FAST MODE DECISION FOR H.264 VIDEO ENCODER BASED ON MB MOTION CHARACTERISTIC

Xiangwen Wang[1], *Jun Sun*[1], *Yunqiang Liu*[2] *and Renjie Li*[1]

[1]Institute of Image Communication & Information Processing, Shanghai Jiao Tong University, Shanghai, China Shanghai Key Laboratory of Digital Media Processing and Transmissions, Shanghai Jiao Tong University, Shanghai 200240, China, [wxw21st, junsun, lrjie]@sjtu.edu.cn
[2] Philips Research East Asia, Shanghai 200070, China, yunqiang.liu@philips.com

ABSTRACT

Seven variable block sizes are adopted for inter-frame MB (macroblock) coding in H.264. This new feature achieves significant coding gain. However, the computation complexity of the mode decision is extremely high when RD (rate distortion) algorithm is used. In this paper, we propose a fast mode decision algorithm with fast coding block size selection based on MB motion characteristic for inter frames. In this algorithm, the residual_MB is first gotten through motion estimation for the MB. Then, the motion characteristic of the MB is extracted through careful analysis of the texture in the residual_MB. At last, the optimization mode for the MB is decided according to the motion characteristic. Experimental results show that the mode decision can save significantly computational complexity at cost of a little degradation of RD performance.

1. INTRODUCTION

In order to select the best block size for one MB, H.264 encoder employs rate distortion optimization (RDO) technique [1] to get the best coding result in terms of maximizing coding quality and minimizing data bits. The reference JM [2] encoder computes the RD-cost for every available block size and the one that leads to the least RD cost is selected. This “try all and select the best” philosophy is optimal in deciding the block size. Nevertheless, this optimal decision is achieved at the expense of high computational complexity. Its computationally intensive nature prevents it from practical video encoder implementations based on embedded processors.

Lots of fast mode decision algorithms containing fast block size selection have been proposed. In [3], a fast mode decision with fast block size selection algorithm is proposed which based on object and texture region movement resolution. The fast inter mode decision algorithm proposed by Wu etc. [4] is based on the MB’s edge direction information. Although some candidate block sizes are eliminated according to the edge direction of the MB, the computation of edge direction introduces a lot of additional complexity.

Most of the former algorithms are based on the MB intrinsic complexity analysis. But we strongly argue that there is no direct relation between the MB intrinsic complexity and the optimized coding mode. For example, a lot of MBs belong to still background and with similar motion characteristic can directly be coded using 16x16 block size though these MBs have complex details. It is obvious that the optimized mode has close relation with the motion MB characteristic. Without knowing the motion characteristic in the MB, false decision will be made.

In this paper, we propose a fast mode decision algorithm with fast block size selection method based on MB motion characteristic for inter frames. In this algorithm, the selection of optimized MB coding mode is combined with the motion estimation. Firstly, the residual_MB is gotten through motion estimation. Then, the texture direction of the residual_MB is extracted through synthesizing the edge directions of 16 4x4 residual blocks. Based on the relation between motion characteristic and residual signal, the MB coding mode is decided according to direction of the residual_MB texture. Experimental results show that this method significantly reduces the computation complexity.

The following sections are organized as follows. In Section 2, we study the relation between object motion and residual texture. We briefly present the fast mode decision with fast block size selection in Section 3. Residual_MB texture direction extraction method and fast MB partition method are presented in Section 4. In Section 5, some experiment results are shown. Section 6 concludes the paper.

2. RELATION BETWEEN OBJECT MOTION AND RESIDUAL TEXTURE

Among the texture characteristics, the texture coarseness and directionality are the two major texture characteristics used by the classification methods. Applied to frame residual signals, the texture coarseness deals with the scene activity within picture frames whereas the texture directionality relates to the direction of motion. In [5], Seferidis applies the texture directionality to decide the motion direction in motion estimation. There is direct relationship between the texture direction and motion direction. This relationship can be demonstrated through one simple example. Consider two successive frames of an image sequence where a dark object moves from left to right in front of a gray background, as shown in Fig. 1. Then a large area of the frame difference signals has small amplitudes, corresponding to the pixels with almost unchanged values between the frames. Few contour-like areas, with large amplitudes are formed from the changed pixels in the regions of the moving edges. Fig. 1c illustrates the frame difference signal of the moving object with the larger values drawn with darker colors. Fig. 1d shows the texture directions of frame difference signal and the object motion direction. From Fig. 1, we can get the following conclusion: if the motion is uniform, the direction of the movement is perpendicular to the dominant direction of the texture in the frame residual picture.

