RELIABLE SEARCH STRATEGY FOR BLOCK MOTION ESTIMATION BY MEASURING THE ERROR SURFACE

Yui-Lam Chan and Wan-Chi Siu

Department of Electronic and Information Engineering
The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

ABSTRACT

The conventional search algorithms for block matching motion estimation reduce the set of possible displacements for locating the motion vector. Nearly all of these algorithms rely on the assumption: the distortion function increases monotonically as the search location moves away from the global minimum. Obviously, this assumption essentially requires that the error surface be unimodal over the search window. Unfortunately, this is usually not true in real-world video signals. In this paper, we formulate a criterion to check the confidence of unimodal error surface over the search window. The proposed Confidence Measure of Error Surface, *CMES*, would be a good measure for identifying whether the searching should continue or not. It is found that this proposed measure is able to strengthen the conventional fast search algorithms for block matching motion estimation. Experimental results show that, as compared to the conventional approach, the new algorithm through the CMES is more robust, produces smaller motion compensation errors, and requires simple computational complexity.

1. INTRODUCTION

Motion estimation is an essential component of all modern video coding standards [1-2]. It is included in these standards to reduce the redundancy between successive frames of a video sequence. The method adopted to estimate the motion between frames is the block matching algorithm (BMA) [3-10]. For the full search algorithm (FSA) of BMA, a matching criterion between every block in a search window from the previous frame and the current block is calculated. The most commonly used matching criterion is the mean absolute difference (MAD) [7]. The FSA evaluates the MAD at all possible locations of the search window to find the optimal motion vector. Hence it is able to find the best-matched block which guarantees to give the minimal MAD. On the other hand, it also demands an enormous amount of computation. Thus a number of fast search algorithms [4-10] have been proposed, which seek to reduce the computation time by searching only a subset of the eligible candidate blocks. These fast block motion estimation algorithms include the n-Step Hierarchical Search algorithm (n-SHS) [7], the conjugate directional search algorithm [8], the new three-step search algorithm [9], the block-based gradient descent search algorithm (BBGDS) [10] and many variations. These algorithms reduce the number of computations required by calculating the MAD matching criterion at positions coarsely spread over the search window according to some pattern and then repeating the procedure with finer resolution around the position with the minimum MAD found from the preceding step. Nearly all of these algorithms rely on the assumption: the MAD distortion function increases monotonically as the search location moves away from the global minimum [4]. Obviously, this assumption essentially requires that the MAD error surface be unimodal over the search window. Unfortunately, this is usually not true in real-world video signals. As a consequence, the minimum MAD found by these methods is frequently higher than that is produced by the FSA. To prevent this, a simple but perhaps the most reliable strategy is to measure the confidence of unimodal error surface over the search window. In this paper, the new Confidence Measure of Error Surface, CMES, is proposed and it becomes a good criterion for determining the continuity for the searching in the block matching motion estimation algorithm. The new algorithm developed in this paper is based on the verification of this newly defined confidence measure, that is used to identify whether the searching would continue or not.

The rest of this paper is organized as follows. In Section 2, we present an in-depth study on the MAD error surface. Based on the studies in Section 2, we formulate the proposed confidence measure into the search window and propose a fast search algorithm through the confidence measure for block matching motion estimation in Section 3. In Section 4, some analysis on the algorithm's complexity and performance will be presented. Finally, conclusions are drawn in Section 5.

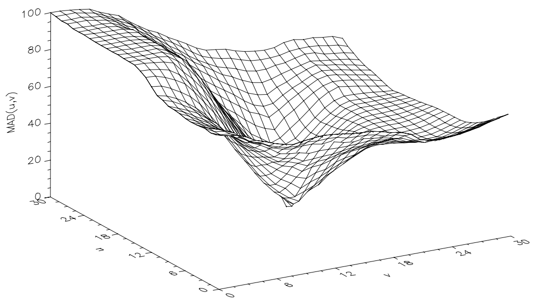
2. THE MAD ERROR SURFACE

Suppose that the maximum motion in the vertical and horizontal directions is $\pm W$, there are thus $(2W+1)^2$ candidates in total to be checked if the full search method is used, each corresponding to a checking point in the search window. The MAD values resulted from these checking points form an error surface

