

Hybrid Block-based Motion Estimation Considering Regional Motion Characteristics

Young Duk Kim¹, Joonyoung Chang² and Moon Gi Kang³

Institute of TMS Information Technology, Yonsei University

134, Shinchon-Dong, Seodaemun-gu, Seoul 120-749, South Korea

E-mail : ¹idea9973@hotmail.com, ²gubung1981@yonsei.ac.kr, ³mkang@yonsei.ac.kr

Abstract: In this paper, we propose a hybrid block-based motion estimation algorithm that considers regional motion characteristics. The proposed method combines multiple search algorithms according to regional motion types to produce more accurate motion vector fields while preventing the increase of the computational cost. For these reasons, the proposed method first divides an entire image into three regions with different motion types. Experimental results show that the proposed motion estimation method improves the accuracy of the estimated motion vector fields over the existing algorithms.

1. Introduction

With the rapid convergence of television, multimedia contents, and personal imaging systems, video format conversion such as de-interlacing or frame rate conversion (FRC) has become an inevitable process. Although a number of de-interlacing or FRC algorithms have been studied with different levels of quality and complexities, motion compensation based algorithms are the most advanced technique with high hardware complexity. They reconstruct missing lines or frames by interpolating pixels along motion trajectories. Motion compensation plays a key role to enhance vertical resolution or temporal resolution, and therefore it produces high quality images, provided that estimated motion information is reliable.

However, it is difficult to ensure the accuracy of motion estimation, and incorrect motion vectors may introduce visible artifacts such as feathering effect or block artifacts, which should be avoided since they are extremely visible to human eyes. Therefore, a number of motion estimation algorithms [1]-[8] have been studied to increase motion estimation accuracy.

Although various types of motion estimation algorithms such as pel-based schemes [3], object-based approaches [4], and hybrid methods with an object-based motion model [8] have been investigated, the block matching algorithm (BMA) [1][5]-[7] has become widely accepted in real-time video applications due to their efficient hardware implementation.

To increase the efficiency of the BMA algorithm further, various search methods including full search, three step search (TSS) [6], hierarchical search [1]-[2][8], and recursive search [3] have been studied. Among them, the recursive search methods provide stable and accurate motion vector fields and require less computational cost. Recursive search algorithms limit the number of candidate vectors that consist of spatially or temporally neighboring results so that they can exploit the spatio-temporal correlation of motion vectors.

However, transient time is necessary to keep track of objects with quite different motions from neighboring blocks due to the limited range of candidates based on the predictions. During the transient time, the recursive search method may produce inaccurate motion vectors.

In this paper, we present a hybrid block-based motion estimation algorithm that considers regional motion characteristics. The proposed method reduces the transition time, which is the main drawback of recursive search, by effectively combining multiple search algorithms including recursive search, three step search with predictions, and single predicted search. The advantages of the proposed method come from consideration of complementary features of multiple search schemes and regional motion characteristics.

The organization of the paper is as follows. In Sec. 2, the proposed motion estimation algorithm is described in detail. In Sec. 3, experimental results of various images are presented and comparisons with other algorithms are given. Finally, conclusions are made in Sec. 4.

2. Proposed Hybrid Motion Estimation Considering Motion Characteristics

The hybrid motion estimation method effectively combines multiple search schemes according to regional motion characteristics to improve accuracy and reduce the computational load. In this paper, three step recursive search (TSRS), three step search with predictions (TSSP), and single predicted search (SPS) are considered as candidates for search process.

The TSRS process considers only limited candidates based on spatial and temporal predictions with additional update vectors. The TSRS method operates in the same way as the conventional three step search (TSS) scheme except for the first step. Figure 1 illustrates the operations of the conventional TSS algorithm. As shown, the TSS method performs search process coarsely in the first step and find a displacement with minimum error. Then, refinements are performed around the displacement in the second and third steps. However, the TSRS method searches only the predicted points in the first step, where the search set is composed of the spatially and temporally predicted motion vectors. The restrictions on the candidate vectors usually improve the accuracy and the smoothness of the motion vectors as well as decrease the computational load. However, it inherently requires transient time to keep track of objects with quite different motions from neighboring blocks due to the limited range of candidates.

