AN ADAPTIVE COMPUTATION-AWARE ALGORITHM FOR MULTI-FRAME VARIABLE BLOCK-SIZE MOTION ESTIMATION IN H.264/AVC

Mariusz Jakubowski and Grzegorz Pastuszak
Institute of Radioelectronics, Warsaw University of Technology, 15/19 Nowowiejska Str., Warsaw, Poland

Keywords: Video compression, Motion estimation, Computational awareness.

Abstract: Block-matching motion estimation (BME) is the most computationally expensive process in every video codec. The algorithm proposed in this paper takes into account almost all key elements of BME including integer-pixel ME (IPME), sub-pixel ME (SPME), variable block-size ME (VBSME) and multiple reference frame ME (MRFME). The algorithm is developed by adding MRFME method to the multi-path adaptive computation-aware ME strategy (MPS) introduced in our previous papers. The algorithm implemented in the H.264/AVC reference software achieves comparable results as the fast full search (FFS) method within less than 3% of execution time required by FFS.

1 INTRODUCTION

Block-matching motion estimation (BME) is an efficient and popular technique for reduction of temporal redundancy within video sequences adopted in various video coding standards, such as ITU-T H.26x and ISO/IEC MPEG-1, -2, and –4 (ITU-T, 2003), (Huang *et al.*, 2006). It is also the most computationally expensive element of video coders. BME always involves the integer-pixel motion estimation (IPME) and usually the sub-pixel motion estimation (SPME) with a half-pixel accuracy. In the H.264/AVC standard, several improvements have been introduced regarding BME (ITU-T, 2003):

- variable block-size motion estimation (VBSME)
- quarter-pixel accuracy motion estimation
- multiple reference frame motion estimation (MRFME)
- weighted prediction

To decrease the computational burden related to each of these elements many fast algorithms have been developed (Huang *et al.*, 2006 and following). In this paper, the multi-path adaptive computation-aware motion estimation strategy (MPS) described in our previous papers (Jakubowski and Pastuszak, 2007), (Jakubowski and Pastuszak, 2008) embedding the MRFME technique is presented as the solution which takes into account almost all these aspects of BME in H.264/AVC (except weighted prediction). The proposed scheme utilizes the efficiency of MPS to determine an optimal reference frame (RF) on the early stage of the search process and is characterized by the ability to adapt to computation-variant conditions (computational awareness) and to achieve similar results as the exhaustive search using only a fraction of execution time, required by the full search (FS) scheme.

The rest of the paper is organized as follows. In Section 2, the MPS algorithm is described. In Section 3, the proposed MRFME method is introduced. Experimental results are presented in Section 4. Section 5 gives a conclusion.

2 MULTI-PATH STRATEGY

Multi-path adaptive computation-aware strategy (MPS) is the motion estimation (ME) algorithm developed and presented in a few our previous papers. In this section all the key elements of this strategy are described.

2.1 Allocation of Computational Resources

Number of search points (SPs) available for the whole frame is divided into two parts. The first one

Jakubowski M. and Pastuszak G. (2009).
AN ADAPTIVE COMPUTATION-AWARE ALGORITHM FOR MULTI-FRAME VARIABLE BLOCK-SIZE MOTION ESTIMATION IN H.264/AVC.
In *Proceedings of the International Conference on Signal Processing and Multimedia Applications*, pages 122-125
DOI: 10.5220/0002231401220125

provides exactly the same number of SPs for each macroblock (MB) for the basic computation. The second one provides some extra points for each MB in proportion to the initial sum of absolute differences (InitSAD) in the starting SP divided by the average minimum SAD of previously processed MBs (AvgMinSAD). The bigger this ratio is, the more extra points are allocated to the MB. If some SPs are left after a given MB processing they are added to the computational pool of the next MB.

2.2 Starting Search Point Selection

The starting SP is chosen from the prediction set which contains motion vectors (MVs) of left, left-upper, upper, and right-upper neighbors, zero motion point, and the co-located block in the previous frame. The vector which gives the smallest SAD is selected.

