A Fast Block Matching Motion Estimation Algorithm using Optimal Search Patterns

Dong-Keun Lim and Yo-Sung Ho

Kwangju Institute of Science and Technology
1 Oryong-Dong Puk-Gu, Kwangju, 500-712, Korea

ABSTRACT

For video compression, motion estimation is popularly employed to exploit temporal correlation existing in video sequences. If we employ the full search block matching algorithm for estimating motion vectors, it requires very heavy computational complexity. Although several fast block matching algorithms have been proposed to solve this problem, they sacrifice their reconstructed image quality. In this paper, we derive optimal search patterns for fast block matching motion estimation. By analyzing the block matching algorithm as a function of the block size and the shape, we find optimal search patterns for initial motion estimation. The proposed idea can provide an analytical ground for the current MPEG-2 proposals. In addition, we propose a new fast motion estimation algorithm using adaptive search patterns, considering matching criteria and statistical properties of object displacement. In order to select an appropriate search pattern, we exploit the relationship between the motion vector and the frame difference of each block. By changing the search pattern adaptively, we can improve the motion prediction accuracy, while reducing the required computational complexity compared to other fast block matching algorithms.

Keywords: Motion Estimation, Fast Block Matching Algorithm, Diamond Search, Advanced Zonal Search

1. INTRODUCTION

In recent days, there has been an increasing demand for real-time video communication services, such as wireless or internet video conferences. The concept that motion information of objects can enhance reconstructed image quality by temporal prediction has had a substantial influence on the progress of image sequence coding. Motion estimation has been widely used to find the motion information in various video coding standards and plays an important role in video compression.

The block matching algorithm (BMA) is one of the popular algorithms for motion estimation. In BMA, we estimate the amount of block displacement based on a simple rigid-body motion model. Because of its efficiency, BMA is widely adopted in most video coding standards, such as H.261, H.263, MPEG-1, and MPEG-2. A full search (FS) BMA method is employed in some video coding systems.

Since FS BMA makes an exhaustive search for an optimal block displacement which minimizes a predefined cost function, it generally provides good performance. However, it is computationally very expensive to find the optimal motion vector in a large search region. Therefore, various fast BMAs, such as 2-D logarithmic search (LOG), three-step search (TSS), conjugate direction search (CDS), dynamic search window adjust and interlaced search (DSWA), and four-step search (4SS), have been developed to reduce the computational burden associated with FS BMA.[1-7] Since most fast BMAs generally take heuristic approaches to reduce the computational complexity, they sacrifice reconstructed image quality.

In this paper, we focus on optimal search patterns for fast block matching motion estimation. Based on correlations of image blocks, we derive the grid shape for fast motion estimation analytically. The derived patterns are verified with various test sequences. The analysis presented in this paper can support several ideas proposed for the MPEG activities. [8-10]

In order to reduce computational complexity of the block matching operation without sacrificing picture quality, we modify the matching criterion for BMA. We also change the search pattern for motion estimation adaptively based on statistical

Correspondence: E-mail: {dklim, hoyo}@kjist.ac.kr; Telephone: +82-62-970-2247; Fax: +82-62-970-2204

properties between the object displacement and the frame difference of each block. The proposed adaptive motion search algorithm can reduce computational burden substantially relative to FS BMA, while producing good motion prediction.

2. OPIMAL SEARCH PATTERNS

2.1. Problem Statement

Since the shape and the size of the initial search pattern in the fast BMA jointly determine the convergence speed and estimation performance, we consider various search patterns: rectangle or cross of different sizes. Fig. 1(a) and Fig. 1(b) show search patterns for the three-step search (TSS) and the 2-D logarithmic search (TDL), respectively.

