Fast block size selection algorithm for inter frame coding in H.264/AVC on TMS320C6416 DSP

Haithem Chaouch, Imen Werda, Amine Samet, Mohamed Ali Ben Ayed and Nouri Masmoudi

Laboratory of Electronics and Information Technologies (LETI)
University of Sfax, National school of Engineering
B.P.W, 3038 Sfax, TUNISIA

Summary
Motion estimation and compensation techniques are widely used for video coding applications but the real-time motion estimation is not easily achieved due to its enormous computations. Therefore, it would be greatly beneficial to optimize as much as possible the motion estimation bloc which is considered to be the most important in terms of computational cost. In this paper, a new fast block size selection DSP-based algorithm is presented, in which computation complexity is greatly reduced when achieving the same video quality. Experimental results show a 45.75% improvement in speed with no major loss in video quality (objective (PSNR) and subjective (SSIM)).

Key words:
H.264, motion estimation, block size selection, DSP, PSNR, SSIM.

1. Introduction

The new H.264/AVC video coding standard [1] can deliver significantly improved compression efficiency compared with previous standards, supporting higher quality video over lower bit rate channels. Due to this improved compression efficiency and increased flexibility of coding and transmission, H.264/AVC has the potential to enable new video services. However, the performance gains of H.264/AVC come at a price of increased computational complexity [2].
The processing overhead required to implement H.264/AVC is likely to be a major problem for real time application. There is therefore a need to develop low complexity implementations of H.264/AVC that offer the performance and flexibility advantages of the standard without an excessive computational cost. An H.264/AVC video encoder typically carries out a number of encoding processes including motion estimation, motion compensation, transform, quantization and entropy coding.In video coding, the high correlation between successive frames can be exploited to improve coding efficiency, which is usually achieved by using motion estimation (ME) and motion compensation technology.

Manuscript received May 5, 2007
Manuscript revised May 20, 2007

Many ME methods have been studied in an effort to reduce the computational complexity of the ME [3-5]. The new video coding standard, H.264/AVC, uses variable block sizes ranging from 4x4 to 16x16 in interframe coding. This new feature has achieved significant coding gain compared to coding a macroblock (MB) using fixed block size. However, this feature results in extremely high computational complexity.
This paper proposes a fast intermode decision algorithm to decide the best mode in interceding process. This pertinent DSP-based approach makes use of the spatial homogeneity and the temporal stationarity characteristics of video objects.
The rest of the paper is structured as follows. Section 2 presents an overview of intercoding in H.264/AVC. Section 3 presents in detail the fast intermode decision algorithm. Experimental results are presented in Section 4. Finally, conclusion will be given in Section 5.

2. Mode selection overview in H.264/AVC

Unlike previous standards, H.264/AVC adopts a tree-based decomposition to partition a MB into smaller sub-blocks of specified sizes illustrated in Figure 1. These different block sizes actually form a two-level hierarchy inside a MB. The first one includes block size of 16x16, 16x8, or 8x16. In the second level, each 8x8 block can be one of the subtypes such as 8x8, 8x4, 4x8 or 4x4. The availability of smaller ME blocks improves the ability of the model to handle fine motion detail and result in better subjective viewing quality because they do not produce large blocking artefacts.

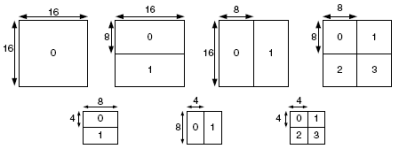

Fig.1. Inter modes with seven different block sizes ranging from 4X4 to 16X16.

Serious experiments on the test video sequences used in JVT Test Model Ad Hoc Group [9] show that there is an

average of 35% homogeneous area in a typical video frame, and these areas are suitable for larger size inter mode coding. Therefore, several cost calculation of small size modes can be saved. Based on this consideration, several mode decision algorithms were proposed to reduce the number of candidate modes [5-8].
In [5], many Fast Variable Block-Size (VBS) Motion Estimation algorithms were tested before proposing the Zoom Motion Estimation (ZME) algorithm structured into 3 steps:

- Step 1:

In a first step, 16x16 block-size motion estimation is used. If the SAD value is smaller than a threshold value V1 that, depend on the quantization step QP, the 16x16 block-size is kipped as a result and terminate the loop. Otherwise, save the SAD value and go to step 2.