2.3 Adaptive Search Strategy Selection

The strategy used in the first step is selected on the basis of a few factors: the number of available SPs, the ratio of InitSAD to AvgMinSAD, and the standard deviation of neighboring MVs around their median. If the standard deviation is greater than 5, the high-motion activity is assumed and three step search (TSS) (Koga *et al.*, 1981) is used. Otherwise, either diamond search (DS) (Zhu and Ma, 1997) or kite-cross-diamond search (KCDS) (Lam *et al.*, 2004) is selected on the basis of the ratio of InitSAD to AvgMinSAD and the number of available SPs. All values of the parameters which affect the strategy selection where adjusted experimentally (Jakubowski and Pastuszak, 2007).

If the amount of resources is sufficient, surroundings of all the points from the prediction set are investigated using TSS. This phase of search is called the multi-path search. If after the multi-path search some resources are still available they can be utilized in the last step by the full search.

2.4 Variable Block-Size and Sub-Pixel ME

Since MVs of all partition modes are usually highly correlated, the probability of finding the optimal MV in the close neighborhood of MV for mode 16×16 is on average larger than 80% (Jakubowski and Pastuszak, 2008). Therefore, similarly like in the fast full search (FFS) method adopted in the H.264/AVC reference software, all modes are checked in parallel in each point of the search path for mode 16x16. Firstly, all SADs of 4×4 blocks are computed and then reused to compose SADs for other modes. However, in FFS, after IPME, each mode gets its own search center for SPME which leads to a substantial increase of computational cost. In our approach, search centers for integer-pixel, half-pixel, and quarter-pixel ME for all modes are the same as for mode 16×16 and the best MV for each mode is selected from among SPs checked for mode 16×16. It makes it possible to check all modes in parallel also during SPME with a relatively small coding efficiency degradation.

3 MULTIPLE REFERENCE FRAME ME

The goal of the MRFME method added to the MPS algorithm is to select the optimal reference frame (RF) at the early stage of ME process. In the MPS algorithm the test of the prediction set and the first strategy are the most crucial elements for the algorithm performance. These two steps require about 30 SPs/MB/Frame on average (including SPME) and give over 90% contribution to the final outcome. Thus, it has been assumed that they are sufficient to determine the optimal RF. Initially, it has been supposed that SPME will not be necessary to select the optimal RF, however, it turned out that SPME has a significant influence on the optimal RF selection and even after half-pixel ME about 20% of selected frames is inconsistent with the optimal ones.

Since the probability that the nearest RF is the optimal one is in general much greater than 50% (Huang *et al.*, 2006), this frame takes priority over the others and gets more resources in the first step. A simplified flowchart of the method is shown in Fig. 1. At the beginning, the prediction set and the first strategy are checked in the closest RF up to the quarter-pixel accuracy. In the next step, the prediction set is checked in the remaining frames up to the integer pixels. If the cost for the best point is smaller than in the previous frame – the first strategy is also checked up to the quarter pixels. This way, the optimal frame is selected and ME is continued in this frame until it is finished or computational resources are exhausted. If some resources are still available, the ME process can be continued in the remaining frames. Additionally, the best point found in a given frame is included to the prediction set of the next frame.

Table 1: Reduction of the execution time and differences in PSNR with reference to FFS with five RFs.

Algorithm	FFS 1 RF		MPS 25 SPs/MB, 5 RFs		MPS 150 SPs/MB, 5 RFs	
Sequence	RET [%]	maxPSNRdiff [dB]	RET [%]	maxPSNRdiff [dB]	RET [%]	maxPSNRdiff [dB]
Mobile	80.00	1.15	99.60	1.10	97.58	0.30
Football	80.00	0.04	99.67	1.02	98.00	0.30
Foreman	80.00	0.80	99.62	1.15	97.74	0.40
Crew	80.00	0.08	99.89	0.25	99.33	0.07
Harbor	80.00	0.12	99.89	0.20	99.31	0.06
Soccer	80.00	0.13	99.89	0.57	99.32	0.15