Now, we apply this conclusion to the residual_MB signals. For a MB, it is possible to contain one or more objects and these objects may not move in the same direction. Therefore using only one motion vector may not be enough to completely describe the motion of all objects in the MB. This results that only part of the MB can have good motion compensation and the resulting texture varies in large difference due to the mismatch in the remaining part of the MB. Those mismatch portions form the texture in the MB residual. We assume that the motion estimation can find a matching block in the reference for each block. If we segment the MB along the texture_MB direction, different objects with different motion activities can be included in different sub-blocks. So we can represent object movement more accurately and reduce the error as the fixed block size prediction is used. This is the primary reason for adopting seven different sizes in H.264/AVC inter coding.

We assume that each frame in Fig. 1 is composed by 16 MBs. The dashed lines in frames are the boundaries of each MB. The moving object spans the middle four MBs in the frame. Based on the analysis above, we segment the MB along the texture_MB direction. The optimal MB partition mode is shown in Fig. 1c with black lines for the two above MBs including moving object. Here no motion estimation is employed, so the residual_MB is just the difference of the same position between two successive frames. In our method, the residual_MB is gotten through motion estimation with the most match motion vector for the MB.

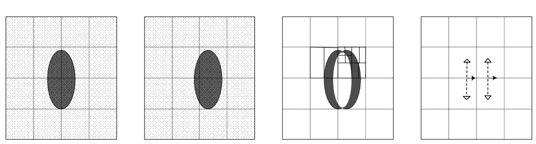


Fig. 1. Texture analysis of the frame difference signal: (a) and (b) successive frames, (c) Frame difference signal with optimal MB partition mode, (d) Texture and motion directions

3. FAST MODE DECISION

The fast block size decision is conducted after the motion estimation for the MB. We assume that the motion estimation can find a matching block in the reference for each MB. The texture analysis and fast MB partition mode are conducted on the residual_MB gotten from the MB motion estimation. Through the motion estimation and texture analysis, the motion correction between the neighboring MBs can be explored and the MBs in stationary region can be detected. Those MBs with similar motion characteristic are more likely to be coded using large block size. If the 16x16 block size is selected for the MB, SKIP detection is employed to detect whether this MB can be SKIP encoded. If the MB is partitioned into two 16x8 or 8x16 blocks, new motion estimations are employed for each block. If the MB is partitioned into 4 8x8 blocks, motion estimation for each block is employed and the redidual_subMB for each 8x8 block is gotten. A further residual_subMB texture direction analysis and subMB partition process as the MB partition method is conducted. The fast mode decision is diagrammed in Fig. 2.

4. RESIDUAL_MB TEXTURE ANALYSIS AND FAST MB PARTITION

To analyze the residual_MB texture, we first divided the residual_MB into four 8x8 blocks. Then each 8x8 block is divided into four 4x4 blocks. At last, each 4x4 block is further segmented into four 2x2 blocks as shown in Fig. 1. The steps to extract the residual_MB texture direction are described as follows.

a) Getting the edge direction inside each 4x4 block through the analysis of four 2x2 blocks in the block;
b) Getting the texture direction of each 8x8 block through synthesizing the edge directions of the four 4x4 blocks in the block;
c) Getting the texture direction of the residual_MB through synthesizing the texture of 4 8x8 blocks.

In the following three sub-sections, we first describe the edge direction analysis method for a 4x4 block. Then, we present the texture direction extraction method. The fast MB partition method is presented at last.

4.1 Edge direction analysis for a 4x4 block

Let A, B, C and D respectively denotes the sum of intensity of all pixels in the corresponding blocks as shown in Fig. 3. A, B, C and D are calculated as follows.

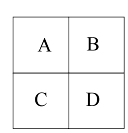


Fig. 3. Division of 4x4 block

$$A=\sum_{y=0}^{1}\sum_{x=0}^{1}P(x,y),\quad B=\sum_{y=0}^{1}\sum_{x=2}^{3}P(x,y) \qquad (1,2)$$

$$C=\sum_{y=2}^{3}\sum_{x=0}^{1}P(x,y),\quad D=\sum_{y=2}^{3}\sum_{x=2}^{3}P(x,y) \qquad (3,4)$$

P(x, y) (0<x<4, 0<y<4) denotes the intensity of the pixel at position (x, y) of the current 4x4 block.

In order to obtain the local edge direction within a 4x4 block, we introduce two edge feature parameters: vertical edge parameter F_v and horizontal edge parameter F_h.

$$F_v=\left\lfloor\frac{(A+C)-(B+D)}{S}\right\rfloor \qquad (5)$$

$$F_h=\left\lfloor\frac{(A+B)-(C+D)}{S}\right\rfloor \qquad (6)$$

Where S is a scaling factor, the scaling factor S is determined according to the quantization factor Qp.

$$S=\begin{cases}8 & Q_P<20\\ 16 & 20\le Q_P<30\\ 32 & 30\le Q_P<40\\ 64 & 40\le Q_P\end{cases} \qquad (7)$$

$\lfloor\ \rfloor$ represents the floor function. The physical meanings of the two parameters are shown pictorially in Fig. 4. F_v and F_h represent intensity differences between the left and right parts and between the upper and lower parts of the block respectively. According to the two parameters, we can obtain the dominant edge direction information within the current 4x4 block [6], which is shown in Table I.