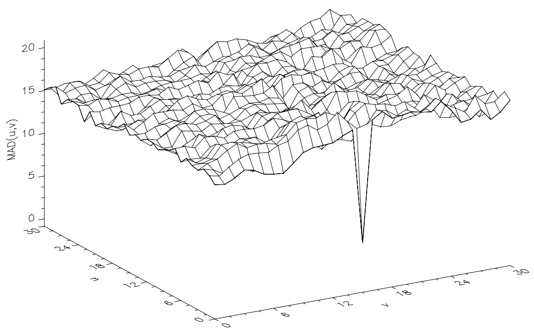
$$MAD(u,v) = \sum_{i=0}^{N-1}\sum_{j=0}^{N-1}\left|I_t(i,j) - S_{t-1}(i+u,\, j+v)\right| \qquad (1)$$

where the block size is taken as $N \times N$, (u,v) denotes the position of the candidate motion vector, and $I_t(\cdot,\cdot)$ and $S_{t-1}(\cdot,\cdot)$ refer to the blocks in the current frame(t^{th} frame and in the reference frame ($(t-1)^{th}$ frame) that are to be compared.

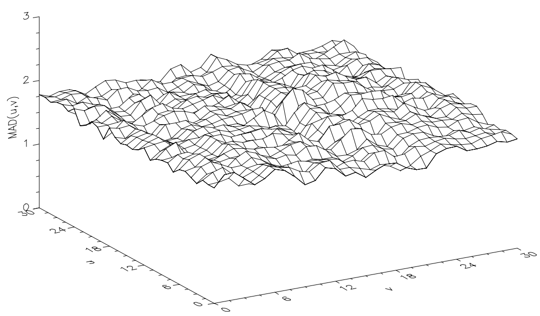
The statistical behaviour of the MAD error surface has a significant impact on the performance of the fast search algorithm for block matching motion estimation. For the surface as shown in Fig. 1(a), the MAD error decreases monotonically as the search location moves toward the global minimum value. It implies that a simple fast search algorithms such as the n-SHS [9] and the BBGDS [10] would require a small number of searches to determine the global optimum position for this block. For the surface as shown in Fig. 1(b), it contains a large number of local minima. Almost all conventional fast algorithms have explicitly or implicitly assumed [4] that the error surface is unimodal over the search window. As a consequence, it is unlikely that the previously described fast search algorithms would converge to the global minimum. In other words, the search would easily be trapped at a local minimum. For the surface in Fig. 1(c), there is no need to find the global minimum position since any of the local minimum positions will correspond to a satisfactory prediction block as *E(u,v)* is uniformly small. The new algorithm presented in this paper explores the property of this important behaviour in order to optimize the performance of the motion estimation.



(a)



(b)



(c)

Figure 1. MAD Error Surface for three different blocks.

3. RELIABLE SEARCH ALGORITHM THROUGH THE CMES

The search algorithm presented in this paper can best be described as an extension of the Block-Based Gradient Descent Search (BBGDS) algorithm [10]. Let us recall that in the first step of the BBGDS algorithm, search is done only around the center checking point. If the optimum is found at the center, the procedure stops. Otherwise, further search is done around the point where the minimum has just been found. The procedure continues until the winning point is a center point of the checking block (3×3 checking points) or the checking block hits the boundary of the predefined search range [10]. The procedure is illustrated in Fig. 2, where the motion vector (3,-4) is found. Of course, the BBGDS algorithm relies on the assumption that the MAD measure decreases monotonically as the search position moves closer to the optimum position. It can easily be trapped into the local minimum if the error surface is similar to Fig. 1(b).

Let us use Fig. 3 to give a clearer account for this phenomenon. In Fig. 3, it shows a nonunimodal surface due to many reasons such as the aperture problem, the textured (periodical) local image content, the inconsistent block segmentation of moving object and background, the luminance change between frames, etc. In the first step of the BBGDS algorithm, the center point in the checking block wins. It will stop the searching process and a local minimum will be found. However, it is seen that the global minimum is located at the far side of the winning point and the MAD value of the winning point is significantly larger than that of the global minimum. It will degrade the quality of the motion-compensated prediction frame. For the new BBGDS algorithm, a similar procedure is conducted. In order to maximize the possibility for finding the global minimum in the situation like Fig. 1(b), it is necessary to determine whether the winning center of the current checking block be identified as the "final winner". Thus, a Confidence Measure of Error Surface (CMES) is proposed to prevent an unsuitable termination of the search being misled by insufficient information. In other words, the CMES is used to determine the continuation of the search by enlarging the checking block according to the superiority of the best-matched center position to others in the current checking block. Let us define the CMES as follows:

