Macroblock Mode Pre-classification Algorithms based on Motion Vectors Filtering for H.264/AVC

Cheng Zhou
College of Electronic Information Engineering, South-Center University for Nationalities, Wuhan, China
Email:zcresearch@gmail.com

Yanhong Lu
Wuhan Polytechnic, Wuhan, China
Email: luyanhong126@gmail.com

Chengyi Xiong
Hubei Key Laboratory of Intelligent Wireless Communications, South-Center University for Nationalities, Wuhan, China
Email: xiongcy@mail.scuec.edu.cn

Zhirong Gao
College of Computer Science, South-Center University for Nationalities, Wuhan, China
Email: gaozhirong@mail.scuec.edu.cn

***Abstract*—Noise robustness of video codec is a big challenge. This paper proposes a novel macroblock mode pre-classification algorithm based on the co-matching criteria, motion vectors spatial and temporal filtering. The proposed algorithm could distinguish the moving object from the background noise. Simulations show that this approach can result in a time savings of over 62.86% for several typical sequences with noise. It also reduces the average Bjontegaard delta bit rate by about 1.67%, and increases the average Bjontegaard delta peak signal-to-noise ratio by about 0.08dB, compared with the algorithm of H.264/AVC.**

***Index Terms*—video coding, noise robustness, motion estimation, mode decision, motion vectors filtering, H.264/AVC**

I. INTRODUCTION

With the rapid growth of network video system, compression and pre-processing analysis of massive digital video becomes more and more important. High efficient video coding technology is the key to solve these problems. However, it appears many problems in the practical application of video system. Firstly, when video noise increases, the video coding efficiency considerably decreases. Secondly, the new generation of video coding standard such as H.264/AVC [1] can obtain higher compression efficiency [2]. Unfortunately, it is difficult to design the hardware codec because of the complex coding control model [3]. Thirdly, as the complexity of the new generation of video coding standard is very high, the pre-processing and analysis of video data become more and more difficult [4].

This paper focuses on how to increase the noise robustness of video encoder. And the main contribution of this thesis is a novel macroblock mode pre-classification method based on co-matching criteria, motion vectors spatial and temporal filtering. The method uses the co-matching criteria to judge the current macroblock in order to eliminate the noise impact and use the temporal and spatial filtering of the motion vector fields of encoded frames to eliminate the noise motion vectors. According to the motion information of current macroblock, the method limits the coding mode of current macroblock. The proposed algorithm can improve the noise robustness of the H.264/AVC encoder. Simulations show that this approach can result in a time savings of over 62.86% for several typical sequences with noise.It also reduces the average Bjontegaard delta bit rate by about 1.67%, and increases the average Bjontegaard delta peak signal-to-noise ratio by about 0.08dB, compared with the algorithm of H.264/AVC. Experiments prove that this algorithm improves the coding performance and coding speed.

II. RELATED WORK

A. Coding Control Model in H.264/AVC

The H.264/AVC standard has been widely applied to real-time video system because it offers significantly better coding efficiency than previous standards [5]. The improvements are mainly obtained from various state-of-the-art techniques such as intra predictions, variable block-size motion compensation and quarter-pixel motion

This work is supported by National Natural Science Foundation of China under Grant # 61201268 to Cheng Zhou and Hubei Natural Science Foundation of China under Grant # 2010CDZ022 to Cheng Zhou.
Corresponding author: Yanhong Lu

doi:10.4304/jmm.8.2.129-136

estimation (ME) [6]. These techniques have been incorporated into a high complexity mode decision (HMD) algorithm framework with rate-distortion optimization (RDO) technique, which ensures the highest efficient macroblock (MB) mode is chosen for final encoding [7].

H.264/AVC also adopts more MB modes than any other video coding standards. In a P-picture, in order to find the best mode for each MB, the HMD algorithm with RDO tries all possible inter and intra modes [8] (SKIP, P16×16, P16×8, P8×16, P8×8, I16×16, I4×4, I8×8, IPCM). Firstly, ME and RDO calculation are performed to find the best motion vector (MV) of each MB inter mode that minimizes:

$$J_{MOTION}(Mode, \lambda_{MOTION}) = SA(T)D(mv, ref) + \lambda_{MOTION} \cdot J_{mv}(Mode) \cdot \quad (1)$$

where MODE denotes one of the potential inter prediction modes, λ_{MOTION} is the Lagrangian multiplier2 which equals to 0.85×2(QP/3), REF indicates the reference picture and JMV(MODE) represents the bits used for coding MV and the best reference picture index. The sum of absolute difference (SAD) is used for integer pixel ME and the sum of absolute transformed difference (SATD) is used for sub-pixel ME. And SAD is defined as follows:

$$SAD(i, j) = \sum_{m=1}^{M} \sum_{n=1}^{N} \left| f_k(m,n) - f_{k-1}(m+i, n+j) \right| \cdot \quad (2)$$

Where $f_k(m,n)$ denotes the MB pixels on the current frame and $f_k(m,n)$ indicates the MB pixels on the reference frame [9].

