Hindawi Publishing Corporation
EURASIP Journal on Image and Video Processing
Volume 2008, Article ID 393727, 14 pages
doi:10.1155/2008/393727

Research Article

Fast Macroblock Mode Selection Algorithm for Multiview Video Coding

Zongju Peng,[1,2] Gangyi Jiang,[1] Mei Yu,[1] and Qionghai Dai[3]

[1] *Faculty of Information Science and Engineering, Ningbo University, Ningbo 315211, China*
[2] *Institute of Computing Technology, Chinese Academy of Science, Beijing 100080, China*
[3] *Broadband Networks & Digital Media Lab, Tsinghua University, Beijing 100084, China*

Correspondence should be addressed to Gangyi Jiang, jianggangyi@126.com

Received 1 March 2008; Revised 7 August 2008; Accepted 14 October 2008

Recommended by Stefano Tubaro

Multiview video coding (MVC) plays an important role in three-dimensional video applications. Joint Video Team developed a joint multiview video model (JMVM) in which full-search algorithm is employed in macroblock mode selection to provide the best rate distortion performance for MVC. However, it results in a considerable increase in encoding complexity. We propose a hybrid fast macroblock mode selection algorithm after analyzing the full-search algorithm of JMVM. For nonanchor frames of the base view, the proposed algorithm halfway stops the macroblock mode search process by designing three dynamic thresholds. When nonanchor frames of the other views are being encoded, the macroblock modes can be predicted from the frames of the neighboring views due to the strong correlations of the macroblock modes. Experimental results show that the proposed hybrid fast macroblock mode selection algorithm promotes the encoding speed by 2.37 ~ 9.97 times without noticeable quality degradation compared with the JMVM.

1. INTRODUCTION

With the advancement in camera and display technologies, a wide variety of three-dimensional (3D) video applications, including free viewpoint video, free viewpoint television, 3D television, 3D telemedicine, 3D teleconference, and surveillance, are emerging. It has been widely recognized that multiview video coding (MVC) is one of the core technologies of 3D video applications [1–4]. The amount of multiview video data is tremendous because it is proportional to the number of cameras by which the multiple viewpoint video signals are captured simultaneously at different positions and angles. In order to transmit and store these signals for practical use, they must be effectively compressed.

The straightforward solution for MVC is to encode all the video signals independently by using state-of-the-art video codec such as H.264/AVC [5–7]. However, multiview video signals contain a large amount of inter-view dependencies, since all cameras capture the same scene from different viewpoints simultaneously [8]. Hence, various exquisitely designed view-temporal prediction structures, such as Hierarchical B Pictures (HBPs), KS_IPP, KS_PIP, and KS_IBP [9], are proposed. These structures efficiently exploit not only the temporal and spatial correlations within a single view, but also the inter-view correlations among different views. Kaup and Fecker analyzed the potential gains from the inter-view prediction [10]. Merkle et al. comparatively analyzed the rate distortion (RD) performances of these prediction structures [9, 11]. Flierl et al. *investigated the RD efficiency of motion and disparity-compensated coding for multiview video* [12, 13].

Standardization of MVC is investigated by *Joint Video Team* (JVT) formed by ISO/IEC MPEG and ITU-T VCEG. Currently, JVT is developing a joint multiview video model (JMVM), based on the video coding standard H.264/AVC [14]. The JMVM serves as a common platform to research on MVC, and uses HBP prediction structure to exploit both temporal and inter-view correlations. In the JMVM, different macroblock modes, including SKIP, Inter16 × 16, Inter16 × 8, Inter8 × 16, Inter8 × 8, Inter8 × 8Frext, Intra16 × 16, Intra8 × 8, and Intra4 × 4, are probed among all temporal and inter-view frames to decide the optimal macroblock mode so as to achieve the best RD performance. It is clear that adopting the full-search scheme to get the motion or disparity vector for each encoding macroblock mode in

each reference frame consumes considerable search time. According to statistics, the motion and disparity estimation consumes approximately 70% of the entire encoding time [15].

Hence, it is necessary to develop a fast algorithm to reduce computational complexity of MVC. The computational burden can be lessened by reducing the search frames or the times of macroblock mode matching. Some fast motion and disparity estimation algorithms for MVC have been proposed [16, 17]. In [16], Y. Kim et al. proposed a fast motion and disparity estimation algorithm to reduce the number of searching points by adaptively controlling a search range considering the reliability of each macroblock. In [17], Ding et al. proposed a new fast motion estimation algorithm which makes use of the coding information such as motion vector of the coded views.

In addition, fast macroblock mode selection algorithm can also be used to accelerate the encoding speed for MVC. Many fast macroblock mode selection algorithms for single-view video coding have been proposed [18–22]. In [18], Yin et al. proposed a coding scheme which jointly optimized motion estimation and mode decision. With this scheme, 85–90% complexity reduction can be achieved versus the H.264/AVC joint model with peak signal-to-noise ratio (PSNR) loss less than 0.2 dB and bit rate increase less than 3% regarding to common intermediate format (CIF) test sequences. In [19], Kuo and Chan proposed a fast macroblock mode selection algorithm in which the motion *field distribution and correlation within a macroblock are* taken into account. In [20], Kim and Kuo proposed a feature-based intra-/intermode decision algorithm. The algorithm decided the macroblock mode by the expected risk of choosing the wrong mode in a multidimensional simple feature space. It achieved a speedup factor of 20–32% without noticeable quality degradation. In [21], Choi et al. proposed a fast algorithm utilizing early SKIP mode decision and selective intramode decision. The algorithm reduced the entire encoding time by about 60% with negligible coding loss. In [22], Yin and Wang proposed a fast intermode selection algorithm. It reduced the encoding time of quarter CIF test sequences by 89.94% on average by making full use of the statistical feature and correlation in spatiotemporal domain.

The fast algorithms for single-view video coding cannot be used directly for MVC because the prediction structures for MVC are different from those of single-view video coding. In this paper, a hybrid fast macroblock mode selection algorithm is developed for MVC. Under the framework of the proposed algorithm, two methods are given to reduce computational complexity of macroblock mode selection in MVC with HBP prediction structure. The first method uses three dynamic thresholds to halfway stop the mode search process of the nonanchor frames in the base view. The second one which is originated from the inter-view and intraframe mode correlations is used for the nonanchor frames in the other views. Full-search algorithm, the same as the JMVM, is used for encoding anchor frames of all views to guarantee the RD performance. The experimental results show that the proposed algorithm promotes the encoding speed greatly without noticeable quality degradation compared with the JMVM.

This paper is organized as follows. Section 2 depicts the framework of the hybrid fast macroblock mode selection algorithm, including two fast mode selection methods for nonanchor frames of the base view and the other views, respectively. These two methods will be described in detail in Sections 3 and 4. Experimental results are given in Section 5 and the work is concluded in Section 6.

2. FRAMEWORK OF THE PROPOSED HYBRID FAST MACROBLOCK MODE SELECTION ALGORITHM

In JMVM, motion and disparity estimations are performed for each macroblock mode, and macroblock mode decision is made by comparing the RD cost of each mode. The mode with minimal RD cost is then selected as the best mode for interframe coding. The RD cost is calculated as

$$J(s,c,MODE \mid \lambda_{MODE}) = SSD(s,c,MODE \mid QP) + \lambda_{MODE} R(s,c,MODE \mid QP), \quad (1)$$

where s and c denote the source and reconstructed signals, respectively, and $MODE$ is the candidate macroblock mode. QP is the macroblock quantization parameter. λ_{MODE} is the Lagrange multiplier for mode decision and given by

$$\lambda_{MODE} = 0.85 \times 2^{QP/3}, \quad (2)$$

where $R(s,c,MODE \mid QP)$ reflects the number of bits produced for header(s) (including $MODE$ indicators), motion vector(s), and coefficients. $SSD(s,c,MODE \mid QP)$is the sum of square differences, which reflects the distortion between the original and reconstructed macroblocks and is calculated by

$$SSD(s,c,MODE \mid QP) = \sum_{i=1,j=1}^{B_1,B_2} \left| s[i,j] - c[i-v_x, j-v_y] \right|^2. \quad (3)$$

The full-search algorithm in the JMVM can obtain the best RD performance. Unfortunately, it consumes too much computational time. Based on the analysis of the macroblock mode selection process of the JMVM, a hybrid fast macroblock mode selection algorithm is proposed to lessen the computational burden.