To reduce transient time, a search scheme capable of responding immediately to abrupt motion discontinuities (such as the TSSP method) is essentially needed for assisting the

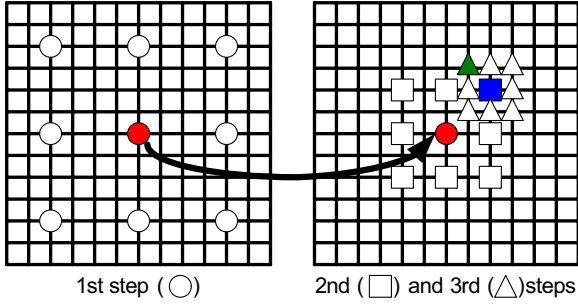


Figure 1. The operations of the conventional TSS algorithm.

TSRS process. The TSSP method operates in the same way as the conventional three step search (TSS) scheme except for the first step. In the first step, it additionally searches the predicted search set besides the original search points used in the conventional TSS method. With the help of search into the predicted search set, the TSSP method improves the accuracy since it reduces the hazards to be trapped into the local minima. As a result, the transient time can be skipped if the TSSP method produces accurate motion vectors for the object boundaries. However, as a large search range is allowed for tracking fast motions, the gap between the estimated and true motion vector fields tends to increase. Hence, to improve accuracy, the TSSP and TSRS methods use motion smoothness constraint besides the SAD for determining the motion vector, such as:

$$\mathbf{mv} = \arg \min_{\mathbf{k}} \epsilon(\mathbf{k}), \quad (1)$$

where \mathbf{k} represents the displacement vector, and $\epsilon(\mathbf{k})$ is the matching cost function determined by:

$$\epsilon(\mathbf{k}) = \text{SAD}(\mathbf{k}) \cdot (1 + \text{MSC}(\mathbf{k})), \quad (2)$$

where $\text{SAD}(\mathbf{k})$ and $\text{MSC}(\mathbf{k})$ represent the SAD and motion smoothness error function, respectively. The SAD is calculated bi-directionally and symmetrically with respect to the current block B as shown in Fig. 2, since it enhances accuracy without causing the hole and overlapping problems [2], such as:

$$\text{SAD}(\mathbf{k}) = \sum_{\mathbf{x} \in B} \left| f(\mathbf{x} + \frac{\mathbf{k}}{2}, n-1) - f(\mathbf{x} - \frac{\mathbf{k}}{2}, n) \right|, \quad (3)$$

where \mathbf{x} and n represent the spatial and temporal indices, respectively, and f represents the frame data.

The motion smoothness constraint (MSC) measures the similarities of a displacement to the neighboring motion vectors that are already estimated, such as:

$$\text{MSC}(\mathbf{k}) = \frac{1}{\alpha} \cdot \sum \left[\frac{|\mathbf{k} - \mathbf{mv}_{\text{neighbor}}|}{Q} \right], \quad (4)$$

where $\mathbf{mv}_{\text{neighbor}}$ represents the motion vectors of neighboring blocks, Q and α represent a quantization factor and a constant value controlling the relative magnitude, respectively, and $[\]$ represents an operator to return the maximum integer less than or equal to the given real number. They are set to the power of two to simplify computations.

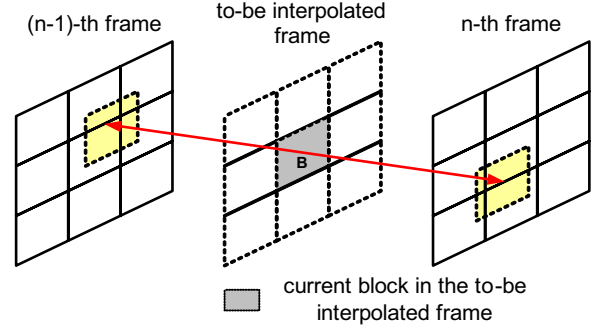


Figure 2. Bi-directional motion estimation method where motion vector is found symmetrically with respect to the current block B in the to-be interpolated frame.