After the first round calculation of the preliminary coding cost of the inter mode, all of inter and intra modes are revaluated by using Lagrange function as:

$$J_{MODE}(Mode \mid QP, \lambda_{MODE}) = SSD(Mode \mid QP) + \lambda_{MODE} \cdot J_{cod}(Mode \mid QP) \cdot \quad (3)$$

where MODE indicates one of the potential inter and intra prediction modes, QP denotes the MB quantization parameter, λ_{MODE} is $\lambda_{MOTION}{}^2$ and J_{COD}(MODE|QP) represents the bits used for mode coding including MB header, MV and DCT coefficients. The sum of squared differences (SSD) is the residue between the original MB and its reconstruction. Finally, the mode that minimizes Lagrange function (3) is selected for the final coding [10].

It is clear that the coding control model of H.264/AVC uses exhaust algorithm which can test all possible coding modes to achieve the best coding efficiency. Moreover, the sophisticated model takes all of coding cost into consideration such factors as motion estimation residuals, motion vectors, reference frame indexes and MB mode indexes. Consequently, the coding control model could minimize the coding cost when the video quality is good [11].

However, when the video quality is not good, the coding efficiency of the coding control model considerably decreases. The reason will be analyzed in the next section.

B. Noise Robustness in Encoder

The digital camera usually works in an all-weather environment where is probably filled with stochastic noise caused by low-level lighting environment such as night rain, snow and fog. This can significantly decrease performance of the video encoder. Consequently, the noise robustness is one of the basic problems of video coding. It can be validated that the performance of the H.264/AVC framework substantial degrades when the video noise increases. One important weak link is the block matching criterion of motion estimation which has decisive effects on the coding efficiency [12].

Essentially, matching criterion of motion estimation is an error metric function which could measure the similarity among the macroblocks. The selection of the matching criteria is directly related to the accuracy of the similarity and the accuracy of motion vectors. And SAD/STAD is not always accurate enough for motion estimation [13].

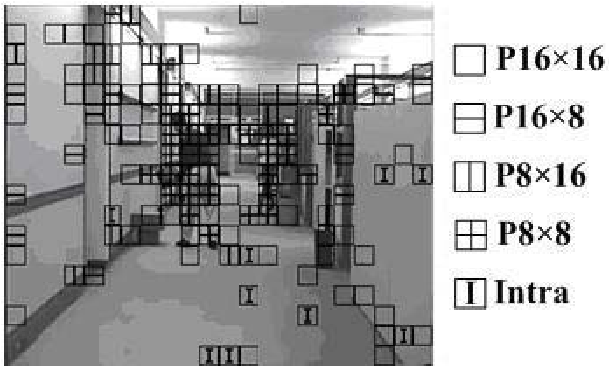

Figure 1. Mode decision results comparison for "Hall_monitor" sequence (except the SKIP mode, using QP=28, 96th frame).

We can take sequence "Hall_monitor" as an example. "Hall_monitor" is a fixed-camera sequence with noise caused by low illumination which contains two man walked through the corridor. We can find that there is no moving object except the man in the 96th frames of "Hall_monitor" sequence. Consequently, the encoder of H.264/AVC should choose SKIP mode for the background to achieve the best coding efficiency because SKIP mode has no further data needed to be transmitted after the skip indication. Figure 1 shows the mode decision results for "Hall_monitor" sequence. It can be seen from the graph that encoder used a lot of sub-mode to describe the 2 man's bodies such as P16×8, P8×16, P8×8, etc. However, figure 1 shows that over 30% best modes of background are not SKIP mode in the 96th frame of "Hall_monitor" sequence, which means that the video noise caused by low illumination severely affects the accuracy of predictive operations and more bit stream are wasted.

Robustness of encoder is one of the most important issues of video coding. However, the noise generated in video capture has not been paid enough attention by research groups [14]. We can see several noise reduction approaches have been presented to overcome this problem.

There are two major methods from relative works. Mode decision based on the background predicts is the first way to improve the motion estimation accuracy [15-22]. Kim [16] and Zhu [17] proposed several algorithms to predict the motion status of macroblock. However, those algorithms only reduce the number of candidate modes of H.264/AVC and don't improve the matching criteria. Although the coding speed of encoder is improved, the coding efficiency of encoder always decreases.

The other way to improve noise robustness of encoder is using spatial filters to wipe out random video noise before coding processing [18-21]. It can be confirmed that those filters can improve subjective quality of video streams, and improve the noise robustness of encoder, but sometimes they also remove image details which are unrecoverable. And the computational complexity of these filters is still very heavy.

Therefore, in this paper, we focus on increasing the coding efficiency by effectively applying a novel macroblock pre-classification algorithm which has a very low-complexity. We use MB-level motion detection and motion vectors filtering technique to evaluate the current MB motion level and classify it. According to the classification, different mode decision strategies are chosen for current MB. And the details of this algorithm will be discussed in next section.

III. Proposed Algorithm

A. Initial Motion Status Mark for Macroblock

To improve the noise robustness of encoder, a more positive and accurate method is needed for mode decision. And the complexity of the method should be acceptable. The proposed algorithm takes full advantage of the temporal and spatial correlation to determine the motion status of current macroblock by using two different matching criteria which is not sensitive to video noise [10].

