LETTER

An All-Zero Block Mode Decision Algorithm for H.264/AVC Optimization

Chaoke PEI[†,††a)], ***Student Member***, **Li GAO**[†b)], **Donghui WANG**[†], ***and*** **Chaohuan HOU**[†], ***Nonmembers***

SUMMARY The H.264/AVC standard achieves significantly high coding efficiency if multiple block size Motion Estimation is adopted. However, the complexity of Motion Estimation and DCT is dramatically increased as a result. In previous work we propose an early mode decision algorithm to control the complexity, based on all-zero-blocks detection in 16×16 size. In this paper, we improve the algorithm. Firstly, we propose to detect all-zero blocks in 16×16, 8×8 and 4×4 sizes to simplify the course of mode decision. Secondly, we define the thresholds which are used to terminate motion estimation and mode decision in advance for these sizes. Last, we present the whole proposed algorithm. Experiments show that about 77% encoding time and 85% motion estimation time can be saved on average, which is better than state-of-the-art approaches.
key words: *H.264/AVC, early mode decision, all-zero block, motion search*

1. Introduction

The newest video compression standard H.264/AVC [1] significantly outperforms previous standards in terms of coding efficiency, using motion estimation (ME), discrete cosine transform (DCT), quantization (Q), inverse quantization (IQ) and inverse discrete cosine transform (IDCT). However, the computation complexity of encoding increases greatly. As a result, reducing the computation of the encoder becomes the top priority in order to decrease encoding time.

Many studies have investigated the course of ME, which occupies more than 70% of the total encoding computation. Another important issue is the cost of DCT (including Q, IQ and IDCT) [4], which approaches 24% of total coding time. For ME, two ways are used to reduce the computation generally. One is to reduce the number of search points. Several fast search algorithms have been proposed [5], [6]. The other way is to terminate inter ME early [9]–[11] or to predict intra-mode fastly [12]. For DCT, it is quite common that all of the DCT coefficients in a block are quantized to zeros. Therefore it is important to detect these all-zero-blocks (AZBs) early in order to skip their transformation and quantization. Some methods in this area are also proposed [7], [8]. AZB detection aims at 4×4 sub-blocks. It compares the cost of sum of absolute difference (SAD) of a sub-block to a predefined threshold and decides to skip DCT or not. Since the amount of AZBs is considerable, the algorithms also reduce the encoding computation.

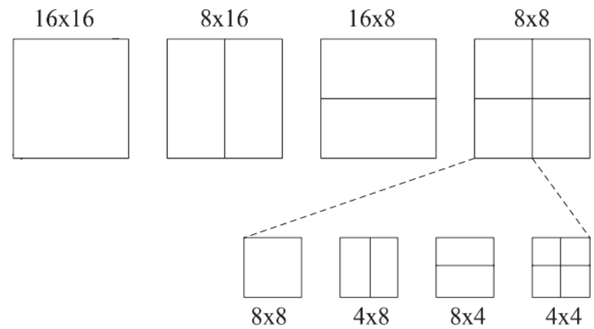


Fig. 1 Seven prediction inter modes in H.264.

As known, there are seven block sizes used for inter prediction in H.264/AVC as shown in Fig. 1. According to these block sizes, every block can be classified to different modes to process ME for P slices. The proposed algorithm takes advantage of the methods of AZB detection for blocks in 16×16, 8×8 and 4×4 sizes, defines thresholds to be used to terminate ME and make mode decision (MD) in advance, and also saves the computation of DCT.

The rest of this paper is organized as follows. In Sect. 2, the course of ME is introduced, an AZB detection method for three kinds of size is analyzed, and an early motion estimation and mode decision algorithm which takes advantage of the AZB detection is proposed. The experimental results are presented in Sect. 3. Finally the conclusion is located in Sect. 4.

2. Proposed Algorithm

In order to reduce the compute complexity, we defined a threshold to compare with SAD of 16×16 blocks to make a decision on whether the block is AZB after DCT (including quantization) [13]. In this paper, we extended the detection to 16×16, 8×8 and 4×4 size blocks.