Besides the TSSP and TSRS methods, the SPS method is also employed as an auxiliary mean to reduce computational cost. The SPS method is a low cost approach that allows only one candidate of the temporally predicted motion vector. Therefore, under the condition that the motion of a block is not likely to change between two successive frames, it becomes a cost-effective motion estimation scheme. However, temporal prediction approaches may be unsuccessful when there are scene changes, non-linear motions, or fast movements. For these reasons, we check the reliability of the predicted motion vector to detect any abnormal cases. If the predicted motion vector is determined to be unreliable, then the search method is altered to the TSSP method. The reliability of the predicted motion vector \mathbf{pmv} is measured by the SAD, such as:

$$\text{rel}(\mathbf{pmv}) = \frac{1}{\text{SAD}(\mathbf{pmv})}. \quad (5)$$

Corresponding to the above three candidates for search strategy, all blocks are classified into three types according to the motion characteristics, i.e., global, local, and boundary. The global motion region occupies the largest area of the same motion in a scene and usually contains motion of constant velocity, which is appropriate for the SPS method.

The boundary motion region is defined as the boundary between the global and non-global motion regions. Thus, motion vector discontinuities usually arise in that region so that the TSSP method requiring no transient time is suitable to this region. The local motion region embraces all the remaining blocks. In this region, the TSRS method derives accurate motion vector fields and does not suffer from transient time by potentially propagating the motion vectors determined by the TSSP method.

Consequently, the proposed hybrid motion estimation method can be summarized as,

$$\text{ME} = \begin{cases} \text{TSRS}, & \text{for local,} \\ \text{SPS}, & \text{for global, } \text{rel}(\mathbf{pmv}) > \text{TH}_r, \\ \text{TSSP}, & \text{otherwise,} \end{cases} \quad (6)$$

where $\text{rel}(\mathbf{pmv})$ represents the reliability of the temporally predicted motion vector \mathbf{pmv} , and TH_r represents the adaptive threshold value that is proportional to the high frequency energy of each block.

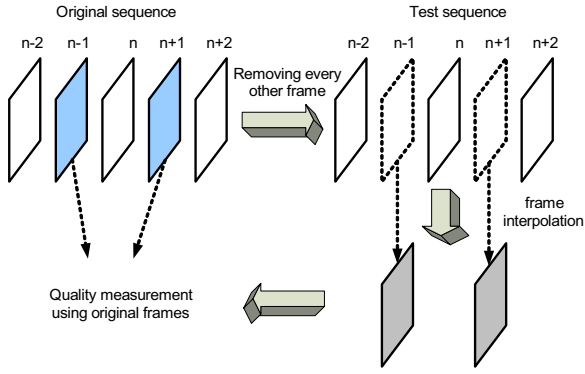


Figure 3. Test methodology for comparing various motion estimation algorithms including the proposed method. All test sequences were first sub-sampled by a factor of two in the temporal direction, and then the missing frames were reconstructed by using various motion estimation methods.

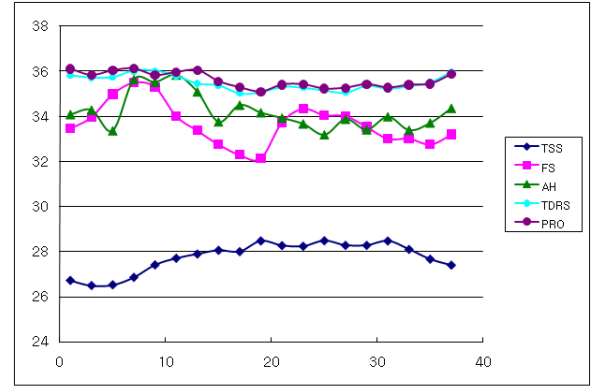
3. Experimental Results

The proposed algorithm was tested with various progressive sequences that contain various picture sizes and motion types. We also created test sequences by synthesizing normal video sequences with characters scrolling at various velocities to increase the complexity of motions and provide abundant occlusions and object boundaries. The test sequences were first sub-sampled by a factor of two in the temporal direction, and then the missing frames were reconstructed by using the tested motion estimation methods, as shown in Fig. 3.