The first matching criterion is absolute value of the sum of difference (ASD), which is defined as follows:

$$ASD = \left| \sum_{m=0}^{15} \sum_{n=0}^{15} \left[f_k(m,n) - f_{k-1}(m,n) \right] \right| . \quad (4)$$

where $f_k(m,n)$ and $f_{k-1}(m,n)$ denote luminance pixel values in the current MB and temporal adjacent MB in the consecutive frame, respectively. Note that, when the noise on the luminance pixel values in the MB obeys white noise distribution with mean zero, the positive noise and negative noise would result in the cumulative error of SAD. For example, there is no moving object in the first two frames of "Hall_monitor" sequence. And by statistical analysis, the average SAD of the current MB and temporal adjacent MB in the first two frames of "Hall_monitor" sequence is 559.21, while the average ASD is 81.75. Obviously, it can be confirmed that the ASD can decimate the effect of white noise.

The second matching criterion is defined as the difference of MB center of gravity (DMCG), which can be calculated by using (5) and (6):

$$X_k = \frac{\sum_{m=0}^{15}\sum_{n=0}^{15} f_k(m,n)\cdot m}{\sum_{m=0}^{15}\sum_{n=0}^{15} f_k(m,n)}, \quad Y_k = \frac{\sum_{m=0}^{15}\sum_{n=0}^{15} f_k(m,n)\cdot n}{\sum_{m=0}^{15}\sum_{n=0}^{15} f_k(m,n)} . \quad (5)$$

$$DMCG = \left| X_k - X_{k-1} \right| + \left| Y_k - Y_{k-1} \right| . \quad (6)$$

where X_k and Y_k indicate the current MB center of gravity on the horizontal and vertical direction, respectively. And X_{k-1} and Y_{k-1} represent the temporal adjacent MB center of gravity in the consecutive frame. And DMCG is quite sensitive to motion and not appreciably affected by video noise.

By using the two matching criteria, initial motion status mark for macroblock can be divided into three categories: background macroblock, macroblock which may contain some moving objects and motion macroblock. And we also use a set of decision thresholds for division: α for ASD, β for DMCG and T for both of them. And those thresholds can be adjusted according to the degree of noise. When video noise increases, they should be appropriately increased, and vice versa should be reduced. According to the level of this two noise robust predictive operations, three different mode decision strategies are chosen adaptively.

Firstly, if the ASD and DMCG of current macroblock satisfy the following condition (7), the algorithm will mark it as a background macroblock, and use the SKIP mode for its coding.

$$ASD \le (QP - T)\cdot \alpha \ \text{ and } \ DMCG \le (QP - T)\cdot \beta . \quad (7)$$

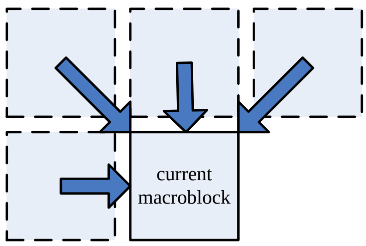

Figure 2. Define of spatially adjacent marcoblocks

Secondly, if ASD and DMCG of current macroblock satisfy the following condition (8) and one of the coded spatially adjacent macroblocks(left, up-left, up and up-right) is motion macroblock, the algorithm will mark it as a macroblock which may contain some moving objects and enable all the inter modes. And figure 2 shows how we define the spatially adjacent macroblocks. In this way moving objects pixel information loss is avoided and the details of moving objects is retained as much as possible.

$$\begin{cases} (QP - T)\cdot \alpha < ASD \le (QP - T)\cdot 2\alpha \\ (QP - T)\cdot \beta < DMCG \le (QP - T)\cdot 2\beta \end{cases} . \quad (8)$$

Finally, if ASD and DMCG of current macroblock do not satisfy the condition (7) and (8), the algorithm will mark it as a motion macroblock and use common coding strategy. And all possible inter and intra modes are available.

The first stage of proposed algorithm can be summarized as follows:

Step 1: Calculate SAD of the current macroblock, if SAD>T_{SAD}, go step 2, otherwise go step 4;
Step 2: Calculate ASD and DMCG of the current macroblock, If ASD and DMCG satisfy the condition (7), go step 3, and if one of the coded spatially adjacent macroblock is motion macroblock, go to Step 4, otherwise, go to Step 5.
Step 3: Mark the current macroblock as a background macroblock and go to Step 6.
Step 4: Mark the current macroblock as a macroblock which may contain some moving objects and go to step 5.
Step 5: Mark the current macroblock as a motion macroblock and go to Step 6.
Step 6: Put all mark result to the second stage of proposed algorithm.

Figure 3 shows macroblock mark results of the 96th frame of "Hall_monitor" sequence. And it obviously shows that the first stage of proposed algorithm successfully improves prediction accuracy of matching criteria.

