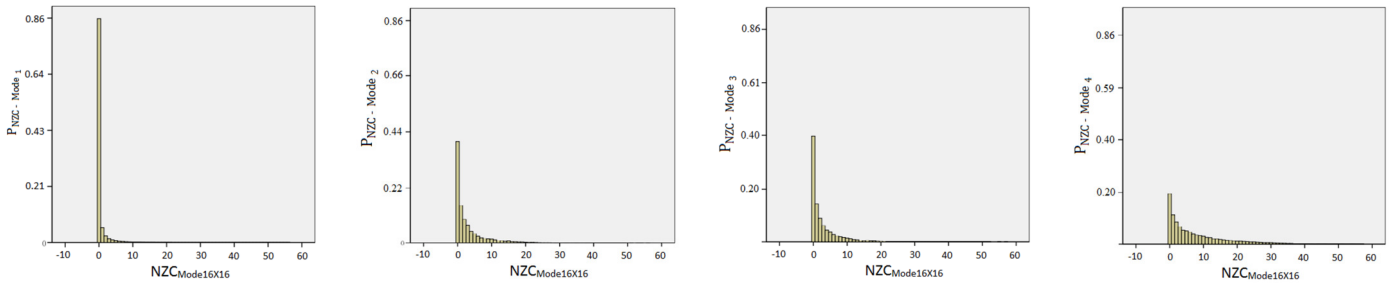
(b) Otherwise, the search range is set to $2\times MVD_{BL}$. If MVD_{BL} is less than 2, then the search range is set to 4 as shown in Eq.(2).

$$SR_{EL} = max\{2\times MVD_{BL}, 4\} \tag{2}$$



(a) (b) (c) (d)

Figure 1. Probabilities of Mode 16×16's MVP in EL for final selected modes with (a) Mode 16×16, (b) Mode 16×8, (c) Mode 8×16 and (d) Mode 8×8.



(a) (b) (c) (d)

Figure 2. Probabilities of Mode 16×16's NZC in EL for final selected modes with (a) Mode 16×16, (b) Mode 16×8, (c) Mode 8×16 and (d) Mode 8×8. (BL：QCIF，EL：CIF)；QP(28/28)

3. PROPOSED ALGORITHM

Our proposed algorithm uses statistic models to find the most probable candidate modes. The purpose is to calculate the best estimator from a known probability density function. From previous analyses of MVP and NZC, we can use the probability density function to help us predict the candidate modes of the current block. By experiment results, these two probability distributions are similar and can be approximated as Laplacian distributions. We model the probability distribution of MVP of Mode 16×16 as shown in Eq. (3) and NZC as shown in Eq. (4).

$$P_{MVP-Mode_i}(MVP) = P_{MVP-Mode_i}(MVP_x) * P_{MVP-Mode_i}(MVP_y)$$
$$= \frac{1}{2\sigma_{i,mvp_x}} exp\left[\frac{-|MVP_x-\theta_{i,mvp_x}|}{\sigma_{i,mvp_x}}\right] * \frac{1}{2\sigma_{i,mvp_y}} exp\left[\frac{-|MVP_y-\theta_{i,mvp_y}|}{\sigma_{i,mvp_y}}\right],\ i=1,2,3,4 \quad (3)$$

$$P_{NZC-Mode_i}(NZC) = \frac{1}{2\sigma_{i,NZC}} exp\left[\frac{-|NZC-\theta_{i,NZC}|}{\sigma_{i,NZC}}\right],\ i=1,2,3,4 \quad (4)$$

where σ_i and θ_i are the standard deviation and the median of all MVP and NZC of Mode 16×16 when $Mode_i$ is selected as the best mode. In Eqs. (3) and (4), $i=1$ represents Mode 16×16, $i=2$ represents Mode 16×8, $i=3$ represents Mode 8×16, $i=4$ represents Mode 8×8. In Eq. 3, σ_i and θ_i are the standard deviation and the median of MVP for Mode 16×16 when $Mode_i$ is selected as the best mode. In the experiments, the σ_i values are as shown in Table III, and θ_i is set to 0.