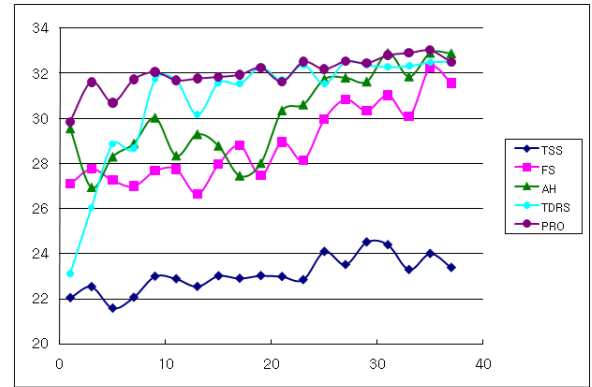
In Fig. 4, we compared the PSNR results of motion compensated images associated with various motion estimation schemes including three step search (TSS), full search (FS), adaptive hierarchical search (AH), three-D recursive search (TDRS), and the proposed method (PRO). The search range was set as 24×16 pixels, and the block size was set as 16×8 pixels, except for the AH [8]. In the AH method, the block size and the search range were set as 8×8 pixels and 3×2 pixels for all the 3-level pyramids, respectively. To ensure fairness, the bi-directional SAD described in Eq. (3) was used for all the tested algorithms.

As shown, the proposed method provided the best performance. The proposed method achieved outstanding performance improvements for the synthesized test sequence depicted in Fig. 4(b). Since "mobile_char" sequence contains fast moving characters, the TDRS method produced lower PSNR values for a few frames from the beginning of the test sequence. The performance improvements were conspicuous especially when the test sequence contains complex translational motions and fast-moving objects.

In Figs. 5 and 6, we present the magnified motion compensated results of various motion estimation methods to compare the subjective quality. The images shown in each figure (a) are the second pictures of the original mobile and mobilechar sequences. The other images (b)-(f) show the reconstructed images by using various motion estimation methods with the first and third pictures. The results of TDRS shown



(a)



(b)

Figure 4. PSNR comparisons of motion compensated results obtained by various motion estimation techniques. (a) "mobile sequence" (b) "mobile_char" sequence that is obtained by synthesizing the mobile sequence with scrolling characters

in the Fig. 6(e) seem worst among the tested algorithms, but it gets better as more pictures are processed. Namely, it provides erroneous results during the transient time since the white characters are scrolling fast, which have quite different motions from their neighboring blocks. However, the visual distortions are significantly reduced by the proposed method. As shown, the proposed method achieves the best visual quality among the tested algorithms.

4. Conclusions

We have proposed a hybrid block-based motion estimation algorithm that considers regional motion characteristics. The proposed algorithm constructed more accurate motion field than the separate search based methods by adequately combining three search strategies. It reduced the transient time, which is the main drawback of recursive search method, by incorporating a kind of three step search algorithm capable of responding immediately to the motion discontinuities. Also, it prevented increase of computational load by employing single predicted search method. The proposed algorithm was applied to various real video sequences with different types of motion for verifying the performance of the proposed algorithm. Simulation results show that the proposed motion esti-

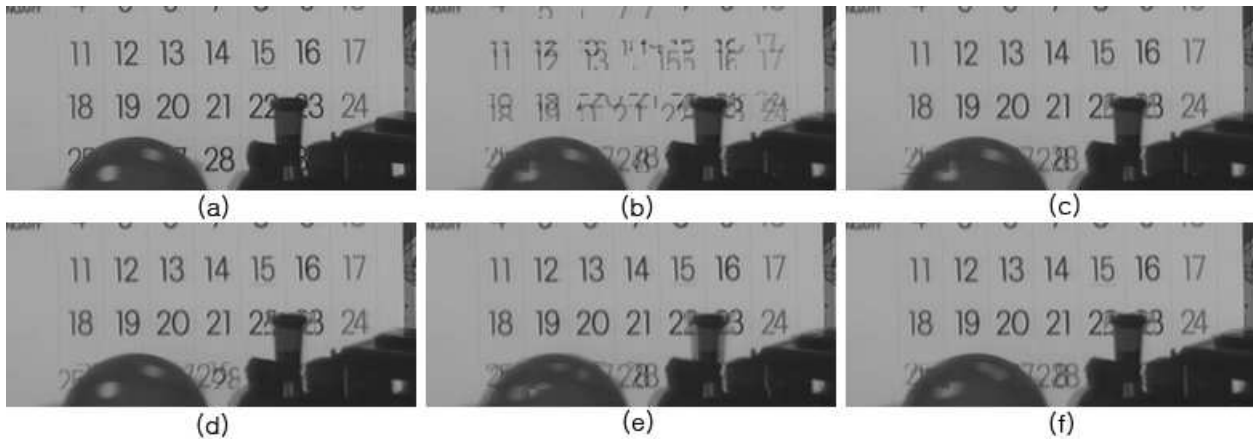


Figure 5. Magnified view of the motion compensated results obtained by various motion estimation techniques for the mobile sequence. (a) Original image (b) TSS method (c) FS method (d) AH method (e) TDRS method (f) Proposed method

