# A Feature-Assisted Search Strategy for Block Motion Estimation

*Yui-Lam Chan and Wan-Chi Siu*
*Centre for Multimedia Signal Processing*
*Department of Electronic and Information Engineering*
*The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong*
*Tel: (852)27666229 Fax: (852)23628439 email: enwcsiu@polyu.edu.hk*

## Abstract

*Block motion estimation using the exhaustive full search is computationally intensive. Previous fast search algorithms tend to reduce the computation by limiting the number of locations to be searched. 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. Unfortunately, this is usually not true in real-world video signals. However, we can reasonably assume that it is monotonic in a small neighbourhood around the global minimum. Consequently, one simple, but perhaps the most efficient and reliable strategy, is to put the checking point as close as possible to the global minimum. In this paper, some image features are suggested to locate the initial search points. Such a guided scheme is based on the location of some feature points. After a feature detecting process was applied to each frame to extract a set of feature points as matching primitives, we studied extensively the statistical behaviour of these matching primitives and found that they are highly correlated with the MAD error surface of real-world motion vectors. These correlation characteristics are extremely useful for fast search algorithms. The results are robust and the implementation could be very efficient.*

## 1. Introduction

Motion estimation plays an important role in today's video coding and processing systems, because motion vectors are critical information for temporal redundancy reduction. The method adopted to estimate the motion between frames is the block matching algorithm (BMA) [1-7]. 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) [2]. 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 [2-7] 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) [2], the block-based gradient descent search algorithm (BBGDS) [7], and many variations. These algorithms reduce the number of computations required by calculating the MAD matching criterion at locations coarsely spread over the search window according to some pattern and then repeating the procedure with finer resolution around the location with the minimum MAD found from the preceding step. Obviously, how to select the initial search pattern is the most crucial task. Nearly all existing fast search algorithms rely on the assumption: the MAD distortion function increases monotonically as the search location moves away from the global minimum [6]. Essentially, this assumption requires that the MAD error surface be unimodal over the search window, which leads to employing a uniform pattern. 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. In this paper, a reliable search strategy through the guide of some image features is proposed to locate one of the initial search points as close as possible to the true motion vector so that the chance of catching the true motion vector is maximized.

## 2. The Statistical Behaviour of 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 algorithm is used, each corresponding to a checking point in the search window. The MAD values resulting from these checking points form an error surface

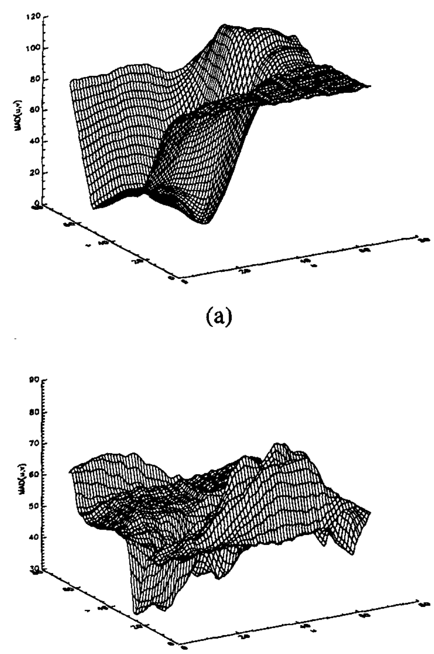


0-7803-5467-2/99/ $10.00 © 1999 IEEE

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

where the block size is taken as $N \times N$, $(u,v)$ denotes the location of the candidate motion vector, and $I_t(\cdot,\cdot)$ and $I_{t-1}(\cdot,\cdot)$ 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 a fast search algorithm for block 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 simple fast search algorithms such as the n-step hierarchical search [2] and the block-based gradient decent search [7] would require a small number of searches to determine the global optimum 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 [6] 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 into a local minimum.