- Step 2:

In this step, the MB 16x16 is split into four 8x8 blocks for further matching.
If SAD16x16 > sum of four SAD8x8 then go to step 3, else 16x16 blocksize is chosen.

- Step 3:

Each block 8x8 will be split into four 4x4 blocks.
If SAD8x8 > sum of four SAD4x4 then 4x4 block-size is chosen, else 8x8 block-size choice is saved.

The ZME algorithm implementation design is as follow: For each position in the search window, 16 4x4 SADs are computed and stored. This will allow the computation of different sub macro blocks. In fact, The best 16x16 block size correspond to the minimum sum of 16 4x4 SADs. In a second step, when extracting the best 8x8 block size, there is no need to re-compute SAD's for all iterations since they were previously computed and stored. We have just to sum respective 4 SAD's to extract the best 8x8 block size.

When adopting this implementation method, we are avoiding the recomputation of SADs for each block size type tested. Unfortunately, we are consuming:

Mc= 16*32*W*H= 512 Kbyte of memory.
Where W=32 and L=32 are respectively width and high of the search window.

3. Proposed block size selection algorithm

A typical DSP-based VBS motion estimation algorithm is proposed in this paper to reduce motion estimation module complexity. The proposal of this algorithm is to extract, using the SAD criteria, the block-size introducing the minimum computational complexity. The SAD is computed as sum of absolute differences between every alternate pixel in the corresponding blocks of reference frame and current frame:

$$SAD\ (x,y) = \sum_{i=0}^{15}\sum_{j=0}^{15} \left| Ycurr\ (i,j) - Yprev\ (x+i, y+j) \right| \quad (1)$$

In this algorithm, only 16x16, 16x8, 8x16 and 8x8 block-size for motion estimation are considered. This approach will reduce extensively the computational complexity since we are eliminating other 3 modes, without affecting the overall visual video quality.
For better understanding, the proposed organigramme explains in detail the different algorithm steps:

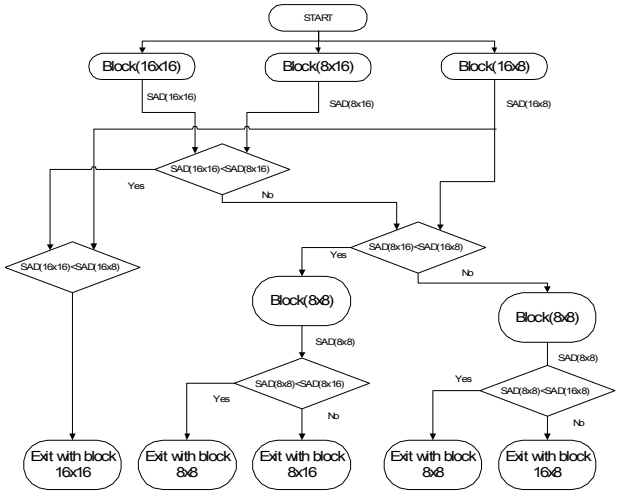

Fig.2. Fast block size selection algorithm (FBSA) organigramme.

SAD calculation contributes to the most of the computational complexity in motion estimation process.
When implementing the following approach, for each iteration, only four 8x8 SAD are computed and stored in the internal memory:
Only 128Kbyte of internal memory, against 512 Kbyte, is used to store SADs.
Only 4 SADs are computed in each iteration, against 16, which reduces computational complexity in motion estimation process.
The best SAD of 16x16 block is computed as sum of SADs of each of its four 8x8 block, the best 16x8 one is the minimum sum of 2 respective 8x8 SADs and the best 8x16 block size is the minimum sum of 2 respective 8x8

SADs. To extract the minimum 8x8 block size, we need just to determinate the minimum 8x8 SADs.

4. Experimental Results

The block-size algorithm implementation, discussed in the present paper has been developed on the TMS320C6416 digital signal processor (DSP) of Texas Instruments [10] based on LETI encoder [11]. JM12.1 H.264/AVC baseline encoder is also used to validate the quality performances.