$$CMES = \frac{\sum_{\substack{i=-l \\ i\neq 0}}^{+l} \sum_{\substack{j=-l \\ j\neq 0}}^{+l} (MAD(u+i, v+j) - MAD_{\min}(u,v))}{\sum_{\substack{i=-l \\ i\neq 0}}^{+l} \sum_{\substack{j=-l \\ j\neq 0}}^{+l} MAD_{\min}(u,v)} \qquad (2)$$

where l is the size of the checking block; $E_{min}(u,v)$ and $E(u+i,v+j)$ are the smallest and other values of the MAD of the checking block, respectively. Values of the CMES can reflect the statistical behaviour of the error surface in the checking block. If the CMES is close to *0*, it means that it is insufficient to make sure that this center point is a winner. That is, the best-matched center position in the checking block is probably a local minimum, and hence the size of the checking block, *l*, is increased to further evaluate the behaviour of this enlarged error surface, as depicted in Fig. 4. On the other hand, if the CMES is far away from *0*, it indicates that the center point is probably located at the global minimum.

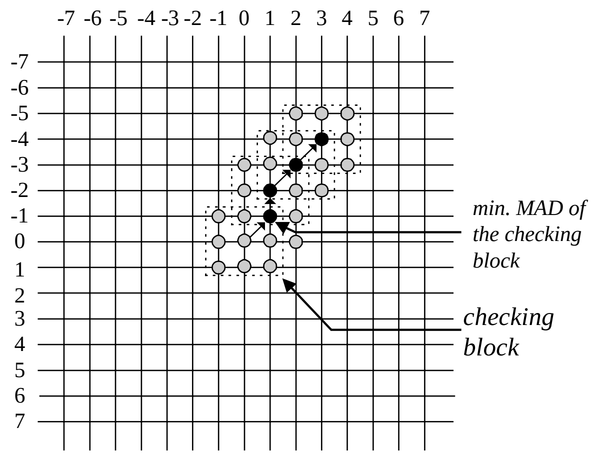


Figure 2. Example of the BBGDS search procedure, where motion vector (3, -4) is found.

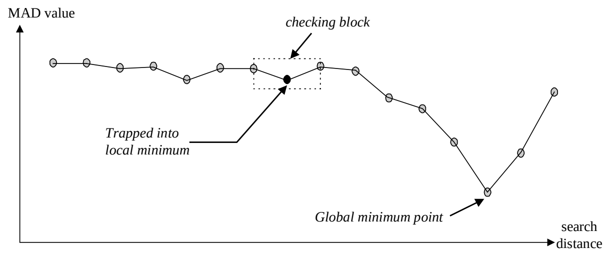


Figure 3. A nonunimodal error surface sampled by checking block.

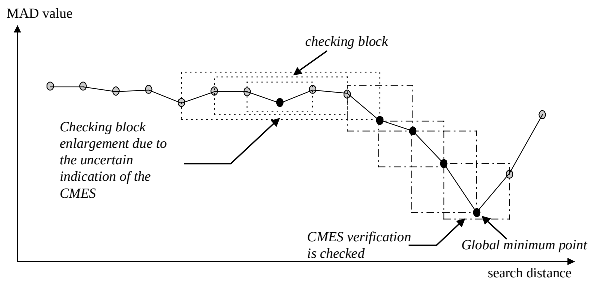


Figure 4. Reliable search through the CMES.

Now by using this CMES, more search positions are allowed in search windows which contain more local minimum values for error surface than in search windows which have monotonically decreasing values of error surface. However, there is still some inefficient use of the search positions. Consider the search window with the MAD error surface shown in Fig. 1(c). The modified BBGDS will find many local minima in this search window, the value calculated for the confidence measure will be small, and consequently, if only the CMES is used, many search positions will be allowed for this search window. It can be seen however that the value of the MAD at all the local minimum positions found will be very small, and hence, any of these positions will correspond to a good prediction for the current block. Therefore, a MAD threshold detector is needed to limit the number of search positions in the search window where the MAD value at the local minimum positions has already reached an acceptably small value.