In JMVM, HBP is used as prediction structure. Figure 1 shows an example of HBP prediction structure with eight views, where S_n denotes the individual view and T_n is the consecutive time instant. For example, S_0T_6 represents the frame locating at the 6th time instant in the view 0. The frames of all views, from T_0 to T_7, are the first group of pictures (GOPs) of the multiview video sequence. The GOP length, the number of frames along the temporal axis, is 8 in Figure 1. The horizontal and vertical arrows denote the inter-view and temporal referring relations, respectively. The frames the arrows point to are referred to by the other

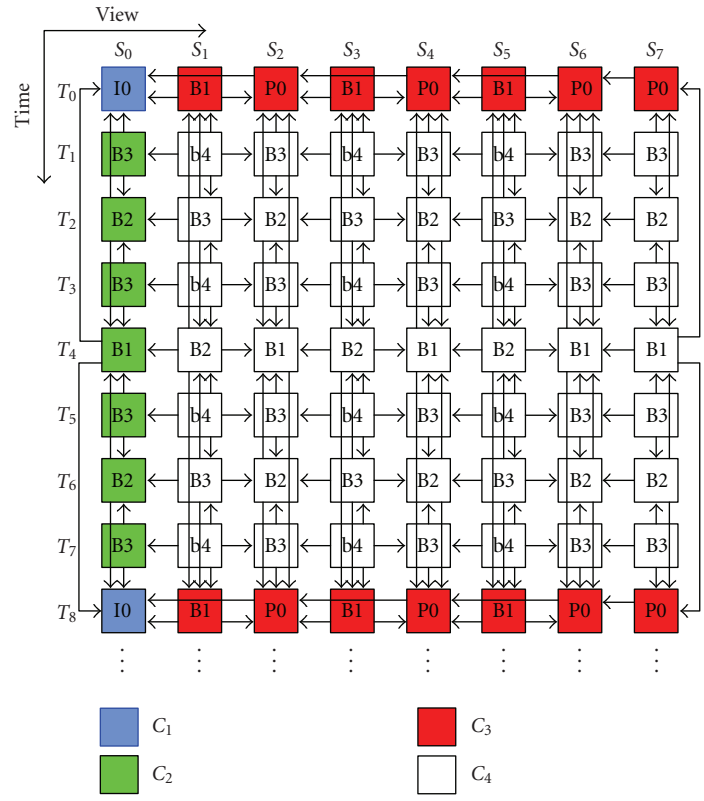


Figure 1: Illustration of frame classification in HBP prediction structure.

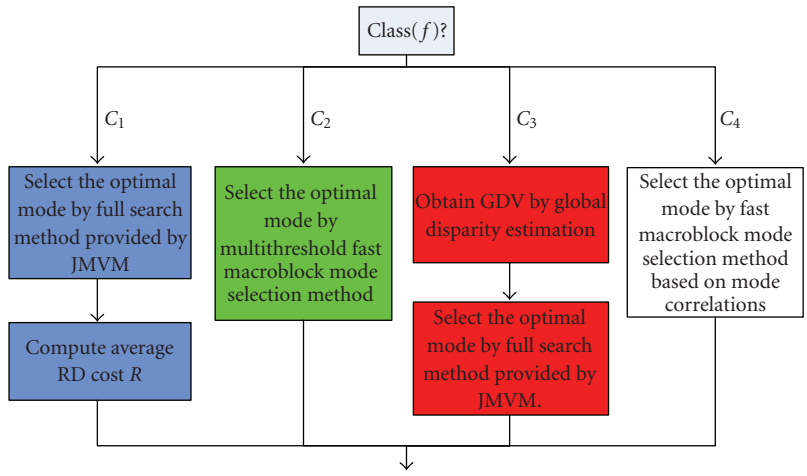


Figure 2: Block diagram of the hybrid fast macroblock mode selection algorithm.

frames. S_0 is the base view within which the frames do not have any inter-view reference frames. All the frames in HBP prediction structure are categorized into four types, that is, C_1, C_2, C_3, and C_4 shown by different colors in Figure 1. C_1 denotes the anchor frames in the base view without any reference frames, C_2 is nonanchor frames in the base view, C_3 and C_4 are the anchor frames and nonanchor frames in other views, respectively. For the GOP shown in Figure 1, the proportions of C_1, C_2, C_3, and C_4 are 1/64, 7/64, 7/64, and 49/64, respectively.

The block diagram of the proposed algorithm is given in Figure 2. Here, *class*(f) denotes the type of the current frame f. If *class*(f) is C_1 or C_3, the optimal macroblock mode is decided by full search which is the same as the JMVM. Since the frames with type C_1 or C_3 are located in high level in reference relationship, it is reasonable for these anchor frames performing the full search to keep the best RD performance. The average RD cost and the global disparity vector (GDV) are also obtained during the encoding process of anchor frames to implement the fast macroblock mode

Table 1: Statistical results of macroblock modes in the frame S_0T_6 of Ballroom.

Category	{SKIP}	{Inter16 × 16}	{Inter16 × 8}	{Inter8 × 16}	{Inter8 × 8, Inter8 × 8Frext}	Other modes
$P(Category)$	3531.08	5875.08	8275.58	8999.61	14675.85	12983.78
$N(Category)$	65.75%	14.33%	5.33%	5.83%	3.83%	4.93%

selection methods of nonanchor frames. If $class(f)$ is C_2 or C_4, the macroblock mode will be selected by the fast macroblock mode selection methods which will be discussed in detail in Sections 3 and 4, respectively.

3. MULTITHRESHOLD FAST MACROBLOCK MODE SELECTION METHOD

In this section, we investigate the full-search process of macroblock mode selection under the JMVM firstly, and find some regularity in the macroblock mode distribution and RD cost of various macroblock modes. Then, a fast macroblock mode selection method for C_2 frames is given and analyzed theoretically in terms of RD performance. Finally, a dynamical updating method of multiple thresholds is devised.

3.1. Analyses of macroblock mode selection process of the JMVM

Before designing the fast macroblock mode selection method, we selected the frame S_0T_6 of Ballroom test sequence provided by Mitsubishi Electric Research Laboratories (MERL, Mass, USA) to investigate the full search method of the JMVM. During the encoding process of the frame, the optimal mode and the RD cost of each traversed mode of every macroblock are recorded. According to the proportion of macroblock mode and RD cost, we find that there are some statistical features in macroblock mode selection. In order to analyze them, two variables $N(M)$ and $P(M)$ are defined to represent proportion and average RD cost of the macroblock mode category M, respectively. They are calculated by

$$N(M) = \frac{\sum_{g=1}^{H\times V}\phi(g,M)}{H\times V}, \qquad \phi(g,M) = \begin{cases} 0, & m \notin M, \\ 1, & m \in M, \end{cases} \tag{4}$$

$$P(M) = \frac{\sum_{g=1}^{H\times V}(\phi(g,M)\times \mathrm{Rd}(g,m))}{\sum_{g=1}^{H\times V}\phi(g,M)}, \qquad \sum_{g=1}^{H\times V}\phi(g,M) \neq 0, \quad \phi(g,M) = \begin{cases} 0, & m \notin M, \\ 1, & m \in M, \end{cases} \tag{5}$$

where H and V are the numbers of macroblocks in horizontal and vertical directions of a frame, respectively, m is the optimal mode, and Rd(g,m) denotes the minimal RD cost of the gth macrobock.

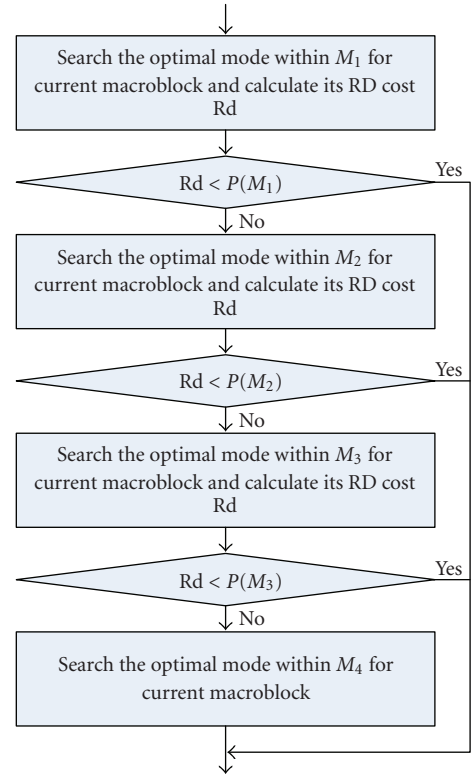


Figure 3: Flowchart of multithreshold fast mode selection method.