To evaluate the fast block size selection algorithm (FBSA) impact on the encoding process, several analysis on three sequences test with different characteristics (Figure 3) are proposed in this section.

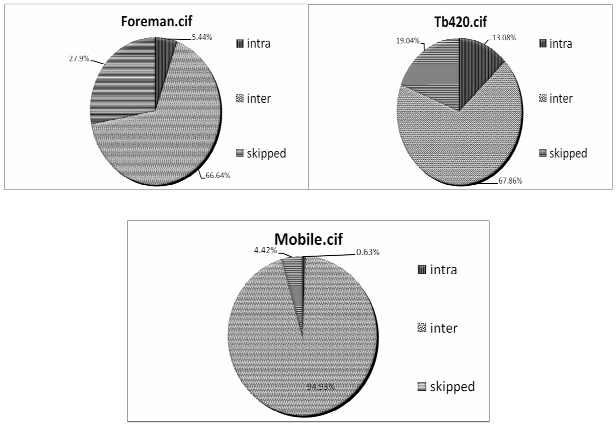

Fig.3. Percentages of ME(inter) intra and skip used in Foreman, Mobile and Tb420 CIF sequences.

Figures 4, 5 and 6 illustrate cycle's consumption of the motion estimation process for Foreman, Mobile and Tb420 at common intermediate format resolution (CIF 352x288) sequences for QP=38. As shown, the implementation of the new block size selection module on the TMS320C6416 DSP yields a significant reduction in terms of cycles account.

Recapitulation results proposed in Table 1 shows that FBSA yields to 44.3% improvement in speed performance compared to the ZME algorithm for Foreman sequence at Qp =38. The same performance is accomplished for Mobile and Tb420 sequence to get respectively 44.4% and 45.7% improvement.

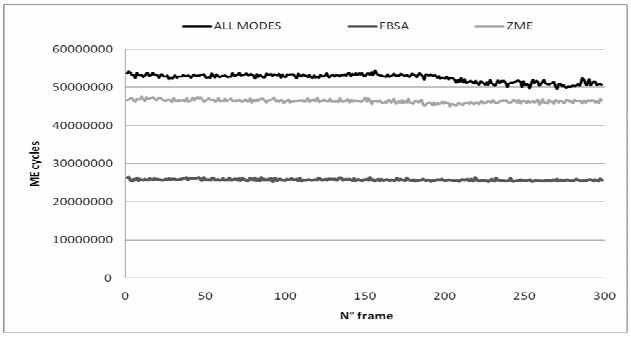

Fig 4. ME process Cycle's consumption for Foreman sequence at QP=38.

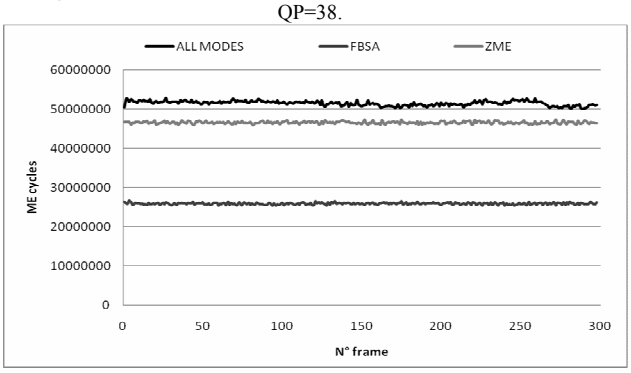

Fig. 5. ME process Cycle's consumption for Mobile sequence at QP=38.

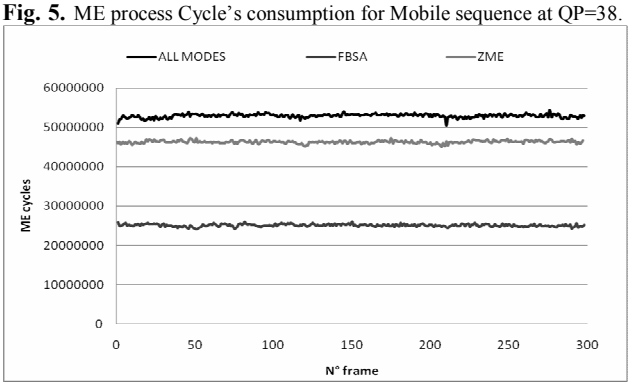

Fig. 6. ME process Cycle's consumption for Tb420 sequence at QP=38.