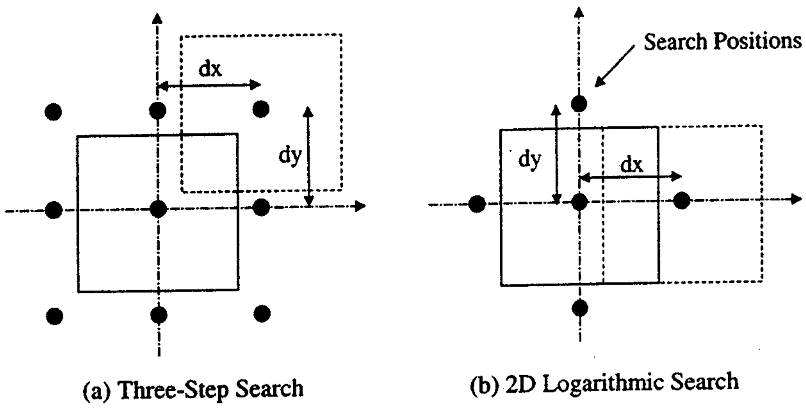


Figure 1. Search Patterns for TSS and TDL

In the first step of TSS, block correlations at nine search positions are examined, as shown in Fig. 1(a). In the second step, eight new search positions are spaced less coarsely around the best matching position in the first step. In the third step, the distances between search positions are further reduced, and the minimum distortion position is selected. In Fig. 1, *dx* and *dy* represent distances between adjacent search positions.

TDL tracks the direction of minimum distortion. In each step, five search positions which are located in a cross-shaped search pattern are examined, as shown in Fig. 1(b). If the minimum distortion position is observed in the center of the search positions or at the boundary of the search region, the distances between search positions are reduced. The final displacement vector is the minimum distortion position among all the search positions of one pixel distance.

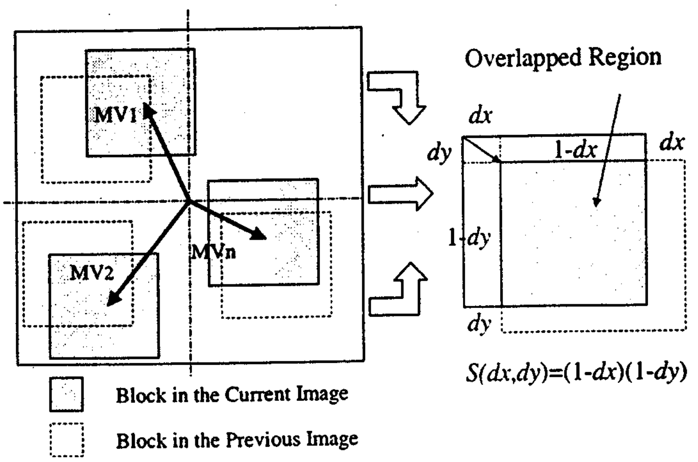


Figure 2. Correlation of Image Blocks

In this section, we analyze search patterns for motion estimation. Matching between two image blocks can be measured by their correlation. The overlapped region in Fig. 2 indicates the correlation between the current and the previous image blocks. Although we do not know the direction and the magnitude of the actual object displacement, we can derive an optimal search pattern for motion estimation by investigating equi-correlation contours as a function of the distances, *dx* and *dy*, between search positions in the search pattern.

2.2. Optimal Search Patterns

In order to make a simple analysis, we consider a rectangular image block whose pixel values are unity and its background has all zero values. The correlation between two rectangular image blocks can be calculated as the normalized area of the overlapped region, as shown in Fig. 2.

$$S(dx,dy) = (1-|dx|)(1-|dy|), \qquad 0 \le |dx| \le 1,\ 0 \le |dy| \le 1 \tag{1}$$

We note that *S(dx,dy)* has the following symmetry property:

$$S(dx,dy) = S(-dx,dy) = S(dx,-dy) = S(-dx,-dy) \tag{2}$$

From Eq. (1), we can find locations of displacement *(dx,dy)* having the same correlation value, and draw equi-correlation contours.

$$|dx \cdot dy| - |dx| - |dy| + 1 - S(dx,dy) = 0 \tag{3}$$

If we represent every position *(dx,dy)* in the block by the following relationship

$$dy = c \cdot dx\ , \qquad c \in \Re \tag{4}$$

We can get

$$dx = \frac{(c+1) - \sqrt{(c-1)^2 + 4c \cdot S(dx,dy)}}{2c} \tag{5}$$

$$dy = \frac{(c+1) - \sqrt{(c-1)^2 + 4c \cdot S(dx,dy)}}{2} \tag{6}$$

By changing the value of *c* from 0 to infinity in Eq. (5) and Eq. (6), we can plot equi-correlation contours of the diamond shape, as shown in Fig. 3. If the contour line is not aligned with a pixel position, we need to perform motion search at the closed pixel position.