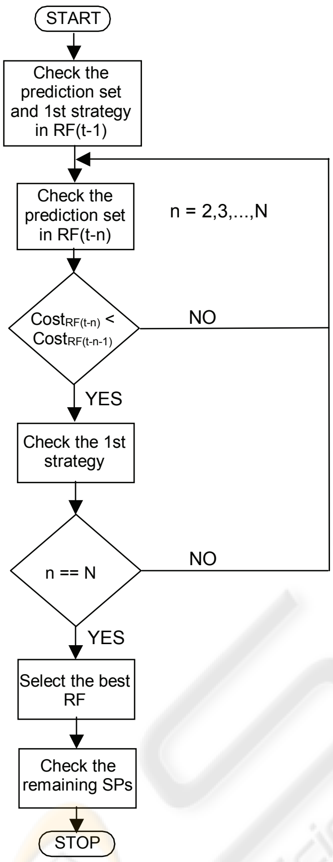

Figure 1: The flowchart of proposed MRFME method.

4 EXPERIMENTAL RESULTS

The algorithm is implemented in the H.264/AVC reference software (JM12) and its performance is compared with FFS with one and five RFs. For MPS, always five RFs are used, however, with two different values of the SPs/MB parameter: 25 and 150, regardless of the spatial resolution of the sequence. In the experiments, three CIF (Foreman, Football, and Mobile) and three 4CIF (Crew, Harbor and Soccer) sequences, 150 frames each, are used. Search range is set at ±15 and ±31 points for CIF and 4CIF sequences, respectively. GOP structure is I-P-P-P. Rate-distortion curves for selected sequences are presented in Fig. 2 and 3. Values of the reduction of execution time and the maximal differences in PSNR with reference to FFS with five RFs are placed in Table 1. RET represents the percentage of reduction of execution time, and maxPSNRdiff represents the maximal difference in PSNR in dB. This difference for MPS with 150 SPs/MB is never greater than 0.4 dB (Foreman). The magnitude of this difference depends mainly on correlation between MVs of different modes. The more they are correlated, the smaller this difference is. Note that for most of the sequences, except Foreman and Mobile, the gain introduced by

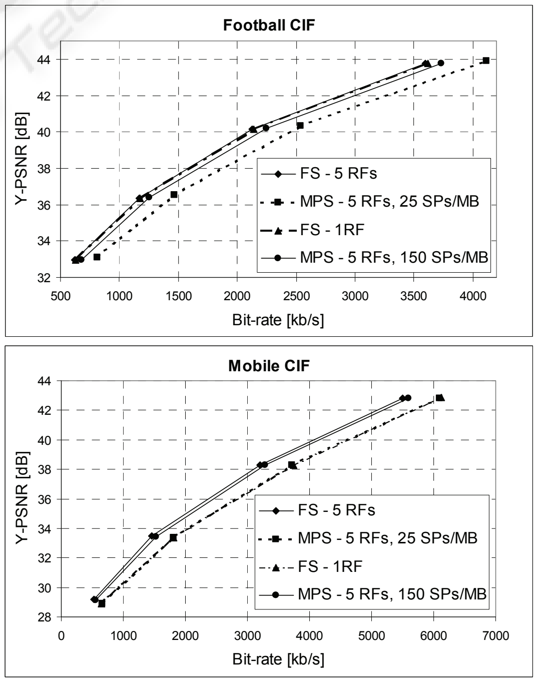

Figure 2: Rate-distortion curves for CIF sequences.