Table 1. Speed performance supported (ME cycles).

Sequences	Block size selection algorithm (Million cycles)			
	ALL MODES	ZME	**FBSA**	Speed FBSA/ZME %
Foreman	524.99	464.11	**258.13**	44.3
Mobile	515.11	465.92	**251.09**	44.4
Tb420	529.46	462.88	**259.02**	45.7

Additional analyses are proposed to provide the process implementation impact on both objective (PSNR) and

subjective (SSIM) [12] quality which are given by the following formula respectively:

$$PSNR = 20\log_{10}(\frac{255}{MSE}) \quad (1)$$

with $$MSE = \frac{\sum (S(i,j) - R(i,j))^2}{N^2} \quad (2)$$

S(i , j) are pixels of source frame, R(i , j) are pixels of reconstructed frame and N is number of pixels in one frame.

Table 2 provides a peak signal-to-noise ratio (PSNR) comparison between original, ZME and new block size selection approaches.

Table 2. PSNR quality performance.

Sequences	Block size selection algorithm		
	ALL MODES	ZME	**FBSA**
Foreman	30,73	30.67	**30.72**
Mobile	26.81	26.72	**26.79**
Tb420	29.50	29.47	**29.49**

Those results show that, from an objective point of view, there is a small quality loss between original and new proposed approach. Moreover, an interesting objective quality improvement is detected compared to the ZME method which support the complexity reduction effectiveness and confirm the reliability of the new approach.

Table 3 shows that these results are subjectively confirmed.

Table 3. SSIM quality performance.

Sequences	Block size selection algorithm		
	ALL MODES	ZME	**FBSA**
Foreman	0.9883	0.9882	**0.9883**
Mobile	0.9824	0.9821	**0.9824**
Tb420	0.9853	0.9852	**0.9853**

5. Conclusions

In this paper, we proposed Fast block size selection algorithm for inter frame coding in H.264/AVC implemented on the TMS320C6416 DSP. Proposed approach is a fast multi-block selection scheme focusing on 16x16, 16x8, 8x16 and 8x8 only, which can efficiently reduce the computational cost while achieving similar visual quality.

In addition to the 25% memory saving, the speed performance (ME cycles) and the quality effectiveness (SSIM and PSNR) analysis show that a major reduction in the computational complexity is obtained while maintaining almost the same quality level. Experimental results demonstrate a 45.75 % improvement in speed performance with no major loss in video quality.Further analysis will be conducting in future search to choose the appropriate search pattern.

References

[1] T. Wiegaid, G. Sullivan, G. Bjontegaard, and A. Luthra, “Overview of the H.264/AVC video coding standard”, IEEE Trans. Circuits Syst. Video Technol, pp. 560-576, July 2003.

[2] M. Horowitz, A. Joch, F. Kossentini, and A. Hallapuro, “H.264/AVC Baseline Profile Decoder Complexity Analysis”, IEEE Trans. Circuits Syst. Video Technol, pp. 704-716, July 2003.

[3] 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, Aug. 1994, pp. 438–442.

[4] S. Zhu and K.-K. Ma, "A new diamond search algorithm for fast block-matching motion estimation", IEEE Trans. Image Processing, vol. 9, Feb. 2000, pp.287–290.

[5] M. Ali Ben AYED, A. SAMET, and N. MASMOUDI , "Toward an optimal block motion estimation algorithm for H.264/AVC", International Journal of Image and Graphics (IJIG), Vol. 7, No. 2, pp. 1–18, 2007.

[6] Y. K. Tu, J. F. Yang, and M. T. Sun, "Fast Variable-size Block Motion Estimation Using Merging Procedure with an Adaptive Threshold", IEEE International Conference on Multimedia and Expo, Baltimore, July 2003, pp. II-789-792.

[7] Z. Zhou, M. T. Sun, and Y. F. Hsu, "Fast variable block-size motion estimation algorithms based on merge and split procedures for H.264/MPEG-4 AVC", IEEE International Symposium on Circuits and Systems, ISCAS, Vancouver, British Columbia, Canada, May 23-26, 2004.