2.1 Motion Estimation and Mode Decision

In JVT reference software JM, ME procedure is carried out to obtain the best block matching for each block size. A cost function is evaluated for each motion vector (MV) within the search range:

Manuscript received June 2, 2010.
Manuscript revised August 20, 2010.
[†]The authors are with the Department of Integrated Digital System, Institute of Acoustics, Chinese Academy of Sciences, 21 Beisihuanxi Road, Haidian, Beijing 100190, China.
[††]The author is with the Graduate School of the Chinese Academy of Sciences, 80 Zhongguancun East Road, Haidian, Beijing 100080, China.
a) E-mail: Pei.Chaoke@gmail.com
b) E-mail: gaoli@mail.ioa.ac.cn
DOI: 10.1587/transinf.E94.D.384

$$J_{motion}(MV) = SAD_{MV} + \lambda_{motion} \times R(MV) \quad (1)$$

In (1) $J_{motion}(MV)$ is the motion cost with the motion vector MV, and SAD is the sum of absolute difference between the original and predicted blocks. λ_{motion} is the Lagrangian multiplier and $R(MV)$ represents the total amount of bits needed for encoding motion information. For each mode, MV is selected to minimize $J_{motion}(MV)$ in (1). Next, MD is carried out to choose the best mode by means of a rate-distortion optimization (RDO) model.

With the similar MV, the more parts the block is divided into, the greater $R(MV)$ is, for every part needs several bits to store the corresponding MV. Because the block of mode SKIP and mode 1 is undivided, if the SAD of the mode SKIP and mode 1 is zero, $J_{motion}(MV)$ is likely to be the least in these modes. As a result, mode SKIP and mode 1 may be chosen as the best mode in mode decision.

2.2 An Improved AZB Mode Decision Algorithm

It is highly probable that the AZBs in 16×16 size will choose mode SKIP or mode 1 as the best mode [13]. However, according to the experiments mode 2 or mode 3 also may be chosen at a certain percentage. As shown in Table 1, about 98% AZBs (in 16×16 size) will choose mode SKIP (mode 0) or mode 1. Almost all AZBs (in 16×16 size) will choose the best mode from mode SKIP, mode 1, mode 2 (16×8) and mode 3 (8×16). Other modes (mode 8 or intra modes) are hardly to be chosen. As a result, when the 16×16 blocks are detected to be AZBs, it is proposed that if the RD cost of mode SKIP is less than that of mode 1 the mode SKIP is chosen, otherwise the best mode will be chosen among mode SKIP, mode 1, mode 2 and mode 3.

Using the AZB detection results of 8×8 blocks, the ME for mode 2 and mode 3 can be simplified. When the two 8×8 blocks in a 8×16 or 16×8 block are both AZBs, the motion vector of the ME for this size could be approximately decided. As shown in (1), when $SAD(MV)$ is small enough, there is no need to go on searching other locations with more $SAD(MV)$. Furthermore, in H.264/AVC the integer transform which is derived from DCT, is performed on 4×4 blocks. If the 4×4 blocks are detected to be AZBs, the computation of integer transform can be saved. Since the coding computation of integer transform is large, the saving computation is also large.

2.3 AZB Detection

In H.264/AVC, DCT is carried out on the 4×4 blocks. Several theories are proposed to detect AZBs in this size level. However, we propose to detect AZBs in 16×16, 8×8 and 4×4 sizes and base on the detection to simplify the courses of ME and DCT. In the previous work we proposed that if the 16×16 blocks are detected to be AZBs, mode SKIP or mode 1 would be chosen to be the best mode. In order to reduce the "side-effect" on PSNR and the bit-rate, mode 2 or mode 3 may also be chosen under some conditions. The AZB detection in 8×8 size blocks can also be used for ME in mode 2 and mode 3. Besides, taking advantage of the AZB detection in 4×4 size blocks, the course of integer transform is able to be skipped.

Table 1 Mode statistics for AZBs.