Table 1 tabulates the statistical results of different macrobock modes in the frame S_0T_6 of Ballroom. Obviously, it is not balanced in proportions of macroblock modes as well as the average RD cost. Most of the macroblocks are encoded with SKIP mode whose average RD cost is the smallest within all macroblock modes. Next to the SKIP mode, Inter16 × 16 mode ranks the second in the proportion and its average RD cost is also the second smallest. The macroblock numbers of Inter16 × 8 and Inter8 × 16 are nearly equivalent. They are less than the SKIP and Inter16 × 16 in quantity, and larger in average RD cost. The other modes, such as Inter8 × 8, Inter8 × 8Frext, Intra16, Intra8, and Intra4, occupying the smallest quantity in a frame, rank highest in the average RD cost. However, these modes are indispensable to MVC. Other test sequences also have similar statistical features [23].

3.2. Multithreshold fast macroblock mode selection method

Based on the analyses above, we divide the macroblock modes into four categories, {SKIP}, {Inter16 × 16}, {Inter16 ×8, Inter8×16}, and {Inter8×8, Inter8×8Frext, Intra16×16,

Intra8 × 8, Intra4 × 4}, denoted by M_1, M_2, M_3, and M_4, respectively.

As tabulated in Table 1, there are great gaps between $P(M_1), P(M_2), P(M_3)$, and $P(M_4)$. They can be utilized to build multiple threshold conditions to halfway stop the process of macroblock mode selection if they are known in advance. Figure 3 illustrates the detailed flowchart of the multithreshold fast macroblock mode selection method. When one macroblock is encoded, the SKIP mode is probed first. If its RD cost is smaller than $P(M_1)$, the mode selection process is halfway stopped and the SKIP mode is selected as the optimal mode. Otherwise, the Inter16 × 16 is tested, and the mode selection process will be ended immediately as the RD cost is smaller than $P(M_2)$. If RD cost is not smaller than $P(M_2)$, the modes in M_3 are traversed one by one, the mode selection is halfway stopped when the RD cost is smaller than $P(M_3)$. If RD cost is not smaller than $P(M_3)$, the modes in M_4 will be searched to find the optimal macrobock mode.

The fast macroblock mode selection method may result in degradation of RD performance because not all macroblocks select optimal modes. The error in mode selection influences not only the RD performance of the current frame, but also that of the frames which refer the current frame directly or indirectly. For one macroblock, suppose the optimal mode selected through the full-search algorithm belongs to M_i while the optimal mode belongs to M_j under the fast mode selection method. If error mode selection happens, it must satisfy the following conditions:

(1) $j < i$,

(2) $\mathrm{Rd}(k, m_i) < \mathrm{Rd}(k, m_j)$,

(3) $\mathrm{Rd}(k, m_l) \geq P(M_l)$, $1 \leq l < j$

In the conditions (2) and (3), m_i, m_j, and m_l are the modes with the least RD cost in M_i, M_j, and M_l, respectively.

The conditions above are about individual macroblock. But for investigating the RD performance, it is important to statistically analyze the error mode selection of all macroblocks of one frame. To estimate the error selection probability of one frame to be encoded, we define a parameter K to express the probability of error mode selection as follows:

$$K = \sum_{g=1}^{4} (\mu_g \times N(M_g)), \tag{6}$$

where $N(M_g)$ and μ_g denote the proportion of macroblock modes and the probability of error mode selection regarding M_g. K should be very small because μ_g is limited by the strict condition listed above. Specially, μ_1 must be zero owing to the condition (1) of the error mode selection. In other words, if M_i equals to M_1, M_j must be M_i and no error selection happens.

Based on the data of the frame S_0T_6 recorded in detail under the JMVM, the thresholds of the multithreshold fast macroblock mode selection method can be estimated. Then, the error selection macroblocks can be filtered out according to the thresholds and their RD costs of all macroblock modes. Figure 4 shows such macroblocks and their increments in RD cost. Each vertical line reflects an RD cost increment caused by an error mode selection. The X-axis represents the number of macroblock. The X-axis of the left line segment is 103. That is to say, the 103rd macroblock is the first one that selected the error macroblock mode. The Y-axes of the upper endpoint and the lower endpoint of the vertical line are the RD costs of the modes selected by the proposed method and the full-search algorithm, respectively. Among 1200 macroblocks in the test frame, only 64 macroblocks select error macroblock mode, namely, $K = 5.33\%$. The average RD cost of all macroblock mode categories $P(M_1 \cup M_2 \cup M_3 \cup M_4)$ with respect to the proposed method only rises 0.29% compared with the full-search algorithm. The degradation in RD performance brought by these error-selected modes can almost be ignored.

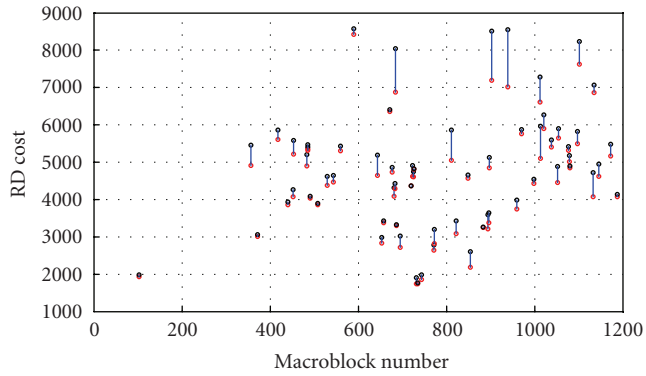


Figure 4: Increments in RD cost resulted from error macroblock mode selection in the frame S_0T_6 of Ballroom.

3.3. Dynamically update the thresholds

The multithreshold fast macroblock mode selection method described in Section 3.2 is based on the hypothesis that the thresholds were already known. Therefore, it is vital to design a feasible threshold computing method. We found that $P(M_i)$ is approximately linear with Lagrange multiplier and average RD cost of all mode categories of the current frame after a lot of experiments and careful observations. Figures 5 and 6 show the approximately linear relationships, where L and R are the Lagrange multiplier and $P(M_1 \cup M_2 \cup M_3 \cup M_4)$. So, the thresholds in the proposed fast macroblock mode selection method can be calculated theoretically by

$$P(M_i) \approx a_i \times L + b_i \times R + c_i \quad (i = 1, 2, 3), \tag{7}$$

where a_i, b_i, and c_i are the parameters of the approximately linear functions. However, it is difficult to calculate the thresholds from (7) directly because R is calculated on the assumption of the current frame having been encoded. Thus, this is a deadlock. In the implementation of the proposed method, the average RD cost of the current frame is estimated approximately from the RD cost of anchor frames in the same GOP. The average RD costs of the nonanchor frames are nearly equivalent owing to the temporal correlation. Unfortunately, the average RD cost of the anchor frames is larger than that of the nonanchor

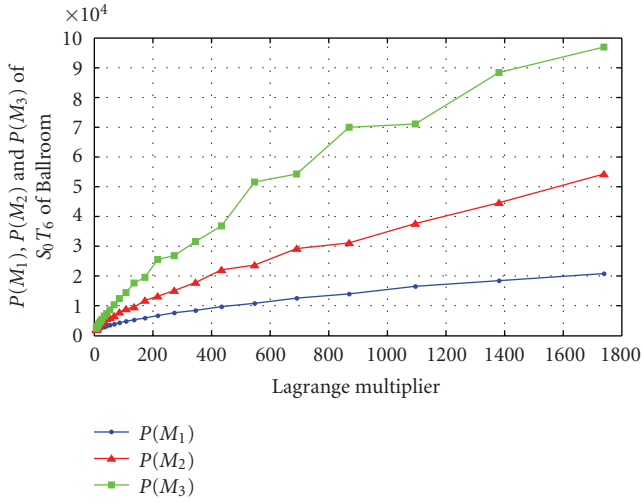


Figure 5: Illustration of approximately linear relationship between $P(M_i)$ and L.