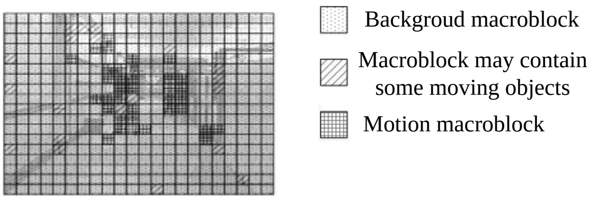

Figure 3. Macroblock mark results of the 96th frame of "Hall_monitor" sequence

B.Macroblock Classification Based on Motion Vectors Filtering

By analyzing the reason why the video noise affects coding efficiency and decline the matching criteria precision, the method that improves the noise robustness of matching criteria of inter prediction is proposed. After the pre initial motion status mark for macroblocks, many background macroblocks are corrected, which are wrongly considered as motion macroblocks. However, the mark results are not accurate enough and need to be further processed. And it's important to note that the computational complexity of improved method should not be too complex.

The proposed algorithm takes advantage of time domain correlation between adjacent frames in video sequences. It uses the motion vector fields in encoded frame to predict the macroblock status in the current coding frame and unite the macroblock marks in the first stage to decide the coding modes of the current macroblock.

Firstly, the coding modes of coded macroblocks are divided into three categories: macroblocks which are coded as SKIP mode, macroblocks whose motion vectors are not changed and macroblocks whose motion vectors are changed.

Macroblocks which are coded as SKIP mode do not have the coding information such as motion vector. Encoder will copy the corresponding macroblocks information for them. Generally, such kinds of macroblocks contain no motion or regular motion portion.

Macroblcoks whose motion vectors are not changed have the coding information. Such kinds of macroblocks usually contain a small part of motion object which lead to the growth of nonzero coefficients in the quantify stage and have be coded by choose different coding modes.

Macroblocks whose motion vectors are changed generally relate to the moving objects in video sequence. And its size describes the intensity of the moving objects. Particularly, these kinds of macroblocks may contain some noise motion vectors which is caused by error inter prediction.

Secondly, it is also necessary to remove the noise motion vectors in the coarse motion vector field by using smoothing filter. The proposed algorithm of useing the space median filter is defined as follows:

$$M_{space}(x,y) = Median_{space}[f_{N\times N}(i,j)]. \qquad (9)$$

where $f_{N\times N}(i,j)$ indicates motion vectors in an N×N window, and $M_{space}(x,y)$ denotes the median motion vector in the window.

The median filter takes advantage of spatial correlation of motion vector fields to correct the error motion vectors caused by video noise. The proposed algorithm also makes use of time-domain correlation of motion vector fields by using motion vectors on the coded frames to predict the motion status of current frame. We use 3 reference frames that are adjacent to the current frame for the prediction which are denoted in figure 4.

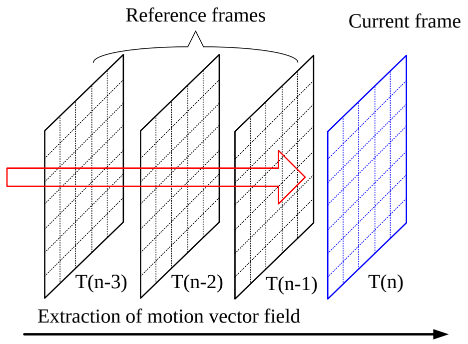

Figure 4. Macroblock status prediction using time-domain correlation

The current frame is defined as T(n) with its motion vector fields $MV_{(n)}$, and the adjacent reference frames are defined as T(n-1), T(n-2) and T(n-3) with their motion vector fields $MV_{(n-1)}$, $MV_{(n-2)}$ and $MV_{(n-3)}$. The proposed algorithm uses the time-domain median filter which is defined as follows:

$$M_{time}(x,y) = Median_{time}[MV_{(n-1)}(i,j), MV_{(n-2)}(i,j), MV_{(n-3)}(i,j)]. \qquad (10)$$

where $MV_{(n)}(i,j)$ denotes the motion vectors in the corresponding frames.

The second stage of proposed algorithm can be summarized as follows:
Step 1: Calculate the predict value of the current macroblock by using the time adjacent macroblocks.
Step 2: Calculate the absolute difference of the motion vector fields predictive values of the 4×4 sub-blocks in each marcroblock, which is defined as $PMV_{SAD}(i,j)$.

Step 3: Calculate the average predicted value of the current frame motion vector fields: $PMV_{mean} = PMV_{SAD}(i, j) / N_{MB}$.

Step 4: According to the average predicted value, set the motion prediction significant threshold: $T_{PMV} = \gamma \cdot PMV_{mean}$

Step 5: According to the motion prediction significant threshold, macroblock motion prediction status is divided into two categories: significant and non-significant.

C. Detailed Steps of Proposed Algorithm

After the two stages of macroblock status prediction, we finally can do a complete determination for macroblock motion status. According to the initial motion status mark and significance mark for Macroblock, the proposed algorithm will set the candidate coding modes for the current macroblock. Table I shows the rule for this setting. And the whole processing of proposed algorithm is shown in Figure 5.