Sequences	Mode 0+1 (%)	Mode 2+3 (%)	Mode 0+1+2+3(%)
Salesman	99.46	0.39	99.85
News	99.01	0.85	99.86
Akiyo	98.87	1.09	99.96
Paris	97.93	1.45	99.38
Silent	98.46	1.28	99.74
Container	99.46	0.31	99.77

In the previous work [13], we present the algorithm to terminate MD in advance and define a threshold for all blocks in 16×16 sizes. In this paper, we develop the algorithm to simplify ME and MD for all blocks in 16×16, 8×8 and 4×4 sizes. The thresholds of all kinds of sizes are obtained as follows.

In this section, we assume S denotes the block sizes such as 16, 8 and 4. The thresholds to predict the AZB can be defined by the same way in [13]:

$$SAD < TH(S) \quad (2)$$

$$TH(S) = \sqrt{2}\alpha S^2/(\gamma\sqrt{\pi(ARA^T)_{u,u}(ARA_T)_{v,v}}) \times Qp \quad (3)$$

Where

$$A(u,x,S) = \sqrt{2/S}\,C(u)\cos((2x+1)u\pi/2S) \quad (4)$$

$$R(s) = \begin{pmatrix} 1 & \rho & \rho^2 & \dots & \rho^{S-1} \\ \rho & \rho^2 & \rho^3 & \dots & \rho^{S-2} \\ \rho^2 & \rho^3 & \rho^4 & \dots & \rho^{S-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{S-1} & \rho^{S-2} & \rho^{S-3} & \dots & 1 \end{pmatrix} \quad (5)$$

$[u,v]$ is the (u,v)th component of the matrix, A is the basis vector of DCT and ρ is the correlation coefficient of the DCT. α is set to be 3.5 and ρ is set to be 0.6 following the convention.

In [13], we assumed the coefficients of the blocks can be modeled by a Gaussian distribution [3]. According to [14], the coefficients may be Laplacian distributed or Cauchy distributed. However, in this paper no matter what kind of distribution the coefficients are, the probability that the values would fall within $(-3\sigma, 3\sigma)$ is much more than 99%. Therefore, the criterion for AZB prediction can be defined as (2)(3). Because the integer transform is derived from DCT, the threshold is also applicable for 4×4 blocks.

Based on the facts and analysis stated above, the improved AZB mode decision algorithm is performed as follows:

Step 1) Motion search for the 16×16 mode, if SAD is less than $TH_{16\times16}$, stop searching for this mode, set $SKIP\_flag = 1$ and go to Step3. Otherwise, complete the search, set $SKIP\_flag = 0$ and go to Step2.

Step 2) Perform motion search for mode 16×8, 8×16 and p8×8. For each block in 16×8 or 8×16 size, if the two

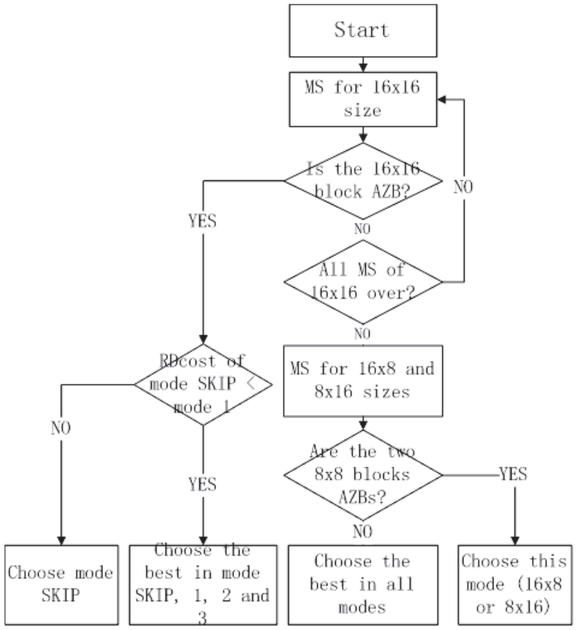


Fig. 2 The flow chart of the algorithm.

8×8 blocks are both AZBs, MV is decided and the search for this block is stopped.

Step 3) If $SKIP\_flag = 0$ choose the best mode among all modes;

Otherwise, if the RD cost of mode SKIP is less than mode 1, choose mode SKIP; or else further perform mode 2, mode 3 and choose the best mode.