(a)

(b)

**Figure 1. Error surface for two different blocks.**

## 3. Reliable Search through Image Features

The search algorithm presented in this paper can best be described as an extension of the Block-Based Gradient Descent Search (BBGDS) algorithm [7], as illustrated in Fig. 2. 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 (*3x3* checking points) or the checking block hits the boundary of the predefined search range [7]. Of course, the BBGDS algorithm relies on the assumption that the MAD measure decreases monotonically as the search point moves closer to the optimum point. It can easily be trapped into the local minimum when the error surface looks like Fig. 1(b). Let us use Fig. 3 to give a clearer account of this phenomenon. Fig. 3 shows a nonunimodal surface due to many reasons such as the aperture problem, 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 frame.

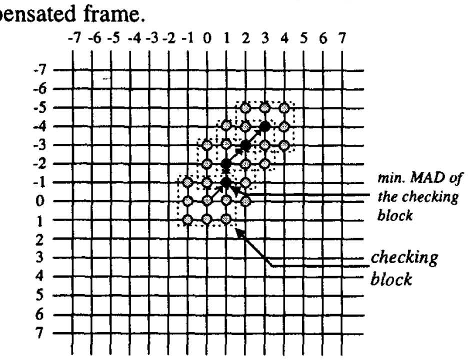



**Figure 2. Example of the BBGDS search procedure.**

Despite the error surface exhibiting uncertainties in large spatial scale, we can reasonably assume that it is monotonic in a small neighbourhood around the global minimum. In the existence of local minima, one simple, but perhaps the most efficient and reliable strategy, is to put the checking block as close as possible to the global minimum, as depicted in Fig. 4. If the initial checking block is close enough to the global minimum, it will be very likely to find the global minimum through a local search. One possible solution is to test more starting points spread across the search window. Fig. 5 shows one of the starting point pattern (SPP) in which the starting points (SP) are distributed evenly over the search window. However, it is inefficient for using such a large amount of SP in this regular SPP. Consider the search window with the MAD surface shown in Fig. 1(a), it is wasteful for using all the starting points as shown in Fig. 5. It is obvious that if the number of starting points is reduced as much as possible and the starting point is as close as possible to the true motion vector, the search algorithm becomes efficient. Hence, we have to adjust the regular SPP among the blocks so that limited SPs have a large chance of catching the global minimum. In this paper, we propose a feature-assisted search algorithm. The adjustment of the regular SPP is a primitive-based approach which generally includes a matching process for tracking feature primitives from frame to frame in a sequence of images. The proposed algorithm first estimates an initial probability of being the global minimum of each possible matching pair between the current block and the block at the regular SPP. Then the regular SPP is updated based on some criteria, such as the feature similarity. Many features, such as edge points, corner points and segmented regions, can be used as matching primitives. This approach has good results if the extraction of primitives is nearly perfect.

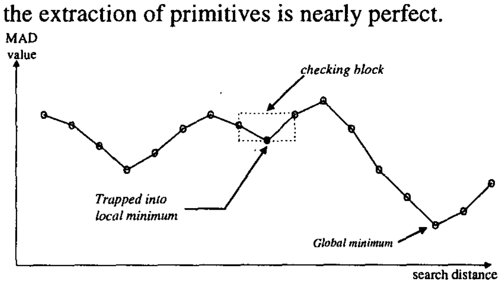



**Figure 3. A nonunimodal error surface sampled by checking block.**

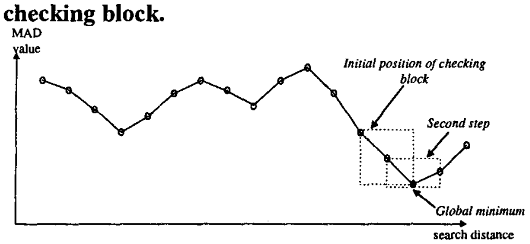



**Figure 4. If the initial checking block is close enough to the global minimum, the global minimum can be successfully found.**

## 4. The Edge-Assisted Search(EAS) Algorithm

In [8], we have shown that edge features can give great improvement on the accuracy of motion estimation. Also, the availability of VLSI edge detection chips [9] makes the possibility of using edges in motion estimation quite realistic and potentially rewarding. In this paper, we try to employ the binary edge feature as an example to illustrate our proposed feature-assisted searching approach and call it as Edge-Assisted Search (EAS) algorithm. In the following, we highlight the main steps of our EAS algorithm.