Table III Standard deviation for the MVP when $Mode_i$ is selected as the best mode.

$Mode_i$	MVP-X	MVP-Y
$Mode_1$	42.4438	16.3960
$Mode_2$	47.0289	22.8491
$Mode_3$	42.8754	22.3743
$Mode_4$	45.4210	20.5889

In Eq. (4), the probability models based on different best modes and QP have corresponding parameters. Table IV shows the parameters for the MVP probability model under different best modes. We have observed that the values of σ and θ vary according to the QP, therefore, we use the σ generated from QP=(28, 32 , 36, 40) to approximate the graph for each mode as shown in Figs. 3 and 4. According to different QPs, the values of σ_{Mode_i} and θ_{Mode_i} for probability models of different modes can be approximated by curve approximations, as shown in Eqs. (5) and (6), and the curve parameters for different modes are shown in Tables V and VI.

$$\sigma_{Mode_i} = A_i QP^2 + B_i QP + C_i \quad (5)$$
$$\theta_{Mode_i} = INT(E_i QP + F_i) \quad (6)$$

Table IV Values of σ and θ for NZC probability models

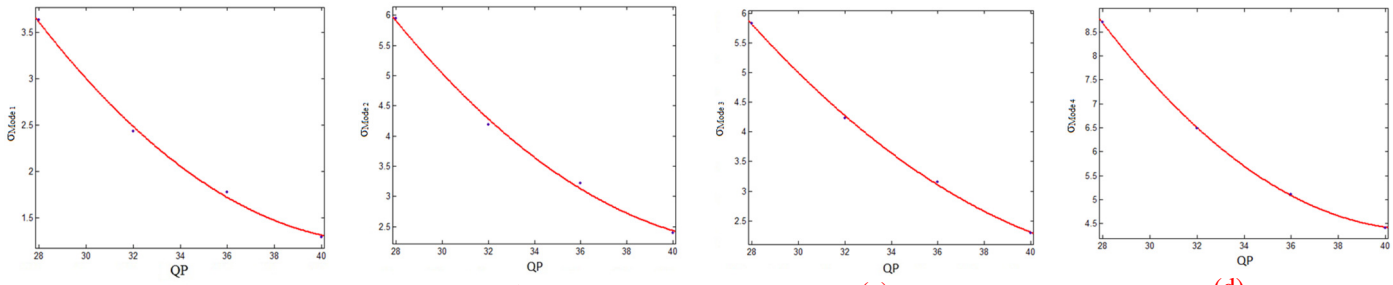
(BL:QCIF/EL:CIF)	QP = 28		QP =32		QP = 36		QP = 40	
	σ	θ	σ	θ	σ	θ	σ	θ
$Mode_1$	3.63	0	1.77	0	2.43	0	1.29	0
$Mode_2$	5.94	1	3.22	0	4.19	0	2.4	0
$Mode_3$	5.83	1	3.15	0	4.23	0	2.29	0
$Mode_4$	8.70	4	5.11	2	6.48	3	4.41	1
(BL:CIF/EL:4CIF)	QP = 28		QP =32		QP = 36		QP = 40	
	σ	θ	σ	θ	σ	θ	σ	θ
$Mode_1$	7.94	0	5.60	0	3.53	0	3.44	0
$Mode_2$	10.58	1	7.99	1	6.08	0	5.05	0
$Mode_3$	11.74	2	8.66	1	6.18	1	5.32	0
$Mode_4$	14.96	7	10.31	5	7.97	3	7.27	1