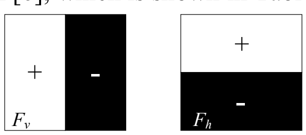


Fig. 4. Physical meanings of F_v and F_h

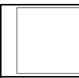
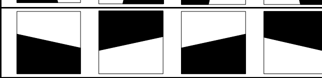
4.2 Residual_MB texture extraction

Each 4x4 block is classified into one of four categories according to the edge direction as diagrammed in Table I. Through synthesizing the edge directions of four 4x4 blocks, we classify each 8x8 block as follow.

a) No obvious texture. Only if all the 4x4 blocks inside are belong to No Obvious Edge.
b) Vertical dominant direction texture. If left two 4x4 blocks or right two 4x4 blocks or only one 4x4 block have vertical dominant edge.
c) Horizontal dominant direction texture. If above two 4x4 blocks or below two 4x4 blocks or only one 4x4 blocks have horizontal dominant edge.
d) Diagonal direction texture for all other combinations.

With the same classification method, the residual_MB is categorized into four classes according to the categories of the four 8x8 residual blocks: no obvious texture, vertical dominant direction texture, horizontal dominant direction texture and diagonal direction texture.

Table I The edge direction Categories inside one 4x4 block

case	Fv & Fh	Edge direction	
I	$\lvert F_v\rvert = \lvert F_h\rvert = 0$	No Obvious Edge	
II	$\lvert F_v\rvert > \lvert F_h\rvert$	Vertical Dominant Edge	
III	$\lvert F_v\rvert < \lvert F_h\rvert$	Horizontal Dominant Edge	
IV	$\lvert F_v\rvert = \lvert F_h\rvert > 0$	Diagonal Edge	

4.3 Fast MB partition method

In our method, we segment the MB along the direction of residual_MB texture as follows.

a) 16x16 size for no obvious texture;
b) 16x8 for horizontal dominant direction texture;
c) 8x16 for vertical dominant direction texture;
d) 8x8 for diagonal direction texture.

Further motion estimation is conducted for all blocks inside the MB except that 16x16 block size is selected. If the MB is split into four 8x8 blocks, a sub-MB partition is employed after the motion estimation for each 8x8 block. With the same classification method, we classify each 8x8 residual block into four classes: a) No obvious texture, b) Vertical dominant direction texture, c) Horizontal dominant direction texture, d) Diagonal direction texture. Then, we segment the sub-MB along the direction of 8x8 residual block texture. Further motion estimation is conducted for all blocks inside the sub-MB except that 8x8 block size is selected.

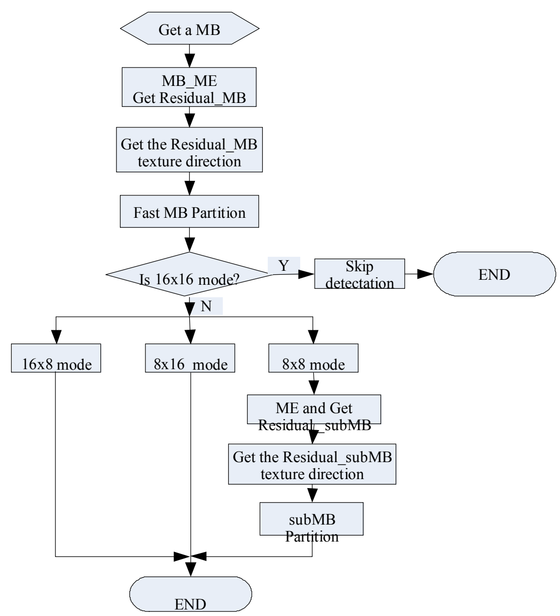


Fig. 2. Fast mode decision

5. EXPERIMENTAL RESULTS

The proposed method was implemented into JM8.6 provided by JVT. Different CIF sequences are selected according to motion and texture information. These sequences are Akiyo, Mobile, Forman, and Paris. The main parameters set in the encoder configure file are list as follows: 100 frames, one I frame followed by all P frames, CAVLC, search range 16, frame rate 30 fps and 1/4 pixel MV resolution. Fast motion estimation algorithm proposed in JVT-F017 [7] is employed. In our algorithm, intra predictions are conducted only if the block size for the MB is not 16x16 and the mode with minimum RD cost is selected as the encoding mode. Table II is a summary of the performance of the proposed MB partition method. The fast block size selection method proposed by Andy etc. in [8] is also listed in the Table. In the table, "Time (%)", "PSNR (dB)" and "Bits (%)" represent time change in percentage, PSNR change, and bit rate change in percentage, respectively, comparing with the JM8.6 encoder. The positive values mean increments while negative values mean decrements.