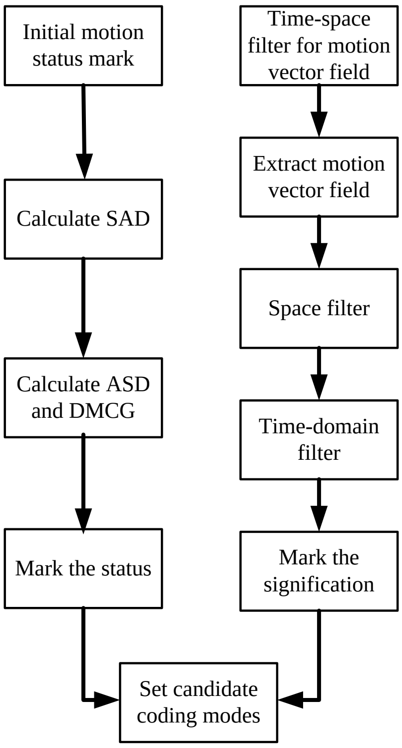

Figure 5. Processing of proposed algorithm

IV. EXPERIMENTS

A. Experiment Conditions

The proposed algorithm is implemented on H.264/AVC reference software version JM10.2 [23]. A Pentium E5200 CPU 2.50 GHz PC with 2GB RAM is used. To avoid the potential performance contribution of artificial noise, we use several common sequences with real noise. We also extent our test on some sequences provided by AVS organization [24]. Those sequences contain real noise caused by rain, fog, snow, night and low light. Each sequence has 100 frames and the simulation conditions are shown in table II.

For performance comparison, we use the BDPSNR and the BDBR [25]. And the time saving (TS) is defined as:

$$TimeSaving = \frac{Time(\text{original}) - Time(proposed)}{Time(\text{original})} \times 100\% \cdot \tag{11}$$

B. Objective Experiment Data Results

This section compares the performances of the proposed algorithm with the fast high-complexity mode decision algorithm in H.264/AVC reference software, which is adopted as fast mode decision algorithms of the H.264/AVC test model [4]. Table III to XII are the test results of each sequences and Table XIII is the comprehensive comparative statistic data.

As can be seen from these forms of experimental statistics, although the proposed algorithm loses some PSNR performance, compared with the algorithm of H.264/AVC at the same QP, it also reduces bitrates and encoding time at the same time. Consequently, table XIII shows the comprehensive performance comparison using BDPSNR and BDBR. That could prove that the proposed algorithm can not only efficiently reduce the encoding time, but also provide better coding efficiency in several sequences. When compared with the algorithm of H.264/AVC, overall time saving of proposed algorithm is about 62.86% with a PSNR gain of 0.08dB and bit rate saving of 1.67% on average, respectively. In particular, the proposed method has much higher coding efficiency than that of H.264/AVC in the "Bridge_far" and "dh_productline" sequences.

TABLE I.
SETTING RULES OF CANDIDATE CODING MODES

No.	Initial motion status mark	Significance mark	Candidate coding modes
1	Backgroud macroblock	Non-significant marcoblock	SKIP
2	Backgroud macroblock	Significant marcoblock	SKIP, P8×8
3	Macroblock may contain some moving objects	Non-significant marcoblock	SKIP, P16×16
4	Macroblock may contain some moving objects	Significant marcoblock	SKIP, P16×16, P8×8
5	Motion macroblock	Non-significant marcoblock	All inter modes
6	Motion macroblock	Significant marcoblock	All modes

TABLE II.
SIMULATION CONDITIONS

Parameters	Value
Frame rate	30.0
Frames to be encoded	100
Profile	Baseline
Number of reference frames	5
Structure of frames	IPPP
Quantization parameters	28, 32, 36 and 40
Entropy coding	CAVLC
RD-optimized mode decision	Fast high complexity mode
Search range	32
Motion estimation search mode	UMHexagon search

TABLE III.
TEST RESULTS OF "HALL_MONTIOR" SEQUENCE

QP	Original algorithm			Proposed algorithm		
	PSNR -Y(dB)	Bitrate (kbit/s)	Coding time(s)	PSNR Y(dB)	Bit-rate (kbit/s)	Coding time(s)
28	37.8	220.4	98.0	37.6	208.6	64.8
32	35.5	110.0	91.9	35.4	105.2	50.8
36	33.1	64.1	89.3	33.0	61.6	40.4
40	30.6	38.8	87.3	30.6	38.1	32.6

TABLE IV.
TEST RESULTS OF "BRIDGE_FAR" SEQUENCE

QP	Original algorithm			Proposed algorithm		
	PSNR -Y(dB)	Bitrate (kbit/s)	Coding time(s)	PSNR Y(dB)	Bit-rate (kbit/s)	Coding time(s)
28	37.3	57.8	89.6	37.2	52.6	31.5
32	36.0	11.7	83.7	36.0	11.1	18.8
36	34.8	6.9	81.7	34.7	6.8	12.0
40	33.3	5.5	80.1	33.3	5.4	7.8