Table V Curve parameters for approximating the σ values of different modes

(BL:QCIF/EL:CIF)			
$Mode_i$	A_i	B_i	C_i
$Mode_1$	0.010	-0.96	21.59
$Mode_2$	0.014	-1.28	30.30
$Mode_3$	0.112	-1.08	26.96
$Mode_4$	0.024	-0.97	45.26
(BL:CIF/EL:4CIF)			
$Mode_i$	A_i	B_i	C_i
$Mode_1$	0.04	-2.78	58.30
$Mode_2$	0.02	-2.12	50.84
$Mode_3$	0.03	-2.90	65.86
$Mode_4$	0.06	-4.77	100.9

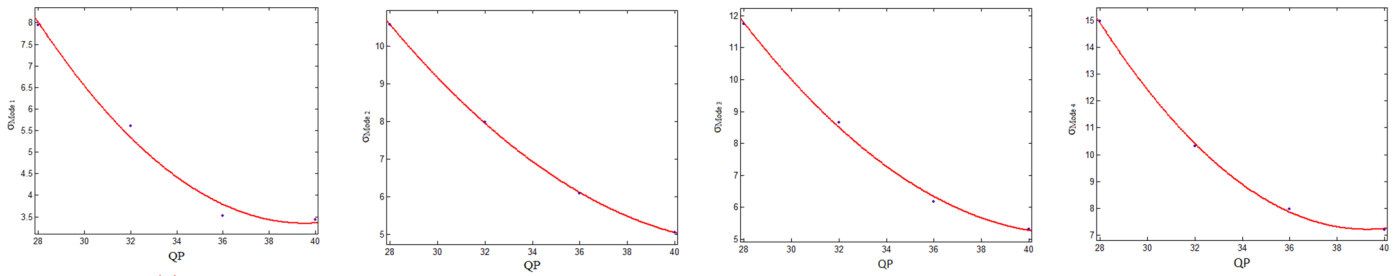
Table VI Curve parameters for approximating the θ values of different modes

(BL:QCIF/EL:CIF)		
$Mode_i$	E_i	F_i
$Mode_1$	0.00	0.0
$Mode_2$	-0.07	2.8
$Mode_3$	-0.07	2.8
$Mode_4$	-0.20	9.3
(BL:CIF/EL:4CIF)		
$Mode_i$	E_i	F_i
$Mode_1$	0.00	0.0
$Mode_2$	-0.10	3.9
$Mode_3$	-0.15	6.1
$Mode_4$	-0.50	21



(a) (b) (c) (d)

Figure 3. Graphs for QP and σ of NZC probability models for different modes (BL:QCIF, EL:CIF)
(a)Mode 16×16 (b)Mode 16×8 (c) Mode 8×16(d)Mode 8×8



(a) (b) (c) (d)

Figure 4. Graphs for QP and σ of NZC probability models for different modes (BL:CIF, EL:4CIF)
(a)Mode 16×16 (b)Mode 16×8 (c) Mode 8×16(d)Mode 8×8

After determining the search range, we code with Mode SKIP and Mode 16×16 and record the MVP (in both X and Y directions) for Mode 16×16 and the number of NZC. The data are input into the probability models in Eqs. (3) and (4) to calculate the probability of each candidate mode for being the best mode. The mode numbers of the most probable candidate modes predicted are represented by i_{MVP} and i_{NZC} as in Eqs. (7) and (8), where i_{MVP} represents the mode with the highest probability as predicted by MVP and i_{NZC} represents the mode with the highest probability as predicted by NZC.

$$i_{MVP} = \arg_i MaxP_{MVP-Mode_i}(MVP)\ ,\ \ i = 1,2,3,4 \tag{7}$$

$$i_{NZC} = \arg_i MaxP_{NZC-Mode_i}(NZC)\ ,\ \ i = 1,2,3,4 \tag{8}$$

If the candidate modes i_{MVP} and i_{NZC} include both Mode 16×16 and Mode 8×8, as in Eq. (9), then the candidate modes are deemed inaccurate; in which case, all modes will be evaluated and the algorithm will not make an early termination. Otherwise, the algorithm proceeds to make an early termination by executing the following steps. The first mode (p=1) obtained by Eq. (1) is selected for coding. The result is compared with the predictions made by Eqs. (3) and (4), if the resultant mode is the same as the predicted mode, then the algorithm makes an early termination. Otherwise, the next mode in the sequence is selected for coding, and the process repeats itself, until modes predicted by Eqs. (3) and (4) have been used for coding. Once modes predicted by Eqs. (3) and (4) have been used for coding, the algorithm makes an early termination to reduce the coding time.