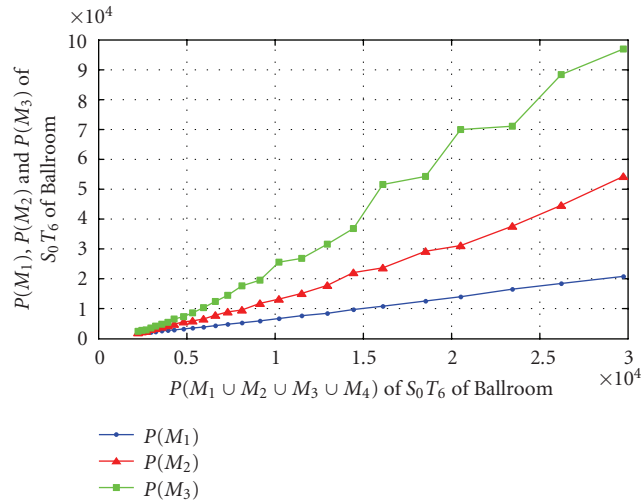


Figure 6: Illustration of approximately linear relationship between $P(M_i)$ and R.

frames, since they are intraframe encoded. In Figure 7, the average RD costs of the anchor frames of Ballroom, whose picture-order counts (POCs) are 0, 12, and 24, are more than 7500. By contrast, the average RD costs of the nonanchor frames are about 5050. Figure 8 also shows the difference of average RD costs between the anchor frames and the nonanchor frames of Exit test sequence. So, (7) is revised as

$$P(M_i) = a_i' \times L + b_i' \times R' + c_i' \quad (i = 1, 2, 3), \tag{8}$$

where L and R' are the Lagrange multiplier and the average RD cost of the anchor frames. $P(M_i)$ is mainly contributed by $b_i' \times R'$ because L is smaller than R' by 10–100 times, while $a_i' \times L + c_i'$ can be used to slightly adjust the threshold $P(M_i)$. In the proposed method, a_i', b_i', and c_i' are set as follows:

$$\begin{aligned} a_1' &= -5, \quad a_2' = -5, \quad a_3' = 30, \\ b_1' &= 0.55, \quad b_2' = 0.80, \quad b_3' = 0.95, \\ c_1' &= 0, \quad c_2' = 0, \quad c_3' = 0. \end{aligned} \tag{9}$$

These parameters are obtained from a large number of experiments. They are suitable for various multiview video sequences. So far, the thresholds are calculated and updated dynamically by the method illustrated in Figure 9. The figure is also a detailed description of C_1 and C_2 subbranch of the block diagram in Figure 2. The method is summarized as follows.

Step 1. Check whether the current frame is an anchor frame or not. If it is an anchor frame then go to Step 2, otherwise, go to Step 3.

Step 2. Select the optimal mode by full search that is the same as JMVM. Then, calculate the average RD cost of all the macroblocks by (5).

Step 3. Calculate $P(M_i)$ by (8), select the optimal mode based on the multithreshold fast macroblock mode selection method.

4. FAST MACROBLOCK MODE SELECTION METHOD BASED ON INTER-VIEW MODE CORRELATIONS

After encoding the base view by the multithreshold fast macroblock mode selection method, the other views are to be dealt with one by one according to the HBP prediction structure. The correlations between two neighboring views may result in strong mode correlations between the current frame and the frames in the neighboring views at the same instant. When the frame with type C_4 is encoded, the mode of the current macroblock may be estimated accurately via macroblock mode correlations. Thus, the mode selection process can be accelerated by making use of mode prediction.

4.1. *Fast macroblock mode selection method based on mode correlation*

The spatial correlation between neighboring views may lead to strong mode correlation. In order to verify the phenomenon, S_0T_6 and S_2T_6 of Ballroom and Exit test sequence, S_0T_7 and S_2T_7 of Race1 test sequence are investigated according to the MVC common test conditions [24]. Exit and Race1 are provided by MERL and KDDI (Japan), respectively. After recording all the macroblock modes of the frames under the full-search algorithm of the JMVM, we draw the macroblock mode distribution maps illustrated by Figures 10, 11, and 12. In these figures, the blocks with red, green, and blue borders denote the macroblocks encoded with the SKIP, Inter- and Intramodes, respectively. It is obvious that the macroblock modes are similar between the frame pairs.

The mode correlation is verified by mode similarity. Because of the mode similarity, the macroblock modes of the encoded frames at the same instant of the neighboring views can be used to predict the modes of the current frame. For example, the HBP structure in Figure 1 has predictive relationships, including view 0 → view 2, view 2 → view 4, view 4 → view 6, view 6 → view 7, view 0 → view 1,

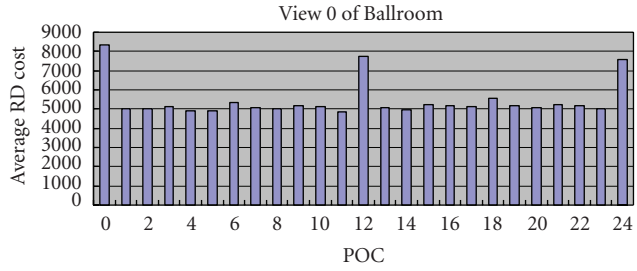


Figure 7: Average RD cost of the frames in Ballroom test sequence.

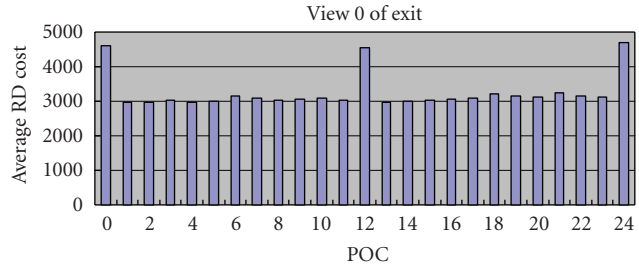


Figure 8: Average RD cost of the frames in Exit test sequence.

view 2 → view 1, view 2 → view 3, view 4 → view 3, view 4 → view 5, and view 6 → view 5, where view i → view j denotes that the macroblock modes of view i are predictive modes of view j. So, multiview video signals are processed more quickly in the order of view 0, view 2, view 1, view 4, view 3, view 6, view 5, and view 7 due to mode similarity.

As for a frame with type C_4, it is unnecessary for the encoder to perform a full search since at least one frame at the same instant of the neighboring views has been encoded. The encoding time can be significantly reduced by only searching the macroblock mode of the corresponding macroblock in the neighboring coded views if specific RD condition is satisfied. The location of the corresponding macroblock can be decided by GDV between the current frame and the frame of the neighboring view. The GDV is measured by the macroblock size of units, and it can be deduced based on Koo's method that has been integrated into the JMVM [25]. GDV is estimated in every anchor picture, and interpolated for nonanchor frames.

As shown in Figure 13, $\mathrm{GDV}_{\mathrm{cur}}$ denotes the location of the corresponding macroblock in the neighboring views on a certain POC. It is derived by

$$\mathrm{GDV}_{\mathrm{cur}} = \mathrm{GDV}_{\mathrm{ahead}} + \left\lfloor \frac{\mathrm{POC}_{\mathrm{cur}} - \mathrm{POC}_{\mathrm{ahead}}}{\mathrm{POC}_{\mathrm{behind}} - \mathrm{POC}_{\mathrm{ahead}}} \times \left(\mathrm{GDV}_{\mathrm{behind}} - \mathrm{GDV}_{\mathrm{ahead}}\right) \right\rfloor, \tag{10}$$

where $\mathrm{GDV}_{\mathrm{ahead}}$ and $\mathrm{GDV}_{\mathrm{behind}}$ are two latest GDVs of anchor frames. $\mathrm{POC}_{\mathrm{cur}}$, $\mathrm{POC}_{\mathrm{ahead}}$, and $\mathrm{POC}_{\mathrm{behind}}$ are POCs along temporal axis.

If the mode of the corresponding macroblock in the frame at the neighboring view is used directly to predict the mode of the current macroblock, the computational complexity is greatly reduced. However, the RD performance may be degraded due to the following reasons.

(1) The global disparity is not the exact disparity between the current macroblock and the corresponding one. There is a deviation between the global disparity and the pixel-wise disparity.

(2) The inter-view mode similarity degree varies from region to region. For background or stationary regions, the macroblock mode in the current view is more similar to that of the neighboring views compared with the foreground or motion regions.

In order to eliminate the ill effects caused by the inaccurate disparity and content dissimilarity, the modes of the corresponding macroblock and its surrounding macroblocks are searched in a nonrepetitive way. For the convenience of narration, we call the macroblocks surrounding the corresponding macroblock in the frame at the same instant in the neighboring view as the corresponding neighboring macroblocks (CNMs). The locations of the current macroblock, the corresponding macroblock, and the CNMs are shown in Figure 13. The proposed method is summarized according to above analyses and depicted by Figure 14. An RD cost will be obtained after the searching operation of the corresponding macroblock and the CNMs in a nonrepetitive way. If the RD cost is smaller than a threshold, the searching process will be stopped immediately. The threshold in the proposed method is determined by the RD cost of the corresponding macroblock and an experimental constant β. The threshold is adopted to identify the macroblocks that cannot be predicted accurately because these macroblocks are usually located in the motion regions and their RD costs often change drastically. Let E_{RD} be the RD cost of the corresponding macroblock, if the RD cost is greater than the threshold $\beta \times E_{\mathrm{RD}}$, the full-search method is used because of the high risk of error mode selection. In the implementation of the proposed method, β is set as 2 empirically.