In terms of PSNR, the proposed method introduces an average degradation of 0.06dB. For Akiyo sequence, there is only little differences in the average PNSR values. As to compression ratio, the proposed algorithm produces slightly higher bit rates than the JM encoder of H.264. The RDO technique is designed to get maximum quality and minimum bit rate, while our proposed algorithm is designed to partition the MB according to the motion property inside the MB which is very close to the RDO. With little loss of PNSR and increment of bit rate, the proposed algorithm significantly reduces the computation complexity. The average time saving of all the test sequences is more than 40%.

From the Table II, we can see that the RD performance of our algorithm is similar to that of the algorithm proposed by Andy in [8]. With the same QP value, the PSNR loss by our algorithms is less than Andy's while compression ratio with our algorithm is a little lower than Andy's. From the Table we can see that time saving by our algorithm more significant than by the algorithms in [8]. This is because the additional complexity induced by the texture direction analysis is neglectable.

Table II. The experimental results of the proposed algorithm and the algorithm in [8] VS. H.264 at QP=28

Sequence	Δ PSNR(dB)		Δ Rate (%)		Time Saving(%)	
	[8]	Prop.	[8]	Prop.	[8]	Prop.
Foreman	-0.086	-0.065	+4.79	+6.74	-25.22	-40.25
Akiyo	-0.027	-0.021	+1.45	+2.68	-35.56	-63.60
Paris	-0.080	-0.055	+2.06	+3.00	-21.93	-40.78
Mobile	-0.104	-0.091	+3.06	+5.32	-19.85	-45.21

6. CONCLUSIONS

In this paper, we proposed a fast mode decision based on the MB motion characteristic. The relation between motion characteristic and residual signal is analyzed distinctively and the principle in our method is to segment the MB according to the motion characteristic in the MB. Experimental results have shown that the proposed algorithm provides significant reduction in computational complexity with little coding loss compared with the brute force full-search algorithm. Experimental results also show our method is more optimized than the block size decision algorithms based on MB intrinsic complexity analysis. This method is especially suitable for practical video encoder implementations based on embedded processors such as DSP and mobile phones.

7. REFERENCES

[1] Joint Video Team (JVT) of ITU-T and ISO/IEC JTC 1, Advanced Video Coding: ITU-T Rec. H.264 and ISO/IEC14496-10, version 1: May 2003, version 2: March 2004, version 3: July 2004, version 4: January 2005.

[2] http://www.vcodex.fsnet.co.uk/h264.html.

[3] T Koga, K Iinuma, A Hirano, Y Iijima and T Ishiguro, "Motion Compensated Interframe Coding for Video Conferencing", Proc NatTelecomm Conf, New Orleans, L A , Nov 29 - Dec 3, 1981, ppG5 3 1-5 3 5

[4] Wu, D.; Pan, F.; Lim, K.P.; Wu, S.; Li, Z.G.; Lin, X.; Rahardja, S.; Ko, C.C.; "Fast intermode decision in H.264/AVC video coding" Circuits and Systems for Video Technology, IEEE Transactions on Volume 15, Issue 7, July 2005 Page(s):953 - 958 Digital Object Identifier 10.1109/TCSVT.2005.848304

[5] Seferidis, V.E.; Ghanbari, M., "Adaptive Motion Estimation based on texture analysis" Communications, IEEE Transactions on Volume 42, Issue 234, Part 2, February-April 1994

[6] Shen, B., and Sethi, 1. K., "Direct feature extraction from compressed images", S P E Vo1.2670 Storage & Retrieval for Image and Video Databases IV, 1996.

[7] Z. Chen, P. Zhou, and Y. He, "Fast integer pel and fractional pel motion estimation for JVT," presented at the 6th JVT-F017 Meeting, Awaji Island, Japan, Dec. 2002.

[8] Andy C.Yu, "Efficient Block-Size Selection Algorithm for Inter-Frame Coding in H.264/MPEG-4 AVC". IEEE ICASSP 2004:III169~III172.

[9] Zhang Yong-dong; Dai Feng; Lin Shou-xun; "Fast 4x4 Intra-prediction Mode Selection for H.264" Multimedia and Expo, 2004. ICME '04. 2004 IEEE International Conference on Volume 2, 27-30 June 2004 Page(s):1151 - 1154 Vol.2