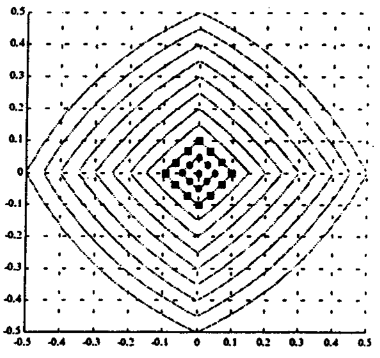


Figure 3. Equi-correlation Contours

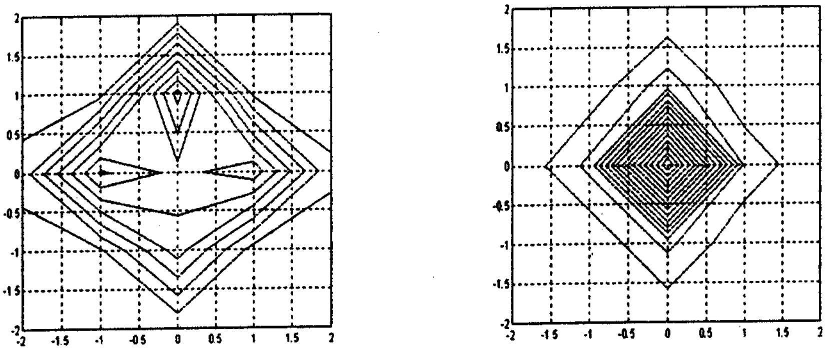
The derived optimal search pattern can be used to analyze conventional block matching algorithms. We note that the diamond search [8,10] and the advanced zonal search [9,10] are of the same characteristics.

2.3. Experimental Results for Optimal Search Patterns

In order to verify the derived search pattern, we perform computer simulations on ITU-T monochrome test sequences: MISS AMERICA, CLAIRE, FOOTBALL, CALENDAR and SALESMAN, each of which consists of 88 frames of the CIF format.

In each simulation, the original image is used as a reference frame to generate a motion-compensated prediction image. Fo FS BMA, the block size is 16×16 and the search region is [-7,+7]. In other words, we set $0 < |dx|, |dy| < 7/16 \approx 1/2$.

While Fig. 3 indicates optimal search patterns for $S(dx,dy) < 1/2$, Fig. 4(a) and Fig. 4(b) show the search patterns of motio vector fields for MISS AMERICA and SALESMAN, respectively. From Fig. 3 and Fig. 4, we observe that the derive optimal search patterns and the experimental ones are all diamond-shaped. It implies that our derivation is valid for th optimal search pattern for BMA with the rectangular block. The diamond search pattern is simple and efficient for fa BMA.



(a) Search Pattern (MISS AMERICA) (b) Search Pattern (SALESMAN)

Figure 4. Experimental Correlation Contours

3. BLOCK MATCHING ALGORITHMS USING OPTIMAL SEARCH PATTERNS

3.1. Criteria for Block Matching

The degree of matching of image blocks can be measured by various criteria[7], including the mean absolute differenc (MAD), the mean square error (MSE), and the matching pel count (MPC). Let $I_t(k,l)$ be the luminance pixel values in th current frame, and $I_{t-1}(k,l)$ be the luminance pixel values in the previous frame. Let (MV_x, MV_y) represent the displacemen vector. A square block of size NxN is used for motion estimation. Definitions of the typical matching criteria for BMA ar summarized below:

(a) Cross-Correlation Function (CCF)

$$CCF(i,j) = \frac{\sum_{k=1}^{N}\sum_{l=1}^{N} I_t(k,l) I_{t-1}(k+i,l+j)}{\left[\sum_{k=1}^{N}\sum_{l=1}^{N} I_t^2(k,l)\right]^{1/2} \left[\sum_{k=1}^{N}\sum_{l=1}^{N} I_{t-1}^2(k+i,l+j)\right]^{1/2}} \quad (7$$