4.2. Analyses on macroblock mode correlations

Mode correlations are the basis of the method proposed above. It is clear that the performance of the proposed method is determined by two factors. The first is the degree of the inter-view mode similarity between the current frame and the view-neighboring frames, and the second is the degree of the mode aggregation of the view-neighboring frame. The first factor affects the accuracy of the mode prediction while the second reflects the macroblock searching times. We call the inter-view mode similarity and the mode aggregation as inter-view mode correlation and intraframe mode correlation, respectively.

Quantitative analyses on the mode correlations are helpful to understand the validity of the proposed method. In the following, we take S_0T_6 and S_2T_6 of Ballroom as an example to investigate the mode correlations. Let S_2T_6 be the current encoding frame, S_0T_6 be the view-neighboring coded frame. The horizontal and vertical components of the GDV between S_0T_6 and S_2T_6 are 2 and 0, respectively. So, the overlapping

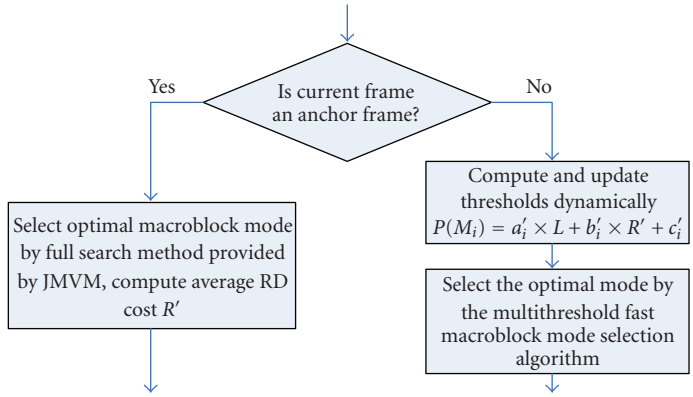


Figure 9: Flowchart of computing and updating multithresholds.

regions in the frames S_0T_6 and S_2T_6 are marked with black borders in Figures 15 and 16. The macroblock modes in Figures 15 and 16 are the optimal modes decided by the full-search method. In order to evaluate the accuracy of the mode prediction, we use the macroblock modes in Figure 15 as the reference of S_2T_6, and the optimal macroblock modes in Figure 16 as the target which the fast macroblock mode selection method described in this section tries to achieve. In the overlapping region of S_2T_6, most of the macroblocks have one corresponding macroblock and eight CNMs. However, the macroblocks, located in the top row, the bottom row, and the right column of the overlapping region, have one corresponding macroblock and different numbers of CNMs. The quantitative variation of CNMs results in difficulty in analyzing the mode correlations. We employed two ways to simplify the discussions.

(1) We only concern the macroblocks of the current frame which have one corresponding macroblock and eight CNMs. They are the macroblocks in the overlapping region of S_2T_6, excluding the top row, the bottom row, and the right column. The number of such macroblocks is 1036.

(2) Slightly different from the multithreshold fast macroblock mode selection method, we divide the macroblock modes into six classes here as follows:

   (a) SKIP;
   (b) Inter16 × 16;
   (c) Inter16 × 8;
   (d) Inter8 × 16;
   (e) Inter8 × 8 and Inter8 × 8Frext;
   (f) Intra16 × 16, Intra4 × 4 and Intra8 × 8.

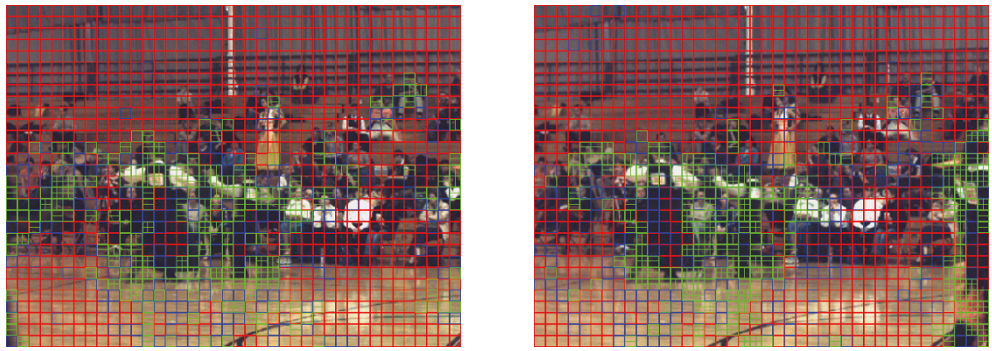
Let (x, y) denote the coordinate of the current macroblock, $g(x, y)$ denote the number of macroblocks among the corresponding macroblock and the eight CNMs which are encoded with the same macroblock mode class as the current macroblock selects, and $h(x, y)$ depict the count of macroblock mode classes of the corresponding macroblock and the CNMs. Then, $f(x, y)$ and $s(x, y)$, representing the inter-view mode correlation of the current macroblock and the intraframe mode correlation of the corresponding macroblock and the CNMs, can be estimated by

$$f(x, y) = \frac{g(x, y)}{9}, \qquad s(x, y) = h(x, y). \tag{11}$$

The bigger $f(x, y)$ is the stronger the inter-view mode correlation is. However, the intraframe mode correlation decreases as $s(x, y)$ rises. Comparing the macroblock modes of the frame S_0T_6 with that of the frame S_2T_6 in Ballroom, it is obvious that both the inter-view mode correlation and the intraframe mode correlation of the background regions are higher than those of the motion regions. The macroblocks at positions (4, 5) and (19, 16) in Figure 16 are located in background and motion regions, respectively. They correspond to the macroblocks at (6, 5) and (21, 16) in Figure 15 according to the GDV. Modes of the macroblock at (4, 5) in S_2T_6, the corresponding macroblock at (6, 5) in S_0T_6, and the eight CNMs are SKIP. So, $g(4, 5) = 9$, $h(4, 5) = 1$, $f(4, 5) = 1$, $s(4, 5) = 1$, and $f(4, 5) = 1$ means that modes of current macroblock, the corresponding macroblock, and the CNMs all belong to the same class. In other words, each mode of the corresponding macroblock and CNMs in S_0T_6 can be used to accurately predict the mode of the current macroblock at (4, 5) in S_2T_6. $s(4, 5) = 1$ indicates that the modes of the corresponding macroblock and the CNMs belong to the same class, so that only searching the modes in one class is enough to obtain the optimal mode for the macroblock at (4, 5) in S_2T_6. As far as the macroblock at (19, 16) is concerned, $f(19, 16) = 0.22$, $s(19, 16) = 4$. Compared with the macroblock in background, the inter-view mode correlation becomes weak and more macroblock modes should be traversed in motion regions.

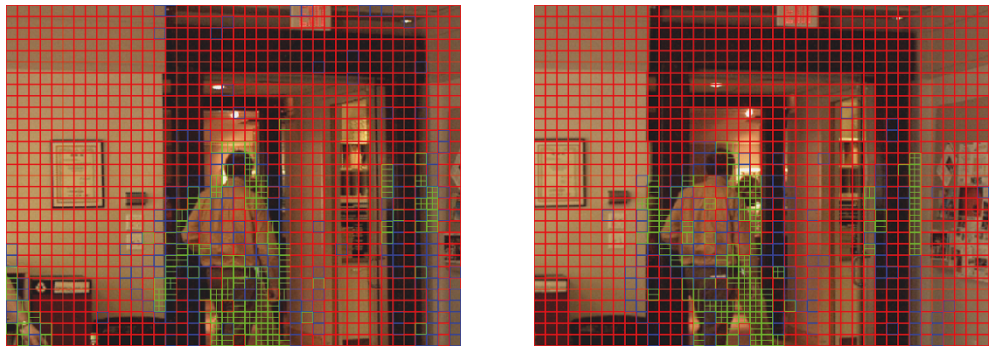
4.3. Discussions on performance of the proposed method

The statistical results of the mode correlations affect the integral performance of the proposed method. Figure 17 shows the inter-view mode correlations between the current macroblocks in S_2T_6 and their corresponding macroblocks



(a) S_0T_6 (b) S_2T_6

Figure 10: Macroblock mode distribution of S_0T_6 and S_2T_6 in Ballroom test sequence.



(a) S_0T_6 (b) S_2T_6

Figure 11: Macroblock mode distribution of S_0T_6 and S_2T_6 in Exit test sequence.