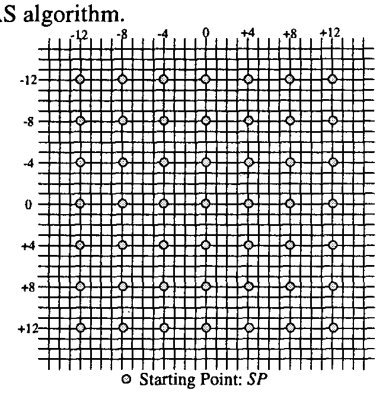



**Figure 5. Regular SPP**

**Step 1: Image pre-processing by edge detection**

An efficient extraction of the binary edge information, $B_t(i,j)$, is an important process for our proposed EAS algorithm. The binary edge detector used in this paper is based on the smoothing process and the $3 \times 3$ Sobel gradient convolution masks as described in [8]. A block is considered to contain an edge if $\sum_{i=0}^{N-1}\sum_{j=0}^{N-1} B_t(i,j) \geq T_{count}$ is satisfied, where $T_{count}$ is the threshold of edge count. This is done in order to prevent false determination of edge blocks due to noise, consequently guaranteeing that motion vectors obtained in the EAS algorithm are more reliable. If the block is classified as an edge block, the search of this block proceeds to step 2; otherwise the conventional BBGDS is applied.

- ***Step 2: Adjustment of the regular SPP***

Once the image feature description is determined, a match evaluation function is required to show the degree of similarity between two descriptions. Usually the similarity of two descriptions is defined in the form of a cost function or a distance function, where these costs are expected to be minimized and are zero only if both descriptions are identical. In this paper, we propose to employ the Edge Matching Score, EMS, as the cost function to adjust the regular SPP. The EMS is defined as the difference between the sum of edge points of each block

$$EMS(u,v) = \left| \sum_{i=0}^{N-1}\sum_{j=0}^{N-1} B_t(i,j) - \sum_{i=0}^{N-1}\sum_{j=0}^{N-1} B_{t-1}(i+u, j+v) \right| \quad \textbf{(2)}$$

where $(u,v)$ denotes the location of the possible motion vector, and $B_t(\cdot,\cdot)$ and $B_{t-1}(\cdot,\cdot)$ refer to the binary edge blocks of the current frame($t^{th}$ frame) and the reference frame $((t-1)^{th}$ frame) that are to be compared.

The adjustment of the regular SPP is based on the measure of how large the probability of being the global minimum of each possible matching pair between the current block and the block at the regular SPP is. The EMS is considered if the numbers of edge points in two blocks are similar. In this case, the block in the regular SPP has a large probability of being closest to the global minimum. In Fig. 6, the MAD surfaces and the EMS surfaces of a block containing a racket moving against a background is plotted. We have found that the correlation between these two surfaces is very high and it further ensures that the motion search algorithm can be guided by the EMS. Thus a block in the regular SPP whose EMS is less than a pre-defined threshold, $T_{EMS}$, will be considered good as an interesting SP. In other words, this SP is reserved in the updated SPP. In order to normalize thresholding, the $T_{EMS}$ is proportional to the number of edge points of the current block. That is,

$$T_{EMS} = \alpha \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} B_t(i,j) \qquad (3)$$

where $\alpha$ is a proportional constant.