(b) Mean Squared Error (MSE)

$$MSE(i,j) = \frac{1}{N^2}\sum_{k=1}^{N}\sum_{l=1}^{N}\left[I_t(k,l) - I_{t-1}(k+i,l+j)\right]^2 \quad (8$$

(c) Mean Absolute Difference (MAD)

$$MAD(i,j) = \frac{1}{N^2}\sum_{k=1}^{N}\sum_{l=1}^{N}\left|I_t(k,l) - I_{t-1}(k+i,l+j)\right| \quad (9$$

(d) Matching Pel Count (MPC)

$$T(k,l,i,j) = \begin{cases} 1, & if \left|I_t(k,l) - I_{t-1}(k+i,l+j)\right| \le THS \\ 0, & otherwise \end{cases} \quad (10$$

$$MPC(i,j)=\sum_{k=1}^{N}\sum_{l=1}^{N}T(k,l,i,j) \tag{11}$$

(e) Minimized Maximum Error (MiniMax)

$$MME(i,j)=\min_{(i,j)}\max_{(k,l)}\left|I_t(k,l)-I_{t-1}(k+i,l+j)\right| \tag{12}$$

While CCF and MPC select the position of the maximum value, the others choose the point of the minimum value. CCF and MSE require multiplications and accumulations, but the others require comparisons and accumulations. Since multiplication generally requires more hardware complexity than comparison, it is known to be more expensive to implement multiplication.

MAD is most widely used due to its lower complexity, and its performance is comparable to that of MSE. Although MPC requires less hardware complexity than MAD, its performance is quite sensitive to the selected threshold value. The computational complexity of the MiniMax criterion is lower than those of any other methods, but its performance is not guaranteed.

In order to reduce the computational complexity of the block matching operation, we define a new matching criterion by combining MAD and MPC.

$$T(k,l,i,j)=\begin{cases}1, & if\ \left|I_t(k,l)-I_{t-1}(k+i,l+j)\right|\geq THS\\ 0, & otherwise\end{cases} \tag{13}$$

$$SAD(i,j)=\sum_{k=1}^{N}\sum_{l=1}^{N}T(k,l,i,j)\left|I_t(k,l)-I_{t-1}(k+i,l+j)\right| \tag{14}$$

The main idea of the new matching criterion is that we only count pixel differences that have significant changes of luminance values. Since the pixels that have small changes of luminance values are not included in the computation, its complexity is reduced substantially. One remaining issue with the new matching criterion is how to select the threshold value THS properly.

Since the human visual system (HVS) is sensitive to large changes of luminance values, we can set a just noticeable difference (JND) as the threshold value for blocks of significant displacement. For blocks of small displacement, we set the threshold value to 0; then, Eq. (14) is equivalent to Eq. (9).

3.2. An Adaptive Motion Search Algorithm

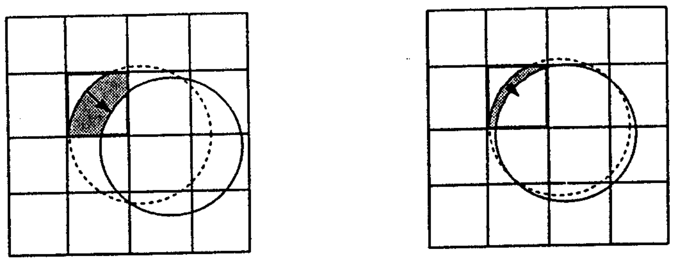
Although various BMAs have been proposed for fast motion estimation,[11] they are suboptimal in the sense that they sacrifice prediction accuracy at the expense of reducing computational requirement.

In the proposed motion estimation algorithm, we calculate SADs of Eq. (14) at search positions located in the diamond shape, that is derived in Section 2.2. This search pattern is more efficient than any other search patterns in conventional fast BMAs. It does not increase the total number of search positions for motion estimation. In the proposed algorithm, the number of search positions in the next step varies depending on the location of the minimum distortion position in the previous step.