$$|i_{MVP} - i_{NZC}| == 3 \tag{9}$$

We utilize the information of Mode 16×16 to calculate the most probable mode as the candidate mode in our algorithm. Table VII shows the hit rate of selecting the correct mode by using different estimators. When only one of these two candidate modes generated by probability models of MVP($\mathrm{Mode}_{i_{MVP}}$) and NZC ($\mathrm{Mode}_{i_{NZC}}$) is used to predict the best mode, the hit rate would be under 66%. If we consider both candidate modes to be the best mode, the hit rate will increase to 77%.

Table VII. Mode hit rate by probability models

Mode	$\mathrm{Mode}_{i_{MVP}}$	$\mathrm{Mode}_{i_{NZC}}$	$\mathrm{Mode}_{i_{MVP}}$ and $\mathrm{Mode}_{i_{NZC}}$
Hit Rate	58.76%	65.23%	77.56%

The flowchart of the proposed fast encoding algorithm is shown in Fig. 5. We use the information of co-located MB in BL and the Mode 16×16 in EL to make an early selection of the best mode of EL in order to reduce the number of the candidate modes. The details of our algorithm are described as follows:

step 1 Check the co-located MB in BL. If it is coded by Mode SKIP, the current MB in EL will only encode with Mode SKIP and Mode 16×16. The search range will be set to 10; otherwise the search range will be set as Eq. (2), then go to step 2.

step 2 Evaluate Mode SKIP and Mode 16×16 first, and calculate the probabilities of candidate modes with the MVP according to Eq.(3) and the number of NZC according to Eq.(4) to generate two of the most likely candidate modes. If the two generated candidate modes include both Mode 16×16 and Mode 8×8, it means the prediction results are not very consistent, evaluate all modes; otherwise, go to step 3.

step 3 Use mode priority defined in Section 2.3 as an order of mode checking. If candidate modes generated from step 2 are all evaluated, then an early termination on the mode decision is made.