According to the above discussion, a reliable solution to terminate the search process in the BBGDS is proposed. The details are given below:

If the minimum MAD point in the search step occurs at the center of checking block and its value is smaller than an acceptable error, MAD_{thr}, stop the search. Let us refer this as *error-acceptable stop*.

If the minimum MAD point in the search step occurs at the center of checking block and the value of its CMES is larger than a confident threshold, α, stop the search. This refers to as *CMES verification stop.*

The block diagram of the new BBGDS is shown in Fig. 5. Clearly, if the CMES verification stop does not occur, the checking block is enlarged as shown in Fig. 4, and it continues this CMES verification of the new checking block until the CMES is larger than α or the minimum MAD point is not in the center. Note that, in the latter case, the size of checking block has to be reset to *1*.

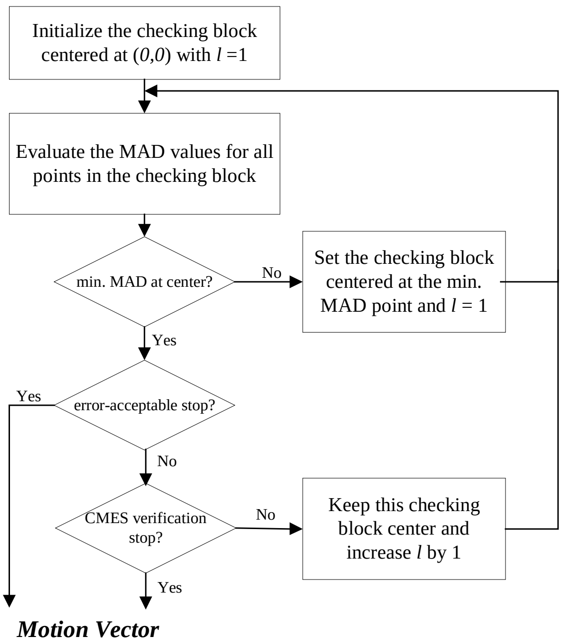


Figure 5. Block diagram of the new BBGDS algorithm

4. SIMULATION RESULTS

The algorithm introduced in this paper has been developed in accordance with the statistical behaviour of error surface. The performance of the proposed algorithm has been tested for a large variety of real image sequences, including "Table Tennis" and "Football". Results of the performance of the block motion vector estimation of the new BBGDS through the CMES and some conventional methods are compared in terms of quality and computational complexity. Parameters MAD_{thr} and α for the stopping criteria of our new BBGDS were set to *3000* and *0.3* respectively. The maximum allowable displacement in both the *x* and *y* directions was set to *15*, and a block size of 16×16 has been used. We have also used the Mean Square Error (MSE) per pixel as the measure of performance.

Fig. 6 shows the results of the MSE of the motion-compensated prediction frames together with some traditional approaches for the comparison. In Fig. 6, there is a great increase in prediction error of the n-SHS and the conventional BBGDS as compared with that of the FSA. It is because the probability of occurring the situation like Fig. 1(b) is more often in the fast moving sequences. This situation makes an inappropriate choice in early steps of the n-SHS, and the unreliable stop in searching of the conventional BBGDS implies that such kind of algorithms are more easily to be trapped in a local minimum. However, our new BBGDS can resolve the misleading stop of the searching by evaluating the confidence measure of error surface, *CMES*. As shown in Fig.6, the new BBGDS through the CMES is significantly better than that of the n-SHS and the conventional BBGDS. Also, we can see that the MSE performance of our approach is very close to the FSA. From Table 1, it is shown that the new BBGDS requires only 2.1% to 2.5% complexity of the FSA. It is much better than the famous n-SHS and has a slight increase in complexity as compared to the conventional BBGDS.

5. SUMMARY

In this paper, we have presented a thorough study on the error surface behaviour of motion vector of video signals. Then, we propose a new measurement for the fast search algorithm design and performance comparison. It has been shown that the Confidence Measure of Error Surface (CMES) is a criterion for measuring the certainty to stop the searching process. As the unimodal error surface is checked in our approach, the searching through the CMES is usually nonuniform so that it is able to best adapt to the statistical behaviour of a particular video sequence. This criterion naturally makes robust and fast motion estimation possible. We have tested the proposed CMES with the BBGDS and found that, a speed-up of about *40-50* times is achievable as compared with the Full Search Algorithm, and both algorithms give similar performance.

6. REFERENCES

[1] ITU-T H.263, "Video coding for low bit rate communication," Mar. 1996.