In teleconferencing video sequences, most image blocks are regarded as stationary or quasi-stationary. Motion vectors for those stationary image blocks are mostly around (0,0). In order to determine whether a block is stationary or not, we consider the following situations. In general, a large object displacement would produce a large block difference (BD) within the search region (SR), as shown in Fig. 5. However, we may have the following exceptional cases.

Case 1: Even if the actual object displacement is large, BD is smaller than the threshold value. This situation can occur when similar blocks in the same image object move to the same direction by the same amount. Since it generates an aperture problem,[4] we cannot solve it satisfactorily. In this case, we simply assume that the block has a small displacement.

Case 2: Even if the actual object displacement is small, BD is larger than the threshold value. This case may happen when the background and the object have large luminance differences. If the number of pixels having significant luminance changes is small, we assume that this case occurred and the block has a small displacement.



(a) Large Motion (b) Small Motion

Figure 5. Object Movement and Block Difference

Fig. 6 explains the overall procedure of the proposed adaptive motion search (AMS) algorithm, where *Ns* denotes the number of significant pixels in the block.

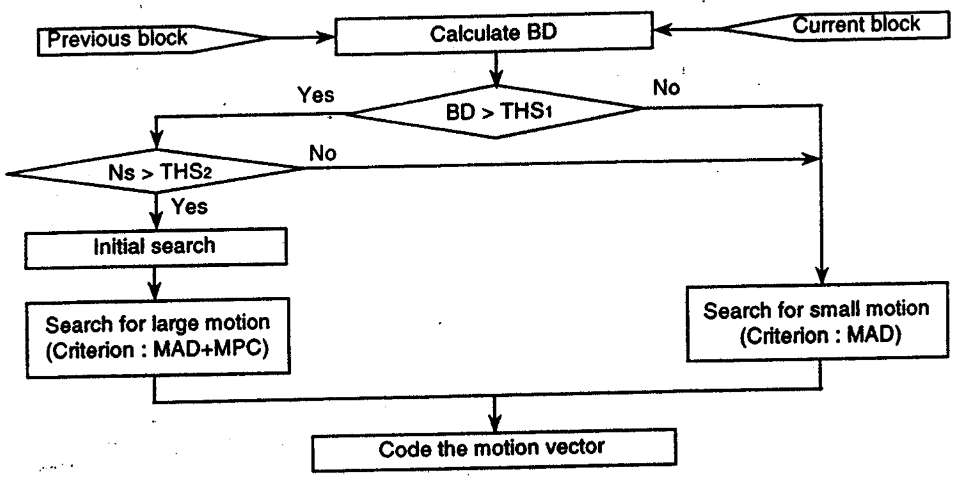
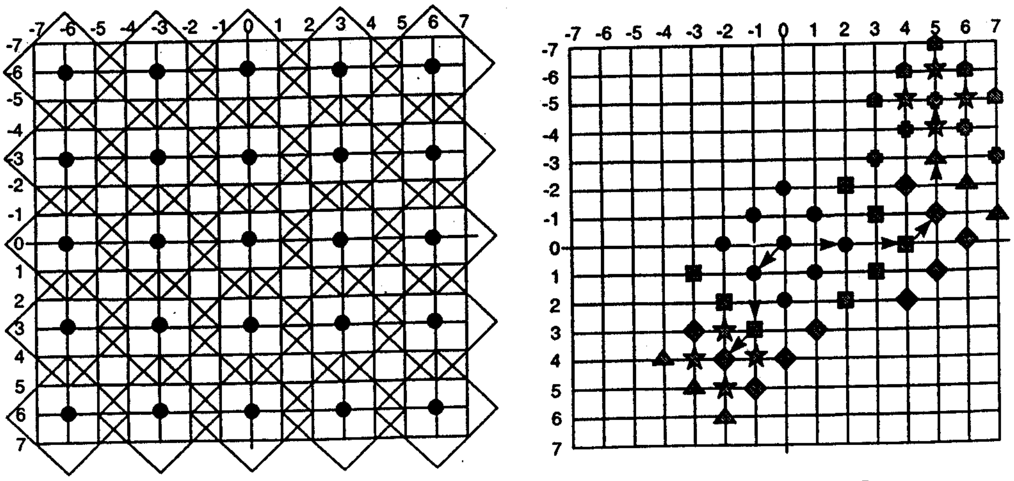