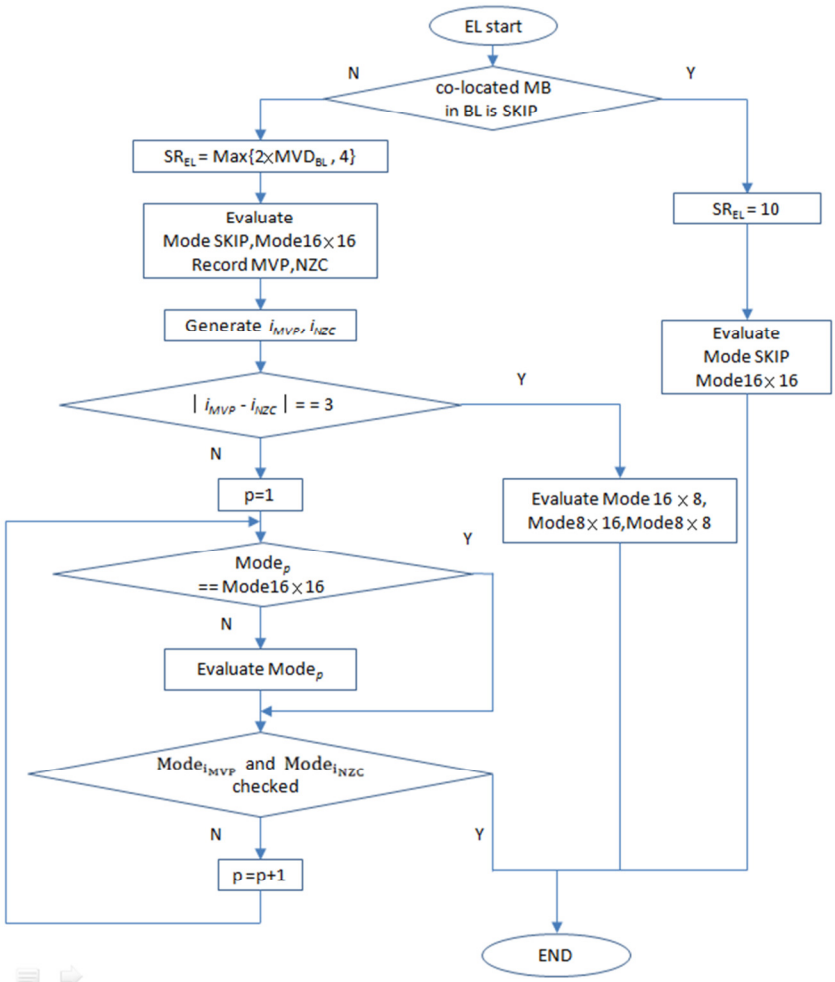


Figure 5. Flowchart of the proposed algorithm

Table VIII shows the probabilities of executing the steps in the proposed algorithm for four different test sequences. The probabilities of executing each step in the proposed algorithm differ according to the motion characteristics and the image complexities. Sequences with less motion have higher probabilities of termination at step 1, whereas sequences with more motion or higher complexities have higher probabilities of termination at step 3.

Table VIII. Probability of executing each step in the proposed algorithm.

	Step 1	Step 2	Step 3
Bus	48.54%	10.54%	40.92%
Foreman	51.35%	8.21%	40.44%
Mobile	55.78%	7.12%	37.10%
M.D.	76.25%	3.52%	20.23%

4. EXPERIMENTAL RESULTS

The proposed algorithm is implemented with the JSVM 9.18 reference software [10]. The training sets used in our work are Akiyo, Foreman, Football, Soccer, Stefan, Coastguard, Table and News, which are different from the test sets. In this work, there are twelve test sets, among the test sets, Bus, City, Crew, Football, Foreman, Harbour, Mobile, and M.D. have resolutions of QCIF(176x144)/CIF(352x288), and City, Crew, Harbour, and Soccer have resolutions of CIF(352x288)/4CIF(704x576). Out of these video sequences, Bus and Mobile have larger motion vectors.

Table IX shows the details of the experimental settings. We compare the performance of our algorithm with Lu and Martin's method in [5]. Five measurements TS, ΔPSNR, ΔBR, BDPSNR and BDBR [11] are used to evaluate the performance of the proposed algorithm. TS, ΔPSNR and ΔBR are metrics defined in Eqs. (10) to (12).

$$Time\ saving\ (TS) = \frac{Time_{JSVM} - Time_{proposed}}{Time_{JSVM}} \times 100\% \tag{10}$$

$$\Delta PSNR = PSNR_{proposed} - PSNR_{JSVM} \tag{11}$$

$$\Delta BR = \frac{BR_{proposed} - BR_{JSVM}}{BR_{JSVM}} \times 100\% \tag{12}$$

Table IX. Experimental environment

JSVM reference software	JSVM. 9.18
PC-CPU	Intel(R) Core(TM) i5-3570 3.4GHZ
Encoding format	Hierarchical B
PC-RAM	4.00GB
GOP	16
Frame encoded	150
Search range	32
Motion estimation	Full search
Entropy coding	CABAC
Frame size	QCIF/CIF , CIF/4CIF

Table X. Performance comparisons of the proposed algorithm with Lu's algorithm in different QPs (QP_{BL}/QP_{EL} = 30/25, 30/30, 30/35) for QCIF/CIF spatial scalability