TABLE V.
TEST RESULTS OF "NIGHT" SEQUENCE

QP	Original algorithm			Proposed algorithm		
	PSNR -Y(dB)	Bitrate (kbit/s)	Coding time(s)	PSNR Y(dB)	Bit-rate (kbit/s)	Coding time(s)
28	36.8	61.9	75.5	60.9	69.7	60.9
32	36.1	12.2	63.6	12.2	58.1	12.2
36	35.2	7.2	61.3	7.1	55.5	7.1
40	33.9	5.6	60.4	5.6	54.2	5.6

TABLE VI.
TEST RESULTS OF "FOG" SEQUENCE

QP	Original algorithm			Proposed algorithm		
	PSNR -Y(dB)	Bitrate (kbit/s)	Coding time(s)	PSNR Y(dB)	Bit-rate (kbit/s)	Coding time(s)
28	37.8	24.2	86.3	37.6	21.1	4.0
32	35.9	14.7	84.0	35.8	14.2	3.0
36	33.8	11.4	82.8	33.7	11.3	2.7
40	31.4	9.0	81.4	31.3	9.0	2.7

TABLE VII.
TEST RESULTS OF "RAIN" SEQUENCE

QP	Original algorithm			Proposed algorithm		
	PSNR -Y(dB)	Bitrate (kbit/s)	Coding time(s)	PSNR Y(dB)	Bit-rate (kbit/s)	Coding time(s)
28	35.3	72.3	94.8	35.2	68.9	23.5
32	33.3	38.5	90.7	33.2	38.0	18.8
36	31.1	23.4	88.2	31.1	23.3	15.5
40	29.1	15.6	86.5	29.1	15.5	12.9

TABLE VIII.
TEST RESULTS OF "DH_WORK" SEQUENCE

QP	Original algorithm			Proposed algorithm		
	PSNR -Y(dB)	Bitrate (kbit/s)	Coding time(s)	PSNR Y(dB)	Bit-rate (kbit/s)	Coding time(s)
28	36.1	251.5	106.4	35.9	241.5	58.4
32	33.4	134.5	102.3	33.3	131.6	51.8
36	30.9	75.1	98.6	30.8	74.1	46.5
40	28.4	41.6	95.4	28.3	40.9	42.1

TABLE IX.
TEST RESULTS OF "DH_PRODUCTLINE" SEQUENCE

QP	Original algorithm			Proposed algorithm		
	PSNR -Y(dB)	Bitrate (kbit/s)	Coding time(s)	PSNR Y(dB)	Bit-rate (kbit/s)	Coding time(s)
28	35.2	528.3	102.1	34.5	338.0	21.6
32	32.3	254.8	98.1	31.9	184.5	17.8
36	29.4	108.2	95.5	29.1	96.1	15.2
40	26.6	54.6	93.8	26.4	53.3	13.5

TABLE X.
TEST RESULTS OF "SNOW_ROAD" SEQUENCE

QP	Original algorithm			Proposed algorithm		
	PSNR -Y(dB)	Bitrate (kbit/s)	Coding time(s)	PSNR Y(dB)	Bit-rate (kbit/s)	Coding time(s)
28	34.4	164.7	102.4	34.3	159.8	56.0
32	31.8	76.2	98.3	31.7	75.4	50.7
36	29.4	40.5	95.1	29.4	40.5	46.7
40	27.3	23.0	92.3	27.3	22.9	43.6

TABLE XI.
TEST RESULTS OF “SNOW_GATE” SEQUENCE

QP	Original algorithm			Proposed algorithm		
	PSNR -Y(dB)	Bitrate (kbit/s)	Coding time(s)	PSNR Y(dB)	Bit-rate (kbit/s)	Coding time(s)
28	35.3	196.6	99.7	35.1	182.6	46.8
32	32.7	85.6	95.5	32.5	82.0	41.0
36	30.2	39.8	92.9	30.2	38.9	37.1
40	27.9	20.6	90.5	27.9	20.1	33.9

TABLE XII.
TEST RESULTS OF “SNOW_WAY” SEQUENCE

QP	Original algorithm			Proposed algorithm		
	PSNR -Y(dB)	Bitrate (kbit/s)	Coding time(s)	PSNR Y(dB)	Bit-rate (kbit/s)	Coding time(s)
28	35.0	112.3	98.3	34.8	103.2	31.2
32	32.3	49.5	95.0	32.2	47.0	26.4
36	29.7	25.9	92.6	29.7	25.3	22.9
40	27.4	15.1	90.4	27.4	14.9	20.5

TABLE XIII.
COMPREHENSIVE PERFORMANCE COMPARISON

Sequence	Resolution	Average ΔPSNR (dB)	Average ΔBitrate (%)	Time Saving (%)
hall_monitor	352×288	0.04	-1.16%	49.01%
bridge-far	352×288	0.12	-0.60%	79.51%
night	320×240	-0.04	0.70%	9.00%
fog	352×288	0.06	-1.68%	96.32%
rain	352×288	0.004	-0.16%	80.52%
dh_work	352×288	-0.04	0.87%	50.84%
dh_productline	352×288	0.61	-12.56%	82.61%
snow_road	352×288	-0.01	0.41%	49.34%
snow_gate	352×288	0.01	-0.37%	58.17%
snow_way	352×288	0.08	-2.11%	73.27%
Average		0.08	-1.67%	62.86%

To illustrate how the proposed fast algorithm could improve the coding performance on several sequences, we have added the curve of bit rate to support this point. Figure 6 and 7 intuitively show that the proposed algorithm require fewer bits than the algorithm of H.264/AVC. This is partly due to the fact that the proposed algorithm saves bit rate in background marcoblocks with noise by using more SKIP and P16×16 modes that require fewer bits than other modes. The results are also consistent with our previous analysis.