Figure 6. Adaptive Motion Search Algorithm

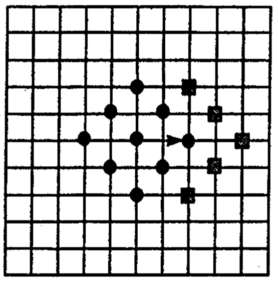
If BD is large, we use the initial search pattern, shown in Fig. 7(a). The initial search pattern consists of the regularly spaced lattice positions that cover the whole search region. Once a position of the minimum distortion is selected in the first search operation, other positions near the minimum distortion position are examined in the next stages, as illustrated in Fig. 7(b).



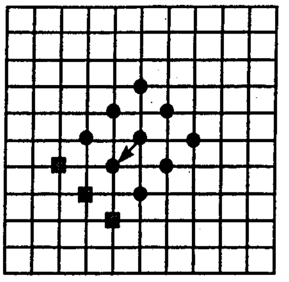
(a) Initial Search Pattern (b) Next Stage Search Patterns

Figure 7. Search Patterns for Large Motion

Fig. 8(a) and Fig. 8(b) explain two different search patterns where the search range is [-7, +7]. Depending on the position of the minimum distortion, we add three or five new search positions.



(a) Horizontal Direction



(b) Diagonal Direction

Figure 8. Search Pattern for Large Motion in Next Stages

In the second stage, nine positions are examined for the minimum distortion. Those nine positions are spaced by two-pixel distance in both the horizontal and the vertical directions, as shown in Fig. 8. This procedure is repeated until we find the minimum distortion in the center of the search pattern or the boundary of the serach region. The final displacement vector is the minimum distortion position among the one-pixel spaced positions around it.

If BD is small, the search region is limited to a local region. The procedure for small motion is similar to that for large motion, but we start with 3×3 neighboring positions, as shown in Fig. 9.