Seq.	QP	Lu [5]			Proposed		
		ΔPSNR	ΔBR	TS	ΔPSNR	ΔBR	TS
Bus	30/25	-0.12	1.51	62.25%	-0.05	0.54	65.86%
	30/30	-0.09	1.08	62.86%	-0.06	0.24	67.25%
	30/35	-0.05	0.45	63.77%	-0.05	-0.09	68.13%
City	30/25	-0.04	1.86	67.84%	-0.10	0.30	76.45%
	30/30	-0.03	1.03	69.04%	-0.12	0.10	78.02%
	30/35	-0.02	0.12	70.47%	-0.08	-0.73	78.03%
Crew	30/25	-0.03	1.96	60.58%	-0.07	0.23	68.15%
	30/30	-0.01	1.02	61.29%	-0.08	-0.42	69.62%
	30/35	-0.03	0.68	62.25%	-0.06	-0.74	71.38%
Football	30/25	-0.05	0.92	57.79%	-0.03	0.27	56.80%
	30/30	-0.05	0.48	58.29%	-0.05	-0.06	58.72%
	30/35	-0.03	0.17	59.22%	-0.02	-0.16	60.45%
Foreman	30/25	-0.14	1.87	63.00%	-0.12	0.41	73.56%
	30/30	-0.07	0.61	64.23%	-0.12	-0.53	74.47%
	30/35	-0.03	-0.24	65.22%	-0.06	-0.42	75.23%
Harbour	30/25	-0.16	0.81	68.19%	-0.04	0.43	75.77%
	30/30	-0.07	0.60	68.77%	-0.05	0.00	75.94%
	30/35	-0.04	0.14	69.95%	-0.06	-0.29	76.25%
Average		-0.06	0.84	64.17%	-0.07	-0.05	70.56%

Table XI. Performance comparisons of the proposed algorithm with Lu's algorithm in the same QPs (QP_{BL}/QP_{EL} = 28/28, 32/32, 36/36, 40/40) for QCIF/CIF spatial scalability

Seq.	Lu [5]			Proposed		
	BDPSNR	BDBR	TS	BDPSNR	BDBR	TS
Bus	-0.10	1.96%	63.83%	-0.10	1.94%	69.62%
Foreman	-0.07	1.44%	67.79%	-0.10	1.94%	76.34%
Mobile	-0.06	1.09%	69.88%	-0.08	1.94%	76.25%
M.D.	-0.02	0.43%	80.19%	-0.01	0.29%	81.85%
Average	-0.06	1.23%	70.42%	-0.07	1.53%	76.02%

Table XII. Performance comparisons of the proposed algorithm with Lu's algorithm for CIF/4CIF spatial scalability

Seq. CIF/4CIF	Lu [5]			Proposed		
	BDPSNR	BDBR	TS	BDPSNR	BDBR	TS
City	-0.02	0.38%	59.13%	-0.03	0.97%	76.93%
Crew	-0.02	0.60%	59.39%	-0.04	1.25%	72.28%
Harbour	-0.04	0.97%	57.21%	-0.04	1.17%	73.77%
Soccer	-0.04	1.05%	59.85%	-0.05	1.24%	72.48%
Average	-0.03	0.75%	58.90%	-0.04	1.16%	73.87%

Table XIII. The actual encoding time of the proposed algorithm and time saving at different QPs (QP_{BL}/QP_{EL} = 30/25, 30/30, 30/35) for QCIF/CIF spatial scalability