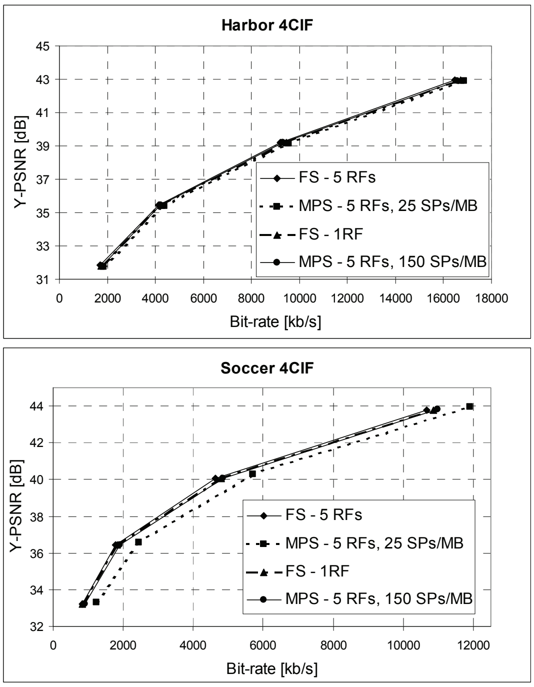

Figure 3: Rate-distortion curves for 4CIF sequences.

MRFME is relatively small. It is especially true for the sequences with low and high motion activity (Crew and Football). In such sequences, the nearest frame is generally the best choice since its resemblance to the next frame is the biggest.

Obtained reduction of execution time is significant especially for 4CIF sequences where exceeds 99% since MPS uses totally 150 SPs/MB both for IPME and SPME in all RFs. Even in case of 4CIF sequences the difference in PSNR remains small without increasing computational resources which demonstrates the efficiency of the MPS algorithm and its insensitivity on changes of resolution.

5 CONCLUSIONS

The adaptive computation-aware MPS strategy in conjunction with MRFME method presented in this paper creates the solution which takes into account almost all major aspects of BME in the H.264/AVC standard. The algorithm was implemented in the JM 12.0 H.264/AVC reference software and compared with FFS method with one and five RFs. Tests showed that within less than 3% of execution time required by FFS, MPS is able to achieve similar results. Additionally, computation-aware feature allows the algorithm to accomplish almost exactly the same results as the exhaustive search if the computational recourses are sufficient.

ACKNOWLEDGEMENTS

The work presented was developed within activities of VISNET II, the European Network of Excellence, (http://www.visnet-noe.org), founded under the European Commission IST 6FP programme.

REFERENCES

ITU-T Recommendation H.264 and ISO/IEC 14496-10 MPEG-4 Part 10, Advanced Video Coding (AVC), 2003.

Huang, Y.W., Chen, C.Y., Tsai, C.H., Shen C.F., and Chen L.G., 2006. Survey on Block Matching Motion Estimation Algorithms and Architectures with New Results. *J. VLSI Signal Proc.*, vol. 42, pp. 297-320.

Koga, T., Iinuma, K., Hirano, A., Iijima, Y., and Ishiguro, T., 1981. Motion Compensated Interframe Coding for Video Conferencing. In *Proc. Nat. Telecom. Conf.*, pp. C9.6.1–C9.6.5.

Zhu, S. and Ma, K. K., 1997. A New Diamond Search Algorithm for Fast Block Matching Motion Estimation. In *Proc. IEEE Int. Conf. Image Processing (ICIP'97)*, pp. 292–296.

Lam, C. W., Po, L. M., Cheung, C. H., 2004. A Novel Kite-Cross-Diamond Search Algorithm for Fast Block Matching Motion Estimation. In *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS'04)*, vol. III, pp. 729–732.

Jakubowski, M., Pastuszak, G., 2007. Multi-Path Adaptive Computation-Aware Search Strategy for Block-Based Motion Estimation. In *Proc. IEEE EUROCON 2007, The International Conference on Computer as a Tool*, pp. 175-181, 2007.

Jakubowski, M., Pastuszak, G., 2008. A Hardware-Oriented Variable Block-Size Motion EstimationMethod for H.264/AVC Video Coding. In *Proc. 12th AES Symposium New Trends in Audio and Video (NTAV) 2008*, pp. 151-156.

Huang, Y. W., Hsieh, B. Y., Chien, S. Y., Ma, S. Y., and Chen, L. G., 2006. Analysis and Complexity Reduction of Multiple Reference Frames Motion Estimation in H.264/AVC. *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, no. 4, pp. 507-522.