A Hierarchical Motion Estimation Algorithm Using Correlation of Motion Fields

B.C. Song, K.W. Lim, and J.B. Ra
(Department of Electrical Engineering, KAIST)

Abstract: *A new three step hierarchical search algorithm for motion estimation is proposed. The proposed algorithm exploits the motion correlation of spatially neighboring blocks and the motion continuity of temporally neighboring blocks to alleviate the local minimum problem in the first step of the three step hierarchical search algorithm (3SHS). Simulation results show that the proposed scheme achieves significant improvements in both estimation accuracy and performance reliability compared with the existing fast block matching algorithms including 3SHS, while maintaining almost the same computational complexity as 3SHS. The proposed scheme also possesses the regularity and simplicity of hardware-oriented features.*

1. INTRODUCTION

In interframe predictive digital video coding system, the temporal redundancy due to high correlation between consecutive frames can be eliminated by motion estimation/compensation techniques. Most of current video communication systems and standards such as CCITT H.261 and MPEG employ a block matching scheme for motion estimation. A straight-forward block matching algorithm (BMA) is a full search BMA (FS-BMA) which matches every possible displaced candidate block within the search area in the reference frame in order to find the block with minimum distortion. However, the implementation of FS-BMA always needs massive computational efforts and huge hardware cost to achieve significant speed. As a result, many fast BMA's, such as the three-step hierarchical search (3SHS) [1], the 2-D logarithmic search (LOGS) [2], the 1-D full search BMA (1DFS) [3], had been developed to alleviate the heavy computation of FS-BMA. Basically, these algorithms reduce the number of searching operations by selectively checking only a small number of positions under the assumption that distortion increases monotonically as the searching location moves away from the best matching position. However, this assumption does not always hold in practice and thus a motion vector often gets trapped in a local minimum. Therefore, none of the fast search algorithms mentioned above will always give reliable performance. In order to alleviate this local minimum problem, variants of 3SHS, such as the center-biased 4 step search (C4SS) and the predictive logarithmic search (PLOGS) [4-5], have been proposed recently. However, C4SS is not suitable for hardware implementation because its number of searching operations per block varies from 17 to 27. PLOGS has a drawback that the probability of being trapped in a local minimum in the first stage is still high.

This paper proposes a new fast search algorithm that solves the local minimum problem in the first step of 3SHS by using the temporal correlation of motion fields as well as the spatial correlation in order to find an approximate motion direction. Unlike PLOGS, our scheme uses a previous motion vector as the approximate motion direction of a current block instead of using the location with minimum mean absolute difference (MAD) among nine locations in the first step of 3SHS. In case of uniform motion, a previous motion vector becomes the motion direction of the current block, while using nine locations in the first step of 3SHS is not suitable for predicting approximate motion direction because of the high possibility of being trapped in a local minimum. Then the center of search for the second step is chosen by comparing the predicted motion direction with four spatially neighboring motion vectors, and search is continued as in 3SHS. In Section II, the proposed algorithm is described. In Section III, simulation results of the proposed scheme are reported and compared with other methods. In Section IV, conclusions are presented.

II. PROPOSED ALGORITHM

The proposed algorithm uses the spatio-temporal correlation of motion fields in order to reduce the risk of being trapped in a local minimum in the first step of 3SHS. In most video sequences, motion is smooth and slow-varying. Motion discontinuity mainly occurs at the boundary of objects moving in different directions. Because moving objects usually cover several blocks, motion vectors of adjacent blocks are highly correlated. Therefore, the motion vector of a current block is predictable from the motion vectors of its neighboring blocks if interblock correlation is properly used. Using this property, the predictive logarithmic motion estimation algorithm (PLOGS) had been proposed. In the initial stage, the motion direction of each block in the current frame is extracted by examining coarsely sampled nine locations. Then, its correlation with motion vectors of the adjacent blocks is checked, and the motion vector of the best correlated block is conditionally chosen as the initial search center of the second stage in a logarithmic search. The search continues until the search plane reduces to a 1 x 1 size. But this scheme has a drawback. Since its first step is based on logarithmic search, which has high possibility of being trapped in a local minimum, it is hard to expect a good motion vector candidate in the video sequences having slow-moving objects. Therefore, PLOGS is not enough to reduce the risk of being trapped in a local minimum even if the spatial correlation with motion vectors of its neighboring blocks is used in addition to the logarithmic search to predict motion direction. In order to predict the motion direction before using spatial correlation, the motion vector of the corresponding block in the previous frame may be more useful than the motion direction predicted from nine searching points in the first step of logarithmic search (as in PLOGS).

Our search algorithm is a two step process. In the initial stage, the approximate motion direction of each block in the current frame is extracted. If moving objects have uniform motion, the motion vector of the previous block will be the same as the motion direction of the current block. Fig. 1 demonstrates how motion continuity is used. Correlation of the predicted motion direction with the motion vectors of adjacent blocks is calculated. Then the motion vector of the best correlated block is chosen as the initial search center of the next stage. In the second stage, the three step hierarchical search continues with the chosen center of search. The proposed method is depicted as follows.