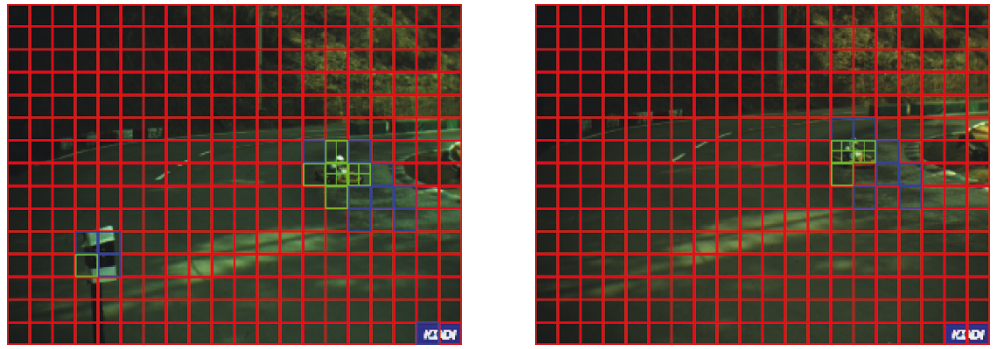
in S_0T_6. Most of the macroblocks in the background are with $f(x, y) = 1$. According to our statistical results, only few macroblocks are completely irrelevant to their corresponding macroblocks and the CNMs in the neighboring view. The average inter-view mode correlation in the overlapping region (exclude the macroblocks in the upper row, the lower row, and the right column) of S_2T_6 amounts to 0.60, which means that $g(x, y)$ of S_2T_6 equals to 5.40 on average. Therefore, most of the macroblock modes of S_0T_6 can be used to predict the macroblock modes of S_2T_6. As long as $f(x, y) > 0$, the optimal mode of macroblock (x, y) can be predicted accurately from the corresponding macroblock and the CNMs. So, the ratio of accurate prediction is even higher than the average inter-view mode correlation. Compared with average inter-view mode correlation with 0.60, the ratio of accurate prediction is up to 91.51% in the same region. Thus, the mode prediction of the proposed method is effective in mode decision. Figure 18 shows the intraframe mode correlation. Similar to the inter-view mode correlation, most of macroblocks in the background are with $s(x, y) = 1$ while $s(x, y)$ is up to 6 for some macroblocks in the motion regions. In general, more macroblock modes should be searched to obtain the optimal mode in motion regions.

Table 2 tabulates the statistical results of mode correlations. Every cell in the table gives the macroblock number/percentage under a specific inter-view and intraframe mode correlation condition. For example, there are 357 macroblocks with $s(x, y) = 1$, in which 336 macroblocks are with $f(x, y) > 0$ and the rest 21 macroblocks with $f(x, y) = 0$. If these macroblocks are encoded by the proposed method, the theoretical times of mode searching are $336 \times 1 + 21 \times 6 = 562$ while $357 \times 6 = 2142$ times are needed for the full-search algorithm. According to all intraframe mode correlation results listed in Table 2, the macroblock mode selection method based on the mode correlation can reduce mode searching times by 2.07 times. The practical speedup ratio is even higher because of the large-scale distribution of the SKIP mode and its ignorable processing time.

5. EXPERIMENTAL RESULTS AND ANALYSIS

To evaluate the performance of the proposed hybrid fast macroblock mode selection algorithm, the experiments are performed complying with the common test conditions for MVC [24]. The detailed parameters and test conditions are listed in Table 3. Figures 19(a) and 19(f) show the first frame in each view of the test sequences. All tests in the experiment are run on the Intel Xeon 3.2 GHz with 12 GB RAM and the OS is Microsoft Windows Server 2003.

Table 4 shows experimental results of encoding time in which *TS* indicates the average time saving in coding process



(a) S_0T_7 (b) S_2T_7

Figure 12: Macroblock mode distribution of S_0T_7 and S_2T_7 in Race1 test sequence.

Table 2: Statistical results of mode correlation.

	$s(x,y)=1$	$s(x,y)=2$	$s(x,y)=3$	$s(x,y)=4$	$s(x,y)=5$	$s(x,y)=6$	Total
$f(x,y)>0$	336/32.43%	157/15.15%	150/14.48%	167/16.12	121/11.68%	17/1.64%	948/91.51%
$f(x,y)=0$	21/2.03%	10/0.97%	16/1.54%	31/2.99%	10/0.97%	0/0.00%	88/8.49%
Total	357/34.46%	167/16.12%	166/16.02%	198/19.11%	131/12.64%	17/1.64%	1036/100%

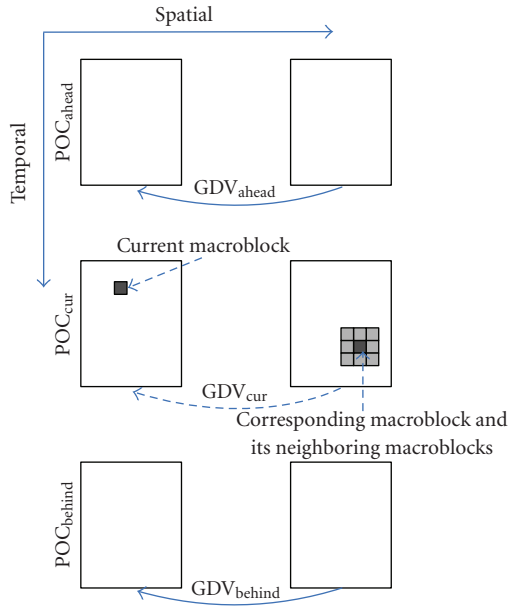


Figure 13: Illustration of interpolation method of GDV.

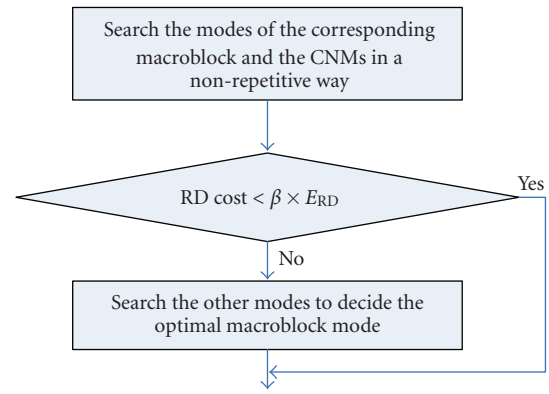


Figure 14: Illustration of fast macroblock mode selection method based on mode correlation.

and it is defined by

$$TS = \frac{T_{\text{JMVM}} - T_{\text{proposed}}}{T_{\text{JMVM}}} \times 100\,[\%], \tag{12}$$

where T_{JMVM} and T_{proposed} are the encoding time of the JMVM and its modified software according to the proposed hybrid algorithm, respectively. Table 4 shows the speedup performance of the proposed two fast macroblock selection methods, respectively. In view 0, the multithreshold fast macroblock mode selection method significantly reduces the encoding time, ranging from 43.10% to 90.27%. In other views, 57.95%–90.76% of the encoding time is saved by fast macroblock mode selection method based on inter-view mode correlations. The real speedup performance of the proposed fast macroblock mode selection methods may be better because the data listed in Table 4 include the encoding time of the anchor frames which the full-search method is adopted. Figure 20 shows the encoding time comparison between the JMVM and the proposed hybrid fast macroblock mode selection algorithm. The total encoding speed is prompted by 2.37–9.97 times.

Table 5 shows the RD performances of the proposed hybrid fast macroblock mode selection algorithm. Every cell shows an average PSNR Y and bit rate of a test sequence with respect to a certain basis QP. Compared with the JMVM, PSNR Y of view 0 decreases less than 0.03 dB and the bit rate nearly keeps the same when the proposed hybrid algorithm is implemented. Similar to view 0, the PSNR Y is decreased by 0.01–0.08 dB, and the bit rate increases a little, or decreases occasionally in views 1–7.

Table 3: Test conditions.

Encoder	Search range	Basis QP	DeltaLayerXQuant						Prediction structure
JMVM 4	±64	22, 27,	0	1	2	3	4	5	HBP
		32, 37	0	3	4	5	6	7	

Test sequence		Resolution	Features	Camera space(cm)	Frame rate(fps)	Property of camera array	GOP length	Encoded frames
MERL	Ballroom	640 × 480	Great disparity and violent motion	19.5	25	1D/parallel	12	49
MERL	Exit	640 × 480	Great disparity	19.5	25	1D/parallel	12	49
KDDI	Race1	320 × 240	Violent motion	20	30	1D/parallel	15	61
Microsoft	Breakdancers	1024 × 768	Violent motion	20	15	1D/arc	15	46
Microsoft	Ballet	1024 × 768	Great disparity and violent motion	20	15	1D/arc	15	46
HHI	Alt Moabit	1024 × 768	Outdoor scene	6.5	16.67	1D/parallel	15	46

Table 4: Speedup performance comparison between the JMVM and the proposed algorithm.