- ***Step 3: The formation of the final SPP***

In order to reduce further the computational complexity, the updated SPP can be refined by using the image intensity. A simply way is to employ the MAD matching criterion. The approach is based upon the MAD values for selecting the best matched SP as compared to other SPs in the updated SPP, and it is defined as

$$G_k = MAD_k - MAD_{\min\_in\_updated\_SPP} \qquad (4)$$

for $k$ means to cover all selected SP of the updated SPP, except the SP with the smallest value of MAD in the updated SPP, where $MAD_{min\_in\_updated\_SPP}$ and $MAD_k$ are the smallest value in the updated SPP and the value of the MAD from the SP in the updated SPP, respectively. First, the SP with smallest value must be reserved as the final SPP. Second, the value of $G_k$ is used to establish the final SPP. If the value of $G_k$ is small enough (smaller than $\beta \times MAD_{min\_in\_updated\_SPP}$, where $\beta$ is also a proportional constant), it implies that the probability of being the global minimum of this SP is high. In other words, this SP must be included in the final SPP; otherwise, the SP is eliminated from the updated SPP. After examining all the SP in the updated SPP, the final SPP is formed.

- ***Step 4: Motion vector estimation using the BBGDS***

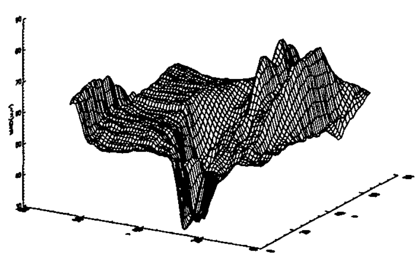
After the establishment of the final SPP, all SPs in the final SPP are served as the starting point of the BBGDS [7]. And finally, searches are conducted to find a minimum value of MAD.

## 5. Simulation Results

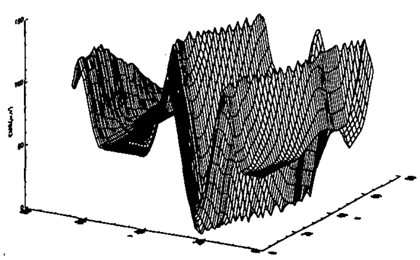
A series of computer simulations have been conducted to evaluate the performance of the proposed EAS algorithm. These include the "Table Tennis" and the "Football" sequences in SIF format. The maximum allowable displacement in both horizontal and vertical directions is *15* with a block size of *16 × 16*. The mean square error (MSE) is used to compare the performance of the proposed algorithm with the related techniques in the literature. For our proposed EAS algorithm, the parameters in the formation of the final SPP $\alpha$ and $\beta$ are set to *0.15* and *0.3*, respectively; while the $T_{count}$ of the edge detection is set to *16*. These values are selected by considering the trade-off between the computational requirement and the quality of most image sequences. The conventional methods for performance comparison are the FSA, the n-SHS [2], and the BBGDS [7].

Current block

reference search window



MAD error surface of the "ball"



EMS surface of the "ball"

**Figure 6. The relationship between the MAD error surface and the proposed EMS surface.**

Fig. 7 shows the results of the MSE of the motion-compensated frames of the proposed algorithm together with some traditional approaches for comparison. In Fig. 7(a) and (b), there is a great increase in the prediction error of the n-SHS and the conventional BBGDS as compared with that of the FSA. It is because the probability of the situation like Fig. 3 occurring is more often in sequences with fast moving objects. This situation makes an inappropriate choice in early steps of the n-SHS, and the unreliable stop in searching of the conventional BBGDS, it implies that these kinds of algorithms are more easily trapped in a local minimum. However, our new EAS algorithm can resolve the problem of the local minimum by placing the checking block closest to the global minimum which is guided by the edge features. As shown in Fig.6, the new EAS 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. Furthermore, by combining all of the operations of our EAS, it has a speed up of over 20 times as fast as the FSA.

## 6. Conclusions

In this paper, a fast search algorithm for block motion estimation has been proposed. The proposed algorithm generally includes a matching process for tracking edge primitives from frame to frame in a sequence of images, hence we can consider it as an edge-assisted search algorithm, EAS. Edge features have been used for the adjustment of start point patterns (SPP) of the search windows such that a limited number of starting points can still give a large chance of catching the global minimum. This method firstly estimates an initial probability of being the global minimum of each possible matching pair between the current block and the block at the SPP. Then the SPP is updated based on the Edge Matching Score, EMS, which is introduced to consider the degree of similarity of edge points between two blocks. We have demonstrated that the correlation between the EMS and the true motion vector is very high and it can ensure that the motion search algorithm can be guided by the EMS. We have tested the proposed EAS using a number of image sequences, including the "Table Tennis" and the "Football" and found that it can reduce the heavy computational burden of the full search algorithm without significantly increasing the prediction error of the motion-compensated frame. The EAS is significantly better than that of the famous search algorithms such as the n-SHS [2] and the BBGDS [7], and shows great improvement in the accuracy of the block motion estimation.