II.1. Finding the center of search using spatio-temporal correlation

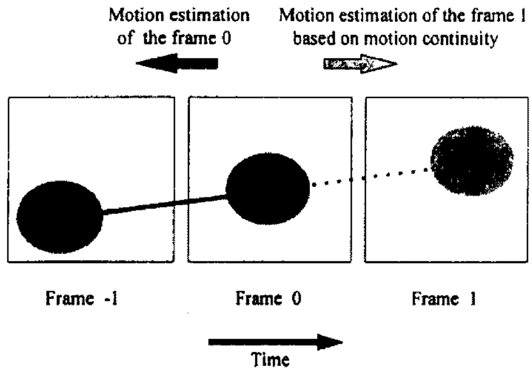


Fig. 1. Use of motion continuity in predicting motion direction.

As shown in Fig 1, the motion vector of a previous block, (m_{prev}, n_{prev}), will be almost the same as the motion vector of a current block, (m_{cur}, n_{cur}), when moving objects move uniformly. Therefore, we choose (m_{prev}, n_{prev}) as the approximate motion direction of the current block. Next, as a spatial correlation extraction process, we compare (m_{prev}, n_{prev}) with motion vectors of the four available adjacent blocks based on the distance measure (D) defined as

$$D = (m - m_{adj})^2 + (n - n_{adj})^2,$$

where (m, n) is the approximate direction of the current block, and (m_{adj}, n_{adj}) is the direction of an adjacent block. When abrupt motion occurs, (0, 0) can be a good candidate motion vector of the current block. Therefore (m, n) can be (m_{prev}, n_{prev}) or (0, 0). A certain threshold, THR, for correlation extraction is to be decided beforehand. If D is larger than or equal to THR, interblock correlation will not be used in the next search step, and (0, 0) will be the center of the next step. Otherwise, the block with the smallest D from the four adjacent blocks will be selected as the best correlated block.

II.2. Three-step hierarchical search

With the above chosen motion vector as the center of search, the three-step hierarchical search goes on for the maximum motion displacement w. For each N x N block centered at (i, j), its corresponding search range R is defined as

$$R=\{(x,y)| -w \le x, y < w\},$$

and a set of searching points is denoted by

$$M(a,b,n)=\{(a,b),(a,b\pm n),(a\pm n,b),(a\pm n,b\pm n)\}.$$

Step 1: Use the motion vector (a, b) of best correlated block as the search center. Find a position (u_1, v_1) with the smallest MAD among $M(a, b, n)$ (n=5 is used here). If any of nine searching points falls outside R, it is offset by $+3n$ or $-3n$ in either or both x, y coordinates so that the new searching point lies within R. (x_{old}, y_{old}) is renewed to (x_{ren}, y_{ren}) as follows.

If $x_{old} < -w$,

$$x_{ren} = x_{old} + 3n,$$

else if $x_{old} \ge w$,

$$x_{ren} = x_{old} - 3n,$$

else

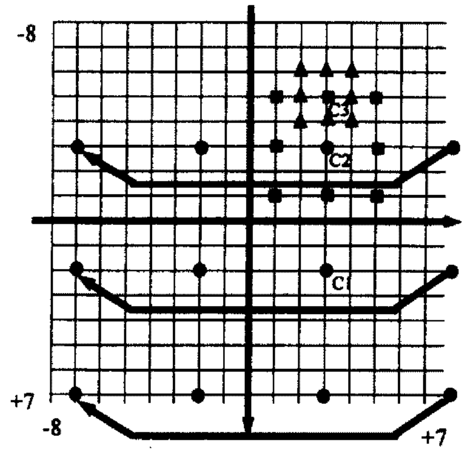
$$x_{ren} = x_{old}.$$

y_{old} is renewed in the same manner as x_{old}.

Step 2: Find the position (u_2, v_2) with minimum MAD among $M(u_1, v_1, 2)$. Use (u_2, v_2) as the search center of the next step.

Step 3: Find the position (u_3, v_3) with minimum MAD among $M(u_2, v_2, 1)$. (u_3, v_3) is chosen as the current block's motion vector.

For a maximum motion displacement of w=8, Fig. 2 shows a graphical representation of the proposed algorithm. *Cn* is the search center of step *n*. How searching points outside R are relocated into R is represented in Fig. 2. As illustrated in Fig. 2, the proposed algorithm is very similar to that of 3SHS except our search center is determined by the spatio-temporal correlation instead of being fixed at (0, 0) in the first step. Computational requirement for determining the search center is negligible compared with that of one MAD calculation, because only 8 distances are calculated beforehand. Therefore, computational complexity of the proposed algorithm is almost the same as that of 3SHS. Table I compares the number of searching points per block, in the worst case, for various fast algorithms with maximum motion displacement of w=8.