[2] ISO/IEC 13818-2, "Information technology -- Generic coding of moving pictures and associated audio information: Video," 1996.

[3] Y.L. Chan and W.C. Siu, "New adaptive pixel decimation for block motion vector estimation," IEEE Trans. on Cir. and Sys. for Video Tech., vol.6, no.1, pp.113 -118, Feb. 1996.

[4] J.R. Jain, and A.K. Jain, "Displacement measurement and it's applications in interframe coding," IEEE Trans. on Comm., vol.29, no.12, pp. 1799-1808, July 1981.

[5] Y.L. Chan and W.C. Siu, "Adaptive multiple-candidate hierarchical search for block matching algorithm," Elect. Letters, vol. 31, no. 19, pp.1637-1639, Sept. 1995.

[6] Y.L. Chan and W.C. Siu, "On block motion estimation using a novel search strategy for an improved adaptive pixel decimation," Journal of Visual Comm. and Image Rep. vol. 9, no. 2, pp. 139-154, June 1998.

[7] T. Koga, K. Iinuma, A. Hirano, Y. Iijima, and T. Ishiguro, "Motion compensated interframe coding for video conferencing," Proceedings of the National Telecomm. Conference, pp.G.5.3.1-5.3.5, Nov 29-Dec 3, 1981.

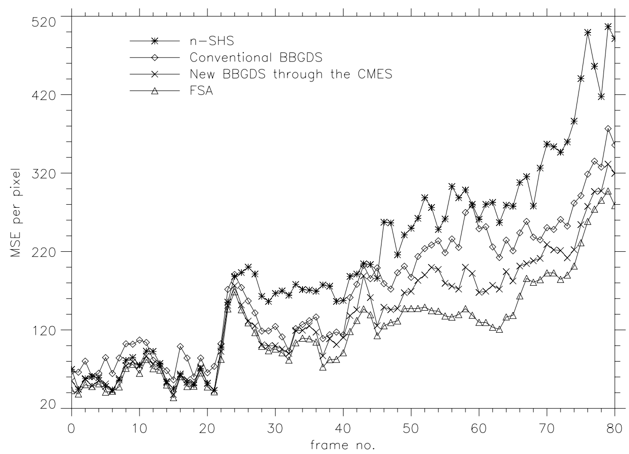
[8] R. Srinivasan and K. R. Rao, "Predictive coding based on efficient motion estimation," IEEE Trans. on Comm., vol. 33, no. 9, pp. 1011-1015, Sept. 1985.

[9] R. Li, B. Zeng, M.L. Liou, "A new three-step search algorithm for block motion estimation," IEEE Trans. on Cir. and Sys, for Video Tech., vol.4, pp. 438-442, Aug. 1994.

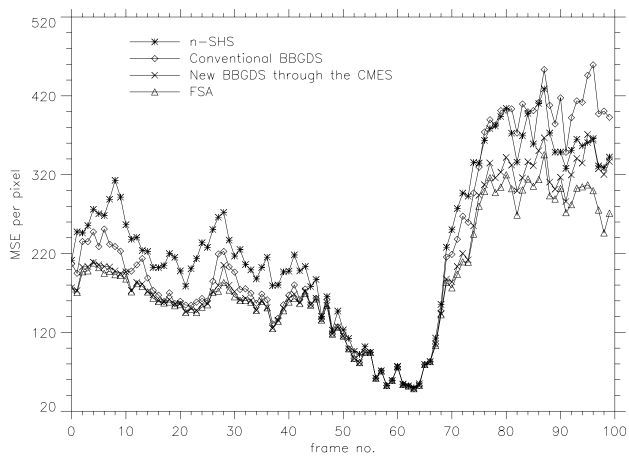
[10] L.K. Liu and E. Feig, "A block-based gradient descent search algorithm for block motion estimation in video coding," IEEE Trans. on Cir. and Sys. for Video Tech., vol.6, no.4, pp. 419-421, Aug. 1996.

Table 1. The complexity of the algorithms.

Algorithms	Table Tennis	Football
FSA	100%	100%
n-SHS	3.19%	3.19%
Conventional BBGDS	1.42%	1.39%
New BBGDS through the CMES	2.13%	2.49%



(a) Table Tennis



(b) Football

Figure 6. MSE produced by different algorithms for image sequences, the "Table Tennis" and the "Football".