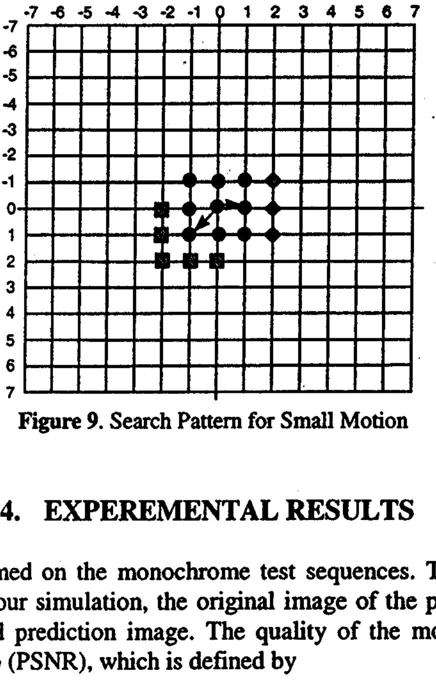


Figure 9. Search Pattern for Small Motion

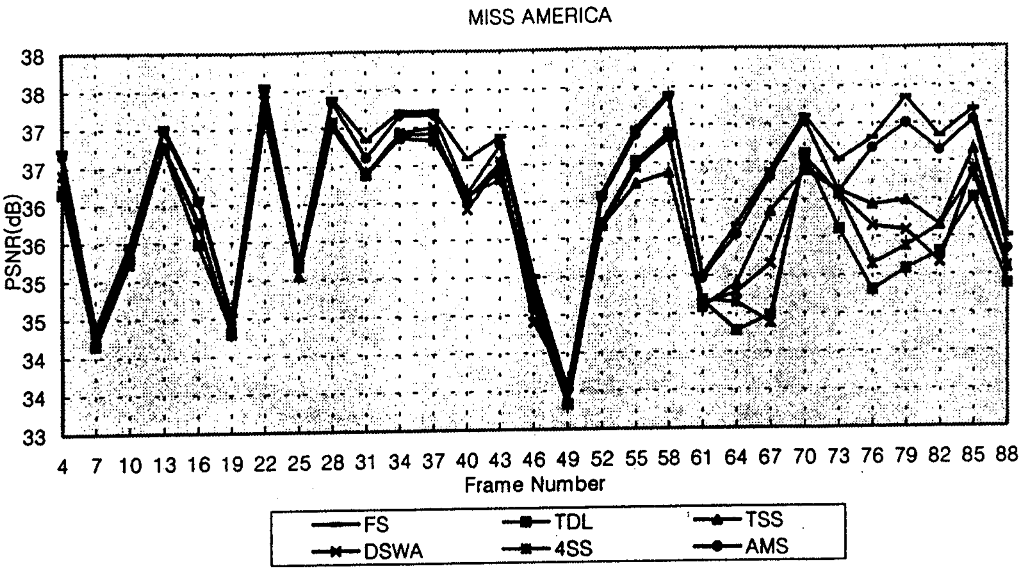
4. EXPEREMENTAL RESULTS

Computer simulations have been performed on the monochrome test sequences. The frame rate is reduced to 10 Hz by selecting 1st, 4th, 7th, ..., 88th frames. In our simulation, the original image of the previous frame was used as a reference frame to generate a motion-compensated prediction image. The quality of the motion-compensated prediction image is measured by the peak signal-to-noise ratio (PSNR), which is defined by

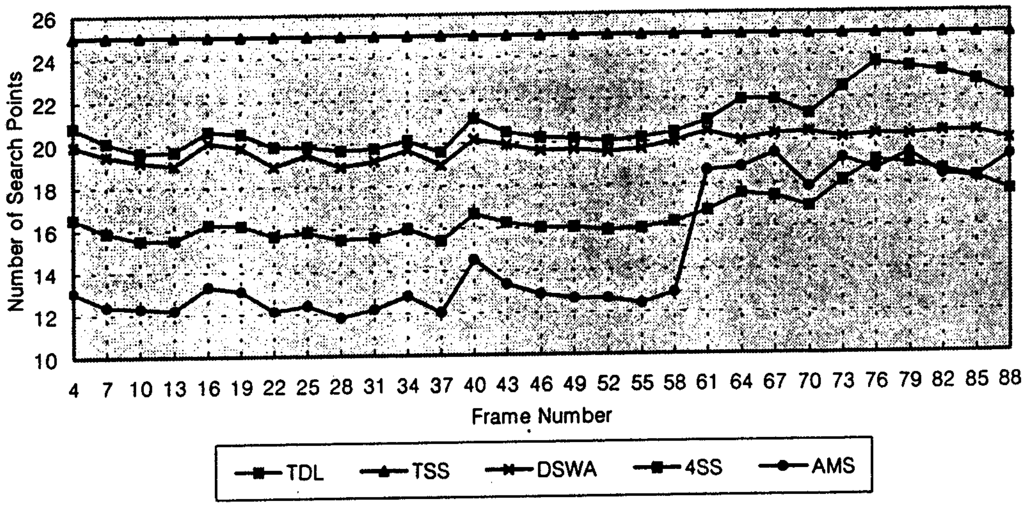
$$MSE = \frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}(I_t(m,n) - \tilde{I}_t(m,n))^2 \quad (15)$$

$$PSNR = 10\log_{10}\frac{255^2}{MSE} \quad [dB] \quad (16)$$

where M=352 and N=288 for the CIF format. I_t denotes the original current image frame, and $\tilde{I}_t$ denotes the motion-compensated prediction image frame.