V. CONCLUSIONS

The video codec faces many challenges such as robustness in noise environment, computational complexity control and pre-processing and analysis of video data. This paper focuses on how to increase the noise robustness of video encoder. A novel macroblock mode pre-classification method based on co-matching criteria and motion vectors spatial and temporal filtering. By taking advantage of those improvements, the method could distinguish the noise macroblock from the background macroblock. The experiment results show that the proposed algorithm saves a significant part of computational time of over 62.86% and offers better coding efficiency in several sequences. And it also reduces the average Bjontegaard delta bit rate by about 1.67% and increases the average Bjontegaard delta peak signal-to-noise ratio by about 0.08dB. The algorithm is very relevant to real-time video system and also suitable for analysis of video data in compressed domain.

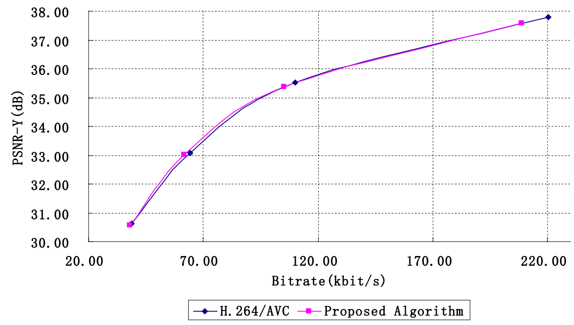

Figure 6. Comparison of Curve of PSNR vs. bitrate for “Hall_monitor” sequence.

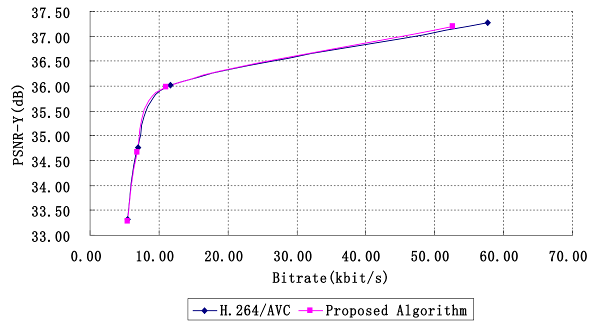

Figure 7. Comparison of Curve of PSNR vs. bitrate for “Bridge_far” sequence.

ACKNOWLEDGMENT

The authors wish to thank AVS organization for providing test sequences. This work is partly supported by National Natural Science Foundation of China under Grant # 61201268 to Cheng Zhou and Hubei Natural Science Foundation of China under Grant # 2010CDZ022 to Cheng Zhou.

REFERENCES

[1] JVT of ISO/IEC MPEG and ITU-T VCEG, “Draft ITU-T Recommendation, and Final Draft International Standard of Joint Video Specification,” ITU-T Rec. H.264, ISO/IEC 14496-10 AVC, 2003.

[2] X. Zhao, J. Sun, etc., “Novel statistical modeling, analysis and implementation of rate-distortion estimation for H.264/AVC coders,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol.20, no.5, pp.647-660, 2010.