Step 4) Go to the next block.

The flow chart is shown in Fig. 2.

Note that in each step, we use the threshold to predict whether every 4×4 block is AZB before DCT. If the block is AZB, the course of DCT is skipped. The computation of DCT for AZBs can be saved.

3. Experimental Results

The proposed algorithm was evaluated by a group of experiments. The simulation was implemented on JM13.2. Several sequences are tested, including QCIF (176×144) which are "akiyo", "news", "salesman", and CIF (352×288) which are "paris", "silent", "container". The full search mode is used and the number of the reference is set to 1. The MV search range is 16 and the MD metric is Hadamard. Quantization factor is set to be 28, and according to the lookup table in H.264/AVC, Qp in the (3) is 16.

To indicate the changes, $\Delta Time$, $\Delta METime$, $\Delta PSNR$ and $\Delta BitRate$ are used. As shown in Table 2, about 77% of the encoding time and about 85% of the ME time can be saved on average. We compare the proposed improved AZB mode decision algorithm under the same conditions with other mode decision algorithms (Moon [7], Ming [9]) and our previous proposed AZB algorithm (AZB [13]), which all use early AZB detection to save computation of ME and DCT. As shown in Table 3, the proposed improved AZB mode decision algorithm can save more computation.

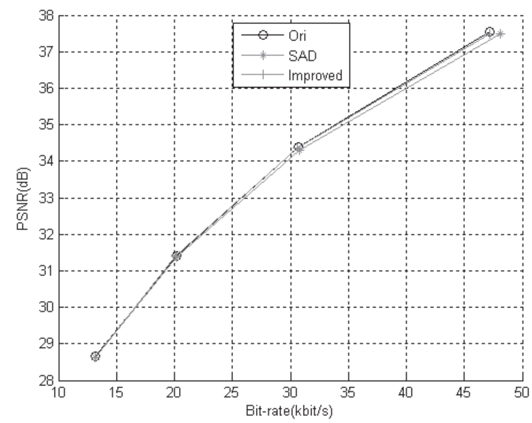
To compare with our previous work, we drew the RD curve and rate-time saving curve of the SAD MD algo-

Table 2 Experimental results compared to JM13.2 full search.

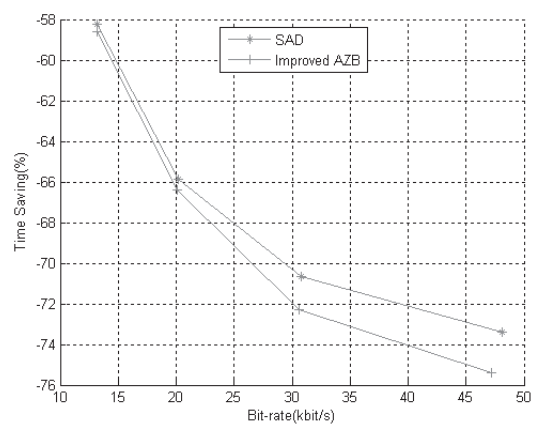
Sequences	$\Delta Time$ (%)	$\Delta METime$ (%)	$\Delta PSNR$ (dB)	$\Delta BitRate$ (%)
Salesman	−79.07	−87.91	−0.05	0.388
News	−75.42	−83.94	−0.05	0.042
Akiyo	−80.51	−90.12	−0.09	1.04
Paris	−65.66	−71.77	−0.05	0.20
Silent	−81.11	−86.91	−0.03	1.23
Container	−84.13	−90.45	−0.04	0.545
Average	−77.65	−85.18	−0.05	0.574

Table 3 Performance comparison.