●: Searching points in step 1 ■: Searching points in step 2
▲: Searching points in step 3 ●: Searching points outside *R*

Fig. 2. An example of the proposed algorithm for w=8. n=5 is used here.

	FS-BMA	3SHS	PLOGS	Proposed
Number of Searching Points	256	25	34	25

Table I. Comparison of the number of searching points per block in the worst case, for various fast search algorithms with maximum motion displacement of w=8.

III. SIMULATION RESULTS

The performance of the proposed algorithm is measured in terms of PSNR. We used *Football, Susie,* and *Mobile* as test sequences which have an image size of 720 x 480 pixels quantized uniformly 8 bits. MAD is used as a distortion measure. We adopted the maximum motion displacement w=8 with a block size of 16 x 16. In the simulation, previous motion vectors are initially set to (0, 0). THR is set to w^2 or 64 according to our simulation results.

Fig. 3 compares performance of 3SHS, FS-BMA, PLOGS, and the proposed algorithm in terms of PSNR for the *Mobile* sequence. The proposed scheme is more robust against being trapped in a local minimum than

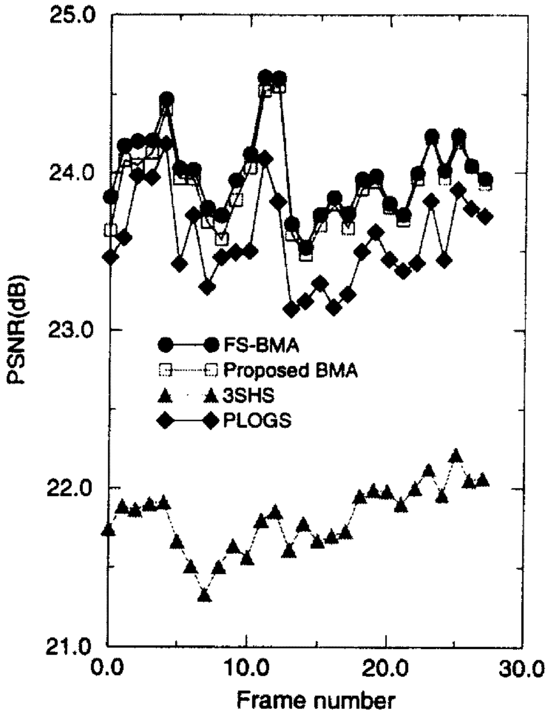


Fig. 3. The PSNR comparison for the *Mobile* sequence.

PLOGS as well as 3SHS, because the center of search is chosen by using motion continuity and spatial correlation of motion fields. It is observed that the performance of the proposed algorithm closely follows that of FS-BMA, and our scheme outperforms recently developed fast search algorithms. Table II gives PSNR comparison of various fast algorithms for several test sequences. The superior performance reliability of our scheme demonstrates that the local minimum problem has been drastically reduced by exploiting spatio-temporal correlation of motion fields effectively.

	Football	*Mobile*	*Susie*
FS-BMA	22.47(dB)	24.00(dB)	37.23(dB)
3SHS	21.87	21.82	36.52
PLOGS	22.00	23.57	36.94
Proposed	22.07	23.93	37.04

Table II. PSNR comparison of proposed algorithm with various fast algorithms with w=8.

These experimental results clearly show that the proposed algorithm always provides better performance with the minimal computational requirement among other fast search algorithms.

IV. CONCLUSIONS

In this paper, a fast predictive 3-step hierarchical search algorithm using spatio-temporal correlation of motion fields is proposed. Motion correlations between adjacent blocks as well as temporal motion continuity are found to be very useful information in locating the initial search center of 3SHS so as to reduce the risk of being trapped in a local minimum. The simulation results show that the proposed scheme can keep the performance up close to FS-BMA, with minor PSNR degradation of 0.07dB in the *Mobile* sequence, while maintaining minimum computational requirement among fast search algorithms.

REFERENCES

[1] T. Koga *et al.*, "Motion compensated interframe coding for video-conferencing," *Proc. Nat. Telecommun. Conf.*, New Orleans, LA, pp. G5.3.1-G5.3.5, 1981.

[2] J.R. Jain and A.K. Jain, "Displacement measurement and its application in interframe image coding," *IEEE Trans. Commun.*, vol. COM-29, no. 12, pp. 1799-1808, Dec. 1981.

[3] M.J, Chen, L.G. Chen, and T.D. Chiueh, "One-dimensional full search motion estimation algorithm for video coding," *IEEE Trans. Circuits and Syst. for Video Technology*, vol. 4, no. 5, pp. 504-509, Oct. 1994.

[4] L.M. Po and W.C. Ma, "A new center-biased search algorithm for block motion estimation," *Proc. ICIP-95*, pp. 410-413, Oct. 1995.

[5] M.C. Chen and A.N. Willson, Jr., "A high accuracy predictive logarithmic motion estimation algorithm for video coding," *ISCAS-95*, pp. 617-620, Apr. 1995.