(a) PSNR values



(b) Average Number of Search Positions

Figure 10. Performance Comparison

Table 1. Average Performance of Block Matching Algorithms

Algorithm	PSNR (dB)	Average Number of Search Positions
Full Search	36.29	225.00
Three Step Search	35.78	25.00
4-Step Search	35.81	20.91
2-D Logarithmic Search	35.62	16.62
Dynamic Search	35.77	19.79
Adaptive Motion Search	36.17	14.78

Note: Block Size =16, Search Region = [-7, +7]

The experimental results with MISS AMERICA are depicted in Fig. 10, where the proposed AMS(Adaptive Motion Searcl algorithm is compared with FS (full search), TDL (two-dimensional logarithmic search)[1], TSS (three-step search)[2], 4S (four-step search)[5], and DSWA (dynamic search window adjust and interlaced search)[6] algorithms. Average performances c simulation results are summarized in Table 1, where we note that AMS improves motion prediction accuracy and reduce the required computational complexity. The proposed AMS method provides better subjective image quality than any othe methods and it reduces the average number of search points. The matching criterion defined in Eq. (14) contributes to large reduction of the searching operation, compared to FS and other fast BMAs.

5. CONCLUSIONS

In this paper, we have proposed a new fast block matching motion estimation algorithm based on an optimal search pattern By a simple analysis, we have shown that the diamond search pattern is compact and optimal in terms of correlation o rectangular blocks. The proposed motion estimation algorithm employs the derived optimal search pattern and a nev matching criterion. It takes adaptive search operations depending on the amount of block differences, and reduces th required computational complexity drastically compared to other fast block matching algorithms, while maintaining gooc prediction accuracy.

ACHNOWLEDGEMENT

This work was supported in part by the Korea Science and Engineering Foundation (KOSEF) through the Ultra-Fast Fibe Optics Networks (UFON) Research Center at Kwangju Institute of Science and Technology (K-JIST), and in part by the Ministry of Education (MOE) through the Brain Korea 21 (BK21) project.

REFERENCES

1. J.R. Jain and A.K. Jain, "Displacement measurement and its application in interframe image coding," *IEEE Trans. Commun.*, vol. COM-29, pp. 1799-1808, Dec. 1981.
2. R. Srinivasan and K.R. Rao, "Predictive coding based on efficient motion estimation," *IEEE Trans. Commun.*, vol. COM-33, no. 8, pp. 888-896, Aug. 1985.
3. M. Ghanbari, "The cross-search algorithm for motion estimation," *IEEE Trans. Commun.*, vol. COM-38, pp. 950-953, July 1990.
4. H. Musmann, P. Pirsh and H. Grallert, "Advances in picture coding," *Proc. IEEE*, vol. 73, no. 4, pp. 523-548, April 1985.
5. L.M. Po and W.C. Ma, "A novel four-step search algorithm for fast block motion etsimation," *IEEE Trans. Circuit and Systems for Video Technology*, vol. 6, no. 3, pp. 313-317, June 1996.
6. L.W. Lee, J.F. Wang, J.Y. Lee and J.D. Shie, "Dynamic search-window adjustment and interlaced search for block-matching algorithm," *IEEE Trans. Circuit and Systems for Video Technology*, vol. 3, no. 1, Feb. 1993.
7. M.J. Chen, L.G. Chen, T.D. Chiueh and Y.P. Lee, "A new block-matching criterion for motion estimation and its implementation," *IEEE Trans. Circuit and Systems for Video Technology*, vol. 5, no. 3, pp. 231-236, June 1995.
8. S. Zhu and K. Ma, "A new diamond search algorithm for fast block matching motion estimation," *IEEE Trans. Image Processing*, vol. 92, pp. 287-290, Feb. 2000.
9. A. Tourapis, O. Au, M. Liou and G. Shen, "An advanced zonal block based algorithm for motion estimation," *Internat'l Conf. on Image Processing*, vol. 2, pp. 610-614, Oct. 1999.
10. T. Chiang and H. Sun, "Report of ad hoc group on encoder optimization," *ISO/IEC/JTC1/SC29/WG11 MPEG99/M5528*, Dec. 1999.
11. D.K. Lim and Y.S. Ho, "A fast block matching motion estimation algorithm based on statistical properties of object displacement," *IEEE Region 10 International Conference*, vol. 1, pp. 138-141, Jan. 1998.