[3] G. L. Wu, C. Y. Chen, etc., “Efficient spatial-temporal error concealment algorithm and hardware architecture design for H.264/AVC,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol.20, no.11, pp.1409-1422, 2010.
[4] Y. H. Sung and J. C. Wang, “Fast Mode Decision for H.264/AVC Based on Rate-Distortion Clustering,” *IEEE Transactions on Multimedia*, vol.14, no.3, pp. 693 – 702, 2012.
[5] T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra, “Overview of the H.264/AVC video coding standard,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol.13, no.7, pp.560-577, 2003.
[6] K.-P. Lim, G. Sullivan, and T. Wiegand, “Text description of joint model reference encoding methods and decoding concealment methods,” ISO/IEC MPEG and ITU-T VCEG Joint Video Team, Doc. JVT-N046, 2005.
[7] B. Jeon, J. Lee, “Fast mode decision for H.264,” In: ISO/IEC MPEG and ITU-T VCEG Joint Video Team, 2003, Doc. JVT-J033.
[8] S. Su, B. Zhou and X. Huang, “A novel fast inter mode decision algorithm for H.264/AVC,” *Journal of Multimedia*, vol.6, no.3, pp.285-292, 2011.
[9] M.S.U. Yusuf and M. Ahmad, “Relative SAD-based inter mode decision algorithm for H.264/AVC video standard,” *Journal of Multimedia*, vol.6 no.5, pp.425-435, 2011.
[10] K. Jong-Ho H. and K. Byung-Gyu Gyu, “Efficient intra-mode decision algorithm for inter-frames in H.264/AVC video coding,” *IET Image Processing*, vol.5, no.3, pp. 286-295, 2011.
[11] S. Tao, J. Zhou and Z. Hu, “Error resilient video transmission over wireless networks based on Wyner-Ziv coding of motion regions,” *Journal of Networks*, vol.4 no.9, pp.905-912, 2009.
[12] J. Chen, S. Su, etc., “An adaptive spatial-temporal error concealment scheme based on H.264/AVC,” *Journal of Multimedia*, vol.6, no.3, pp.293-299, 2011.
[13] M. E. Eduardo, J.M. Amaya, etc., “A two-level classification-based approach to inter mode decision in H.264/AVC,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol.21, no.11, pp.1719-1732, 2011.
[14] Y. Liu, L. Sun and S. Yang, “Wireless video transmission scheme based on virtual MIMO and network coding technology,” *Journal of Networks*, vol.6, no.1, pp.88-95, 2011.
[15] J. Jiang, Yang Liu, etc., “An improved selective encryption for H.264 video based on intra prediction mode scrambling,” *Journal of Multimedia*, vol.5, no.5, pp.464-472, Oct 2010.
[16] K. Lim, S. Kim, etc., “Fast block size and mode decision algorithm for intra prediction in H.264/AVC”, *IEEE Transactions on Consumer Electronics*, vol.58, no.2 pp. 654-660, 2012.
[17] X. Zhu and H. Fang. “Fast mode decision used in video surveillance system,” *Journal of Computer Applications*. vol.27, no.12, pp.3069-3071, 2007.
[18] Z. Yang, M. Li and D. Liu, “Fast inter prediction mode decision algorithm for AVS oriented to video surveillance,” *Application Research of Computers*, vol.25, no.8, pp.2407-2408, 2008,
[19] L. Sendur and I. W. Selesnick, “Bivariate shrinkage functions for wavelet-based denoising exploring interscale dependency,” *IEEE Trans. Signal Proc.* vol.50, no.11, pp. 2744-2756, 2002.
[20] H.B. Yin, X.Z. Fang, Z. Wei, and X.K.Yang, “An improved motion-compensated 3-D LLMMSE filter with spatio-temporal adaptive filtering support,” *IEEE Trans. Circuits Syst. Video Technol*, vol.17, no.12, pp. 1714-1727 ,2007.
[21] K. Hachicha, and P. Garda, “Noise-robustness improvement of the H.264 video coder,” *J. of Electronic Imaging*, vol.17, no.3, pp.033019 1-11, 2008.
[22] C. Zhou, C. Xiong and Y. Lu, “A noise robust mode decision algorithm based on macroblock-level motion detection for H.264/AVC,” *Physics Procedia (Elsevier)*, Vol.25, pp.1589-1594, 2012.
[23] K. Sühring, “H.264/AVC Reference Software Version JM10.2,” http://iphome.hhi.de/suehring/tml/download/, Joint Video Team, 2003.
[24] AVS test sequences. URL at http://www.avs.org.cn/
[25] G. Bjontegaard, “Calculation of average PSNR differences between RD-curves,” ISO/IEC MPEG and ITU-T VCEG Joint Video Team, Doc. VCEG-M33, 2001.

Cheng Zhou was born in Wuhan, April 2nd, 1979. He received the B.S. degree in electronics and information engineering from Hubei University, Wuhan, China, in 2001, and the M.S. and the Ph.D. degrees in pattern recognition and intelligent systems and in control science and control engineering from Huazhong University of Science and Technology, Wuhan, China, in 2007 and 2010, respectively.

He is currently an Instructor with the department of electronics and information engineering, South-Central University for Nationalities (SCUEC), Wuhan, China, and he has been engaged in the research of multimedia signal processing, particularly, video coding for MPEG-4, H.264, AVS (China) and HEVC standards since 2004.

Yanhong Lu was born in Hubei, April 30th, 1977. She received the B.S. degree in electronics and information engineering from South-Central University for Nationalities, Wuhan, China, in 1998, and the M.S. degree in communication and information systems from Huazhong University of Science and Technology, Wuhan, China, in 2003.

She is currently an Associate Professor with the department of electronics and information Engineering, Wuhan polytechnic, Wuhan, China. Her current research interests include image and video compression.

Chengyi Xiong received the B.S. degree in electronics engineering from the University of Electronic Science and Technology of China in 1992, the M.S. degree in electronics engineering from the SCUEC, Wuhan, China, in 2000, and the Ph.D. degree in pattern recognition and intelligent systems at Institute of Pattern Recognition and Artificial Intelligence, Huazhong University of Science and Technology, Wuhan, in 2006.

He is currently a Professor with SCUEC, China. His research interests mainly include digital signal and image processing, wavelet, image/video compression, and VLSI design.

Zhirong Gao received the B.S. degree in computer science from the SCUEC Wuhan, China, in 1997, the M.S. degree in computer science from the Huazhong University of Science and Technology, Wuhan, China in 2004.

She is currently an Associate Professor with SCUEC, China. Her research interests mainly include digital signal and image processing, image/video coding, and VLSI design.