## 7. Acknowledgments

This work was supported by the POLYU 351/844

### References


[1] M. Bierling, "Displacement Estimation by Hierarchical Block Matching," in SPIE, Visual Comm. Image Processing, vol. 1001, pp. 942-951, 1988.

[2] T. Koga, K. Iinuma, A. Hirano, Y. Iijima, and T. Ishiguro, "Motion Compensated Interframe Coding for Video Conferencing," in Proc. Nat. Telecommun. Conf., pp. G.5.3.1-5.3.5, Nov.29-Dec.3, 1981.

[3] M.J. Chen, L.G. Chen and T.D. Chiueh, "One-Dimensional Full Search Motion Estimation Algorithm for video coding," IEEE Trans. on Circuits and Systems for Video Tech., vol. 4, no. 5, pp. 378-385, October, 1994.

[4] Y.L. Chan and W.C. Siu, "Adaptive Multiple-Candidate Hierarchical Search for Block Matching Algorithm," Electronics Letters, vol. 31, no. 19, pp.1637-39, Sept. 1995.

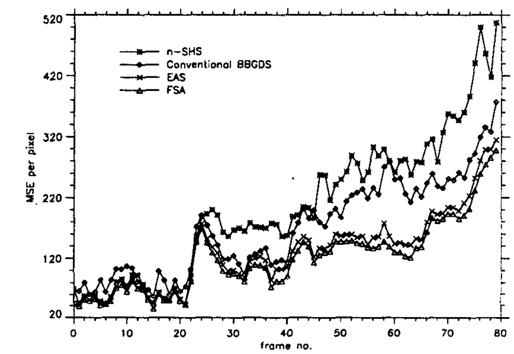
[5] Y.L. Chan and W.C. Siu, "On Block Motion Estimation using a Novel Search Strategy for an Improved Adaptive Pixel Decimation," Journal of Visual Comm. and Image Representation, vol. 9, no. 2, pp. 139-154, June 1998.

[6] 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.

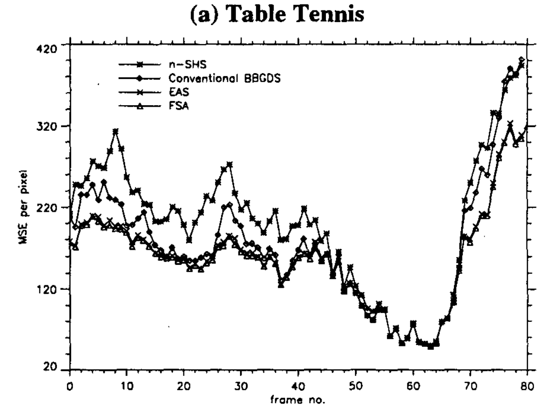
[7] L.K. Liu and E. Feig, "A Block-Based Gradient Descent Search Algorithm for Block Motion Estimation in Video Coding," IEEE Trans. on Circuits and Systems for Video Tech., vol.6, no.4, pp. 419-421, August 1996.

[8] Y.L. Chan and W.C. Siu, "Edge Oriented Block Motion Estimation for Video Coding," IEE Proceedings: Vision, Image and Signal Processing, vol. 144, no. 3, pp. 136-144, June 1997.

[9] C. Lee, F.V.M. Catthoor and H.J.De Man, "An Efficient ASIC Architecture for a Time Edge Detection," IEEE Trans. on Circuits and Systems, vol. 36, no. 10, pp. 1350-60, 1989.


**(a) Table Tennis**




**(b) Football**

**Figure 7. The MSE produce by different algorithms, the "Table Tennis" and the "Football."**