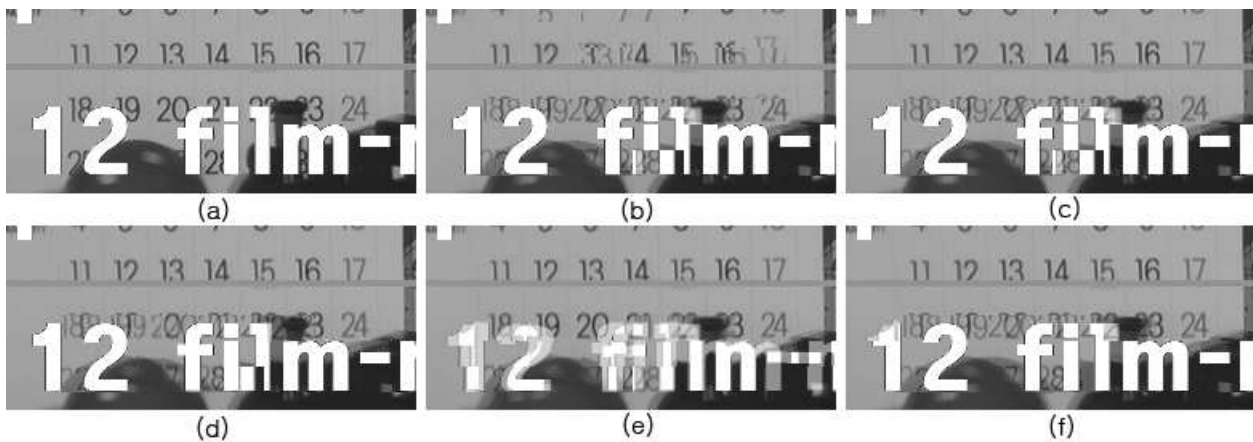


Figure 6. Magnified view of the motion compensated results obtained by various motion estimation techniques for the mobile character sequence. (a) Original image (b) TSS method (c) FS method (d) AH method (e) TDRS method (f) Proposed method

mation algorithm outperforms the conventional algorithms in both the objective and visual criteria.

Acknowledgement

This work was supported in part by Seoul Future Contents Convergence (SFCC) Cluster established by Seoul Industry-Academy-Research Cooperation Project at Yonsei University.

References

- [1] K. Bugwadia, E. Petajan, and N. Puri, "Progressive-scan rate up-conversion of 24/30Hz source materials for HDTV" *IEEE Transaction on Consumer Electronics.*, Vol. 42, No. 3, Aug. 1996.
- [2] B. Choi, S. Lee, S. Ko, "New frame rate up-conversion using bi-directional motion estimation" *IEEE Transaction on Consumer Electronics.*, Vol. 46, No. 3, Aug. 2000.
- [3] R. Castagno, P. Haavisto, and G. Ramponi, "A method for motion adaptive frame rate up-conversion" *IEEE Transaction on Circuit and Systems for Video Technology.*, Vol. 6, No. 5, Oct. 1996.
- [4] R. Schutten and G. Haan, "Real-time 2-3 pull-down elimination applying motion estimation/compensation in a programmable device" *IEEE Transaction on Consumer Electronics.*, Vol. 44, No. 3, Aug. 1998.
- [5] G. de Haan, P. W.A.C. Biezen, H. Huijgen, and O. A. Ojo, "True-Motion Estimation with 3-D Recursive Search Block Matching" *IEEE Transaction on Circuits and Systems for Video Technology.*, Vol. 3, No. 5, Oct. 1993,
- [6] A. M. Tekalp, *Digital Video Processing, Prentice Hall Signal Processing Series, Upper Saddle River, 1995.*
- [7] F. Dufaux and F. Moscheni, "Motion estimation techniques for digital TV: A review and a new contribution" *Proceedings of the IEEE.*, Vol. 83, No. 6, Jun. 1995.
- [8] D. Wang and D. Lauzon, "Hybrid algorithm for estimating true motion fields" *SPIE Optical Engineering.*, Vol. 39, No. 11, Nov. 2000.