[8] Y. Jiang, S. Li, S. Goto, "A Low Complexity Variable Block Size Motion Estimation Algorithm for Video Telephony Communication", 47th IEEE International Midwest Symposium on Circuits and Systems, July 2004, pp. II-465 - II-468.

[9] Evaluation sheet for motion estimation, JVT Test Model Ad Hoc Group,Feb. 19, 2003. Draft version 4.

[10] Texas Instruments, “TMS320C6000 CPU and Instruction Set Reference Guide”, Literature Number: SPRU189F October 2000.

[11] Circuit and System Group,LETI Laboratory, http://www.csgroup.tunet.tn.

[12] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity", IEEE Transactions on Image Processing, vol. 13, no. 4, Apr. 2004.

Haithem CHAOUCH was born in Riyadh, Arabia Saudi, in 1982. He received his degree in Electrical Engineering, from Sfax National School of Engineering (ENIS), Tunisia, in 2006. Currently, he is working toward the Master, in Electronic Engineering at the ENIS. He is a member of the Sfax Laboratory of Electronics and Information Technology. His research interests include digital signal processing, image and video coding with emphasis on H.263 and H.264 standards and motion estimation algorithms.

Imen WERDA was born in Sfax, Tunisia, in 1978. She received his degree in Electrical Engineering and his DEA in Electronic Engineering from Sfax National School of Engineering (ENIS), Tunisia, in 2003 and 2004 respectively.
In 2004, she joined the Ubvideo Tunisia Inc. as an R&D engineer.
In 2006, she joined the High Institute of technology of Sousse, Tunisia, as an Assistant Technologue.
She is a member of the Sfax Laboratory of Electronics and Information Technology. Her current research interests include DSP and hardware implementation of H.264 video coding standard for video conference applications.

Amine SAMET was born in Tunis, Tunisia, in 1977. He received his degree of Diplôme d'Ingénieur in Electrical Engineering, his DEA in Electronic Engineering, and his Ph. D. in Electronic Engineering at Sfax National School of Engineering (ENIS), Tunisia, in 2001, 2002, and 2006 respectively.
In 2002, he joined the High Institute of Electronics and Communication of Sfax, Tunisia, as an Assistant.
He is a member of the Sfax Laboratory of Electronics and Information Technology. His research interests include digital signal processing, image and video coding with emphasis on H.263 and H.264 standards and motion estimation algorithms.

Mohamed Ali BEN AYED was born in Sfax, Tunisia, in 1966. He received his B.S. degree in computer engineering from Oregon State University in 1988, his M.S. degree in electrical engineering from Georgia Institute of Technology in 1990, his DEA degree and Ph. D. in electronic engineering from Sfax National School of Engineering, Tunisia, in 1998 and 2004 respectively.
He was a co-fonder ob Ubvideo Tunisia, a wholly owned subsidiary of Ubvideo Inc. in 2003.
He is currently an Assistant Professor at High Institute of Electronics and Communication of Sfax.
His current research interests include DSP and hardware implementation of digital algorithms for multimedia services.

Nouri MASMOUDI was born in Sfax, Tunisia, in 1955. He received the electrical engineering degree from the Faculté des Sciences et Techniques de Sfax, Tunisia, in 1982, the DEA degree from the Institut National des Sciences Appliquées de Lyon and Université Claude Bernard de Lyon, France in 1982. From 1986 to 1990, he prepared his thesis at the laboratory of Power Electronics (LEP) at the Ecole Nationale d'Ingénieurs de Sfax (ENIS). He then received the 'Thèse de 3 ème Cycle' at the Ecole Nationale d'Ingénieurs de Tunis (ENIT), Tunisia in 1990.
From 1990 to 2000, he was an assistant professor at the electrical engineering department at the ENIS. Since 2000, he has been an Associate Professor and head of the group 'Circuits and Systems' of the Sfax Laboratory of Electronics and Information Technology. Currently, he is responsible for the Electronic Master Program at ENIS.
His research activities have been devoted to several topics: power system, embedded systems, DSP and FPGA implementation of digital algorithms, and information techno