Sequences		Ballroom			Exit			Race1		
View	QP	T_{JMVM} [s]	T_{proposed} [s]	*TS* [%]	T_{JMVM} [s]	T_{proposed} [s]	*TS* [%]	T_{JMVM} [s]	T_{proposed} [s]	*TS* [%]
0	22	696.63	309.63	55.55	530.58	204.72	61.42	136.50	13.28	90.27
	27	663.13	245.13	63.03	488.77	134.41	72.50	135.25	16.64	87.70
	32	632.06	200.52	68.28	469.03	106.64	77.26	133.88	17.36	87.03
	37	600.41	183.94	69.36	454.95	96.80	78.72	132.77	18.74	85.89
1–7	22	7960.74	3347.81	57.95	7519.10	2462.09	67.26	1917.17	331.17	82.73
	27	7511.86	2709.00	63.93	6978.16	1970.25	71.77	1817.22	325.78	82.07
	32	7056.66	2214.78	68.61	6592.98	1746.74	73.50	1709.58	327.14	80.86
	37	6562.53	1894.16	71.14	6204.53	1638.94	73.58	1594.48	320.02	79.93
Sequences		Breakdancers			Ballet			Alt Moabit		
View	QP	T_{JMVM} [s]	T_{proposed} [s]	*TS* [%]	T_{JMVM} [s]	T_{proposed} [s]	*TS* [%]	T_{JMVM} [s]	T_{proposed} [s]	*TS* [%]
0	22	2273.11	1293.36	43.10	1384.50	484.67	64.99	1337.24	321.63	75.95
	27	2000.3	941.72	52.92	1312.30	375.02	71.42	1296.24	254.17	80.39
	32	1797.0	737.17	58.98	1254.20	327.17	73.91	1262.75	230.72	81.73
	37	1620.11	698.83	56.87	1197.19	297.61	75.14	1224.48	218.67	82.14
1–7	22	19975.47	5193.25	74.00	16885.28	3266.44	80.66	13817.16	1604.19	88.39
	27	17785.75	3650.03	79.48	15591.39	2947.91	81.09	13127.58	1312.80	90.00
	32	16064.52	3771.36	76.52	14389.78	2937.91	79.58	12577.98	1196.71	90.49
	37	14633.30	3504.45	76.05	13309.55	2959.20	77.77	12090.08	1116.85	90.76

The speedup performances of the proposed hybrid algorithm are different for test sequences owing to their different features. Compared with other test sequences, the mode distribution in Race1 is severe imbalance, and most of the macroblocks selected SKIP as their optimal mode under the JMVM [23]. According to (6), the risk of error mode selection is very small in this case. Therefore, the performance of multithreshold fast macroblock mode selection method is more effective for Race1. Additionally, the intraframe mode correlation is higher because the modes are more aggregative, so higher speedup performance is gained under the fast macroblock mode selection method based on the mode correlation. Thus, the proposed hybrid algorithm is more effective for Race1 in terms of the encoding time and the RD performance. As for Alt Moabit test sequence, the camera space is narrow so that the GDV is small and the overlapping region occupies large region of the frame, which means that more macroblocks can be predicted from the neighboring views. Therefore, the higher speedup performance is also achieved for Alt Moabit.

6. CONCLUSION AND FUTURE WORK

MVC is one of the core technologies in 3D video applications. It is essential to design a fast macroblock mode selection algorithm to reduce the complexity of MVC.

Table 5: RD performance comparison between the JMVM and the proposed algorithm [dB/Kbps].

Sequences		Ballroom		Exit		Race1	
View	QP	JMVM	Proposed	JMVM	Proposed	JMVM	Proposed
0	22	39.48/1650.24	39.45/1647.53	40.30/808.32	40.28/806.78	43.15/181.57	43.13/181.34
	27	37.23/886.16	37.22/885.61	38.98/359.51	38.98/359.44	40.02/104.01	40.02/104.00
	32	34.73/490.05	34.71/490.39	37.24/191.67	37.24/191.44	37.06/59.37	37.06/59.35
	37	32.16/284.01	32.16/284.18	35.12/114.05	35.12/114.11	34.42/35.76	35.41/35.76
1–7	22	39.36/1532.87	39.33/1540.11	40.01/920.17	39.97/919.54	43.39/121.76	43.34/121.51
	27	37.19/756.09	37.16/761.28	38.52/382.33	39.49/382.96	40.27/66.04	40.23/66.01
	32	34.58/394.37	34.54/397.66	36.65/192.38	36.62/192.63	36.98/36.20	36.96/36.25
	37	31.83/221.90	31.80/223.13	34.34/110.78	34.31/110.73	34.10/22.25	34.09/22.43
Seqences		Breakdancers		Ballet		Alt Moabit	
View	QP	JMVM	Proposed	JMVM	Proposed	JMVM	Proposed
0	22	38.80/1839.84	39.79/1840.22	41.32/649.03	41.32/648.16	40.65/1099.35	40.64/1096.74
	27	37.59/729.02	37.59/729.09	40.27/309.33	40.27/309.56	38.95/599.77	38.95/599.67
	32	36.28/369.06	36.27/369.01	38.70/178.38	38.70/178.52	36.52/341.77	36.51/341.80
	37	34.52/213.84	34.52/213.65	36.63/109.00	36.63/109.02	33.86/202.40	33.86/202.57
1–7	22	39.23/1472.68	39.21/1487.22	41.35/564.19	41.31/562.20	41.08/773.89	41.02/779.71
	27	38.00/569.15	37.97/574.23	40.27/257.80	40.24/257.81	39.29/352.70	39.21/356.27
	32	36.50/283.06	36.47/285.43	38.55/143.72	38.52/143.62	36.62/172.69	36.57/175.16
	37	34.50/161.55	34.46/162.18	36.28/86.57	36.26/86.56	33.95/95.99	33.93/98.00

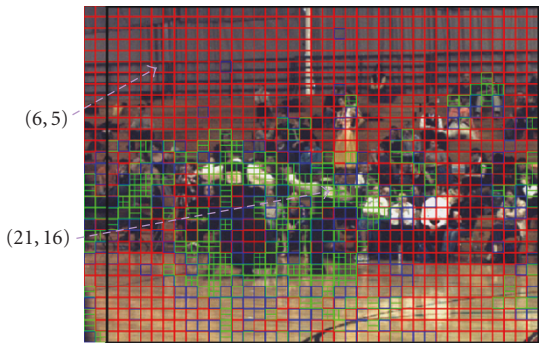


Figure 15: Illustration of overlapping region of the frame S_0T_6.

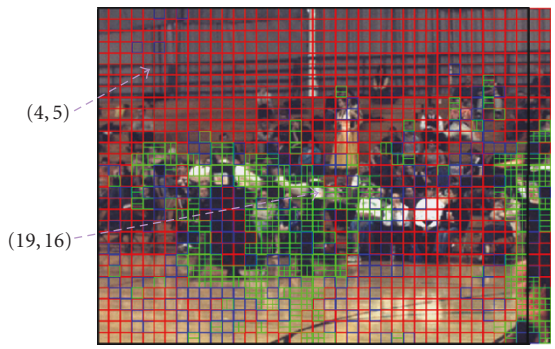


Figure 16: Illustration of overlapping region of the frame S_2T_6.

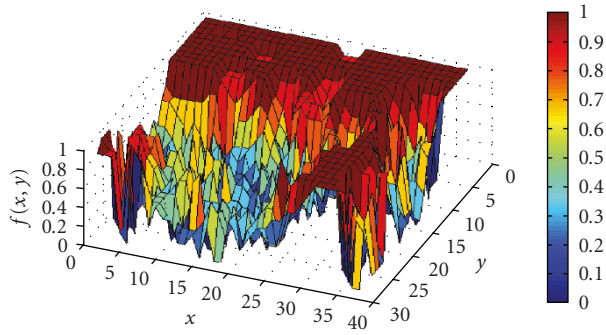


Figure 17: Illustration of inter-view mode correlation.

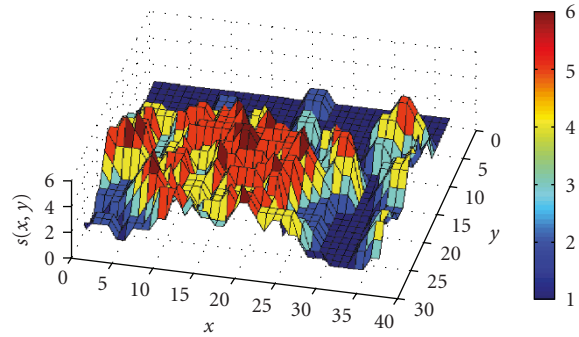


Figure 18: Illustration of intraframe mode correlation.