$\Delta Time$(%)	Moon	Ming	AZB	Improved AZB
Foreman	−42.37	−61.47	−52.49	−62.42
Grandma	−62.75	−79.05	−78.99	−78.57
Mthr&Dtr	−47.39	−65.54	−80.11	−79.14
Salesman	−56.55	−75.30	−79.83	−79.07
Paris	−49.18	−67.88	−63.32	−75.11
Container	−56.81	−78.60	−79.98	−84.13
Average	−52.51	−71.31	−72.45	−76.41



(a) RD curve



(b) Bit rate-Time Saving curve

Fig. 3 Performance comparison.

rithm (SAD) [13] and this improved AZB MD algorithm (Improved) for various Qps (28, 32, 36, 40). The experiment is implemented on the sequence "news_qcif.yuv". The adverse impact of PSNR and bit-rate is improved. The decrease of PSNR is 0.05 and the increase of bit-rate is 0.57%. As shown in Fig. 3, the RD curve of the improved AZB algorithm is nearly the same with the original H.264/AVC JM approach. The improved algorithm can save more coding time and at the same time the increment of bit-rate is much less.

4. Conclusion

In this paper we extended the previous work and presented an improved AZB mode decision algorithm for H.264/AVC optimization to reduce the encoding computation. We propose to detect AZBs in all 16×16, 8×8 and 4×4 blocks, terminate motion search early and simplify the mode decision. The simulation results reveal that the proposed algorithm is effective.

References

[1] T. Wiegand, G.J. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," IEEE Trans. Circuits Syst. Video Technol., vol.13, no.7, pp.560–576, Aug. 2003.

[2] A.K. Jain, Fundamentals of Digital Image Processing, Englewood Cliffs, Prentice Hall, 1983.

[3] E.Y. Lam, "Analysis of the DCT coefficient distributions for document coding," IEEE Signal Process. Lett., vol.11, no.2, pp.97–100, Feb. 2004.

[4] H.S. Malvar, "Low-complexity transform and quantization in H.264/AVC," IEEE Trans. Circuits Syst. Video Technol., vol.13, no.7, pp.598–603, July 2003.

[5] R. Li, B. Zeng, and M.L. Liou, "A new three-step search algorithm for block motion estimation," IEEE Trans. Circuits Syst. Video Technol.. vol.4, no.4, pp.438–442, Aug. 1994.

[6] K. Ng, L. Po, K. Wong, C. Ting, and K. Cheung, "A search patterns switching algorithm for block motion estimation," IEEE Trans. Circuits Syst. Video Technol., vol.19, no.5, pp.753–759, 2009.

[7] Y.H. Moon, G.Y. Kim, and J.H. Kim, "An improved early detection algorithm for all-zero-blocks in H.264 video coding," IEEE Trans. Circuits Syst. Video Technol., vol.15, no.8, pp.1053–1057, Aug. 2005.

[8] H. Wang, "Hybrid model to detect zero quantized DCT coefficients in H.264," IEEE Trans. Multimed., vol.9, no.4, pp.728–735, June 2007.

[9] Y.M. Lee and Y. Lin, "An improved zero-block mode decision algorithm for H.264/AVC," Proc. ICIP, vol.5, pp.293–296, Sept. 2007.

[10] H. Zeng, C. Cai, and K. Ma, "Fast mode decision for H.264/AVC based on macroblock motion activity," IEEE Trans. Circuits Syst. Video Technol., vol.19, no.4, pp.491–499, April 2009.

[11] F. Pan, X. Lin, S. Rahardja, K.P. Lim, Z.G. Li, D. Wu, and S. Wu, "Fast mode decision algorithm for intraprediction in H.264/AVC video coding," IEEE Trans. Circuits Syst. Video Technol., vol.15, no.7, pp.816–822, July 2005.

[12] J. Kim, B.G. Kim and C.S. Cho, "Fast intra-mode prediction algorithm in H.264/AVC video coding," IEICE Trans. Inf. & Syst., vol.E90-D, no.8, pp.1320–1323, Aug. 2007.

[13] C.K. Pei, S. Yang, L. Gao, and W. Ma, "An early mode decision algorithm for H.264 optimization," Proc. Picture Coding Symposium, pp.1–4, May 2009.

[14] I. Pao and M. Sun, "Modeling DCT coefficients for fast video encoding," IEEE Trans. Circuits Syst. Video Technol., vol.9, no.4, pp.608–616, June 1999.