Seq.	QP	Original coding time (Seconds)	Proposed coding time (Seconds)	Time Saving
Bus	30/25	26891	9178	65.86%
	30/30	27097	8871	67.25%
	30/35	26835	8550	68.13%
City	30/25	24714	5819	76.45%
	30/30	24709	5428	78.02%
	30/35	24504	5381	78.03%
Crew	30/25	27822	8861	68.15%
	30/30	27825	8453	69.62%
	30/35	27714	7929	71.38%
Football	30/25	28257	12205	56.80%
	30/30	28404	11724	58.72%
	30/35	28152	11132	60.45%
Foreman	30/25	25532	6748	73.56%
	30/30	25663	6519	74.47%
	30/35	25424	6297	75.23%
Harbour	30/25	28042	6792	75.77%
	30/30	27793	6686	75.94%
	30/35	27783	6597	76.25%

The performance of our proposed algorithm is compared with Lu and Martin's method shown in Tables X to XII, and the actual encoding time of the proposed algorithm is shown in Table XIII. Regardless of the QP's between BL and EL, higher coding efficiency with lower bitrates can be achieved by our algorithm. The results show that the proposed algorithm can efficiently and accurately select the best mode. In high motion sequences such as Bus, Football, or Soccer, the abundance in block information leads to better prediction and better performance. For higher resolution sequences, it is clear that our algorithm can achieve higher time saving than Lu and Martin's method. Nevertheless, the performance of our proposed algorithm might still be ineffective for R-D performance of high resolution videos. Therefore, in the future, we will analyze the characteristics of high resolution videos and focus on a solution to improve the performance of the proposed method.

5. CONCLUSIONS

This paper proposes an algorithm for fast mode decision based on statistic models. The proposed algorithm utilizes the probability models generated by motion vector predictor (MVP) and the number of non-zero coefficients (NZC) of Mode 16×16 in EL to determine the most probable modes. As experiment results demonstrate, in addition to being able to efficiently and accurately select the best mode, the proposed algorithm is able to adaptively reduce the search range for motion estimation. Overall, the proposed method achieves up to 76% time saving on average with better coding performance in terms of coding time when compared to prior work.

6. 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-X20*1, 2007.

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

[3] L. Shen, Y. Sun, Z. Liu, and Z. Zhang, "Efficient SKIP mode detection for coarse grain quality scalable video coding," *IEEE Signal Processing Letters*, vol. 17, no. 10, pp.887-890, October 2010.

[4] T.H. Wang, M.J. Chen, M.C. Chi, S.F. Huang and C.H. Yeh, "Computation-scalable algorithm for scalable video coding," *IEEE Transactions on Consumer Electronics*, vol. 57, no. 3, pp.1194-1202, August 2011.

[5] X. Lu, and G.R. Martin, "Fast mode decision algorithm for the H.264/AVC scalable video coding extension," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 23, no. 5, pp. 846-855, May 2013.

[6] T.Y. Kuo and C.H. Chan, "Fast variable block size motion estimation for H.264 using likelihood and correlation of motion field," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 16, no. 10, pp.1185-1195, October 2006.

[7] S.C. Huang, C.T. Hsu, M.J. Chen, "Efficient motion re-estimation for H.264 B- and P-Frame transcoding by using maximum likelihood," *in Proceedings of International Conference on Electronics and Information Engineering*, pp.V1-557 - V1-561, 2010.

[8] J. Lee, M. Choi, Y. Cho, J. Kim and W.K. Cho, "Fast H.264/AVC motion estimation algorithm using adaptive search range," *in Proceedings of the12thInternational Symposium on Integrated Circuits,* pp. 336-339, 2009.

[9] B.G. Kim, K. Reddy, and K.W. Lim, "Dynamic search range control algorithm for inter-frame coding in scalable video coding," *in Proceedings of IEEE International Conference on Multimedia and Expo*, pp. 225-228, 2009.

[10] JSVM (Joint Scalable Video Model) reference software for SVC. [Online].Available:http://ip.hhi.de/imagecom/G1/savce/downloads/SVCReference-Software.htm

[11] G. Bjϕntegaard, "Calculation of average PSNR differences between RD curves," *ITU-T VCEG-M33*, VCEG 13th meeting, Austin, Texas, USA, April 2001.