The current fast macroblock mode selection algorithms for the traditional single-view video coding cannot be applied directly to MVC due to the different prediction structures. After careful observations and detailed analyses, the features of the mode distribution, the RD cost of various modes, the inter-view mode correlation, and the intraframe mode correlation are exploited as the bases of the proposed algorithm.

(a) Eight views of Ballroom sequence

(b) Eight views of Exit sequence

(c) Eight views of Race1 sequence

(d) Eight views of Breakdancers sequence

(e) Eight views of Ballet sequence

(f) Eight views of Alt Moabit sequence

Figure 19: Multiview video test sequences.

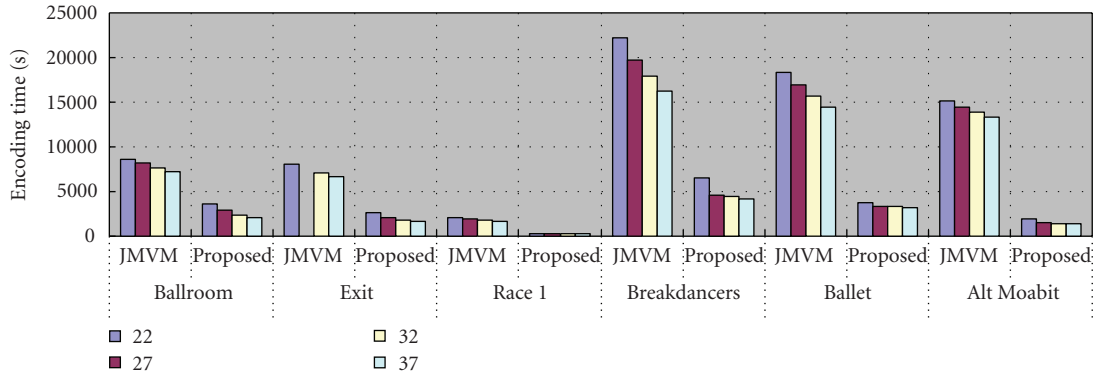


Figure 20: Encoding time comparison between the JMVM and the proposed algorithm.

The proposed hybrid fast macroblock mode selection algorithm consists of two methods, multithreshold fast macroblock mode selection method for nonanchor frames in the base view and the fast macroblock mode selection method based on mode correlations for nonanchor frames in the other views. The first method accelerates the encoding by halfway stopping the mode selection process via the thresholds which are updated for each frame. The second method utilizes the modes of the frame in neighboring views to predict the modes of the frame in the current view. The mode searching time is reduced because of the mode aggregation. Experimental results show that the proposed algorithm promotes the encoding speed by 2.37 ~ 9.97 times in comparison with the JMVM, the testing benchmark, while the algorithm hardly influences the RD performance.

Our current work is focusing on fast macroblock mode selection algorithm for MVC. If the proposed algorithm goes collaboratively with the other fast encoding algorithm, such as fast disparity and motion estimation, they can complement with each other to speedup the encoding process. There are many literatures about fast disparity and motion estimation for MVC. It is an interesting future work to design an integral algorithm to incorporate fast disparity and motion estimation algorithm into our current work for overall performance promotion.

ACKNOWLEDGMENTS

This work was supported by Natural Science Foundation of China (Grants 60472100, 60672073, 60872094), the Program for New Century Excellent Talents in University (NCET-06-0537), Natural Science Foundation of Ningbo (2007A 610037), and Scientific Research Fund of Zhejiang Provincial Education Department (20070954).

REFERENCES

[1] A. Smolic, K. Müeller, N. Stefanoski, et al., "Coding algorithms for 3DTV: a survey," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17, no. 11, pp. 1606–1620, 2007.

[2] A. Vetro, S. Yea, M. Zwicker, W. Matusik, and H. Pfister, "Overview of multiview video coding and anti-aliasing for 3D displays," in *Proceedings of the 14th IEEE International Conference on Image Processing (ICIP '07)*, vol. 1, pp. 17–20, San Antonio, Tex, USA, September 2007.

[3] A. Smolic, K. Müeller, P. Merkle, et al., "3D video and free viewpoint video—technologies, applications and MPEG standards," in *Proceedings of IEEE International Conference on Multimedia and Expo (ICME '06)*, pp. 2161–2164, Toronto, Canada, July 2006.

[4] M. Tanimoto, "Free viewpoint television—FTV," in *Proceedings of Picture Coding Symposium*, pp. 289–294, San Francisco, Calif, USA, December 2004.

[5] ITU-T Rec. H.264-ISO/IEC 14496-10 AVC, "Advanced video coding for generic audiovisual services," ITU-T and ISO/IEC Joint Video Team, 2005.

[6] T. Wiegand, G. J. Sullivan, G. Bjøntegaard, 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. 560–576, 2003.

[7] G. J. Sullivan and T. Wiegand, "Video compression-from concepts to the H.264/AVC standard," *Proceedings of the IEEE*, vol. 93, no. 1, pp. 18–31, 2005.

[8] Y. Zhang, G. Jiang, W. Yi, M. Yu, Z. Jiang, and Y. D. Kim, "An approach to multi-modal multi-view video coding," in *Proceedings of the 8th International Conference on Signal Processing (CSP '06)*, vol. 2, pp. 1401–1404, Guilin, China, November 2006.

[9] P. Merkle, A. Smolic, K. Müller, and T. Wiegand, "Efficient prediction structures for multiview video coding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17, no. 11, pp. 1461–1473, 2007.

[10] A. Kaup and U. Fecker, "Analysis of multi-reference block matching for multi-view video coding," in *Proceedings of the 7th Workshop Digital Broadcasting*, pp. 33–39, Erlangen, Germany, September 2006.

[11] ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6, "Comparative study of MVC prediction structures," *JVT-V132*, Marrakech, Morocco, January 2007.

[12] M. Flierl, A. Mavlankar, and B. Girod, "Motion and disparity compensated coding for multiview video," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17, no. 11, pp. 1474–1484, 2007.

[13] M. Flierl and B. Girod, "Multiview video compression," *IEEE Signal Processing Magazine*, vol. 24, no. 6, pp. 66–76, 2007.

[14] ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6, "Joint multiview video model (JMVM) 4.0," *JVT-W207*, San Jose, Calif, USA, April 2007.

[15] M.-J. Chen, G.-L. Li, Y.-Y. Chiang, and C.-T. Hsu, "Fast multiframe motion estimation algorithms by motion vector composition for the MPEG-4/AVC/H.264 standard," *IEEE Transactions on Multimedia*, vol. 8, no. 3, pp. 478–487, 2006.

[16] Y. Kim, J. Kim, and K. Sohn, "Fast disparity and motion estimation for multi-view video coding," *IEEE Transactions on Consumer Electronics*, vol. 53, no. 2, pp. 712–719, 2007.

[17] L.-F. Ding, P.-K. Tsung, W.-Y. Chen, S.-Y. Chien, and L.-G. Chen, "East motion estimation with inter-view motion vector prediction for stereo and multiview video coding," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '08)*, pp. 1373–1376, Las Vegas, Nev, USA, March 2008.

[18] P. Yin, H.-Y. C. Tourapis, A. M. Tourapis, and J. Boyce, "Fast mode decision and motion estimation for JVT/H.264," in *Proceedings of IEEE International Conference on Image Processing (ICIP '03)*, vol. 3, pp. 853–856, Barcelona, Spain, September 2003.

[19] 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, 2006.

[20] C. Kim and C.-C. J. Kuo, "Feature-based intra-/intercoding mode selection for H.264/AVC," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17, no. 4, pp. 441–453, 2007.

[21] I. Choi, J. Lee, and B. Jeon, "Fast coding mode selection with rate-distortion optimization for MPEG-4 Part-10 AVC/H.264," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 16, no. 12, pp. 1557–1561, 2006.

[22] M. Yin and H.-Y. Wang, "An improvement fast INTER mode selection for H.264 joint with spatio-temporal correlation," in *Proceedings of International Conference on Wireless Communications, Networking and Mobile Computing (WCNM '05)*, vol. 2, pp. 1237–1240, Wuhan, China, September 2005.

[23] ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6, "Statistical analysis of macroblock mode selection in JMVM," *JVT-Y026*, Shenzhen, China, October 2007.

[24] ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6, "Common test conditions for multiview video coding," *JVT-T207*, Klagenfurt, Austria, July 2006.

[25] ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6, "MVC motionskip mode," *JVT-W081*, San Jose, Calif, USA, April 2007.