INTERNATIONAL JOURNAL OF ENGINEERING AND MANAGEMENT SCIENCES

www.scienceandnature.org

TWO LEVEL BLOCK MATCHING CRITERION FOR MOTION ESTIMATION USING INTEGRAL FRAMES FEATURES

*Ravindra Kumar Purwar, Navin Rajpal
School of Information Technology,
GGSIP University, Kashmere Gate, Delhi – 110403, India

ABSTRACT
In block based motion estimation algorithms, computation is reduced by limiting the number of candidate search points within the search window or simplifying the distortion measurement criterion for block matching. Integral frame based block motion estimation algorithms have been suggested in literature which drastically reduce computation cost. However, these algorithms have a serious drawback of spurious block matching possibility, leading to poor quality results. In this manuscript, a two level block matching criterion based on integral frame concept is proposed to minimize this drawback. Experimental results show that an increase up to 12% in terms of peak signal to noise ratio (PSNR) (dB) has been achieved than conventional integral frame based block matching criterion with almost same execution time. Further, in terms of quality/computation ratio where quality and computation has been measured in terms of PSNR and execution time respectively, proposed method has 25-26% gain over existing integral frame based techniques.

KEYWORDS: Integral frame, motion estimation, mean absolute error, block matching criterion, video coding

INTRODUCTION

In video encoders, motion estimation has significant role in achieving high compression as it is used to remove temporal redundancy present in video sequences [12]. In literature, block based motion estimation algorithms (BMEA) are used for it because of their efficiency to predict motion of an individual object precisely. A video may consists of multiple objects moving independently. Each frame is divided in to a number of blocks and motion of each block is studied. As shown in figure 1, a current frame block is searched for its best matching block in the reference frame within a predefined search window and the location of the block which has best match with the current frame block is called the *motion vector* (MV) of the current block

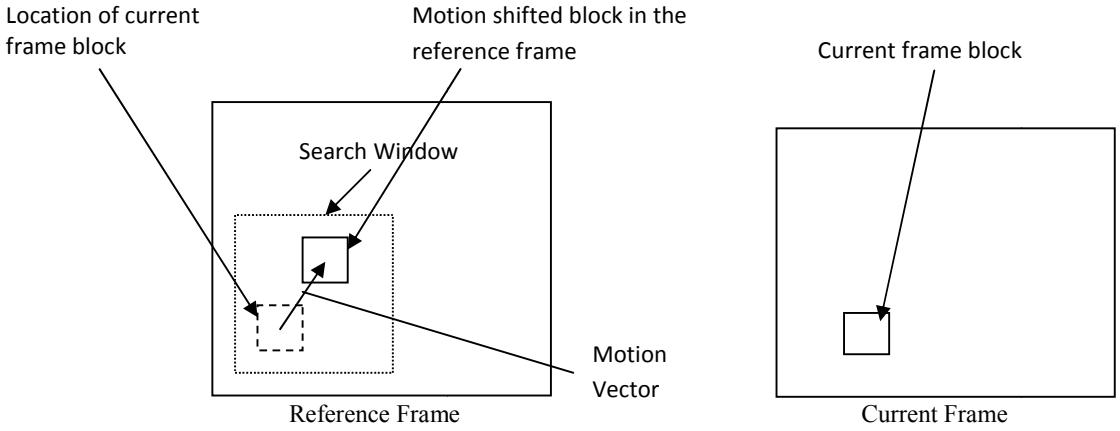

Fig. 1 Block based motion estimation

A number of BMEAs have been proposed in literature and full search algorithm produces optimal results as it explores all $(2p+1)^2$ search points for a search window of size ±p. Being exhaustive in nature, its computation cost is very high and it can not be used in real life applications. Therefore, many fast algorithms have been proposed which explore limited search points within the search window like 2-D Logarithmic Search [2], New Three Step Search [3], Four Step Search [8], Gradient Descent Search [4], Diamond Search [16,1], Hexagonal Search [14,15], Adaptive Rood Pattern Search [7] and Dynamic Pattern Search [11] etc. These BMEAs are faster and they use mean of sum of absolute difference (SAD) (also called conventional mean absolute error (MAE)) as distortion measure criterion for block matching. Eq. 1 gives the expression to compute SAD of a block centered at (i,j).

$$SAD(i,j) = \sum_{(x,y)} | c(x,y) - r(x+i,y+j) | \quad (1)$$

where, $-p \leq i,j \leq +p$ for search window size $\pm p$ and c(x,y) and r(x,y) are pixel values at position (x,y) in the current and reference frames respectively.

Sum of absolute difference in eq. 1 is computed using each individual pixel within the block causing more computation. For an NxN block, it requires N^2 subtraction operations, N^2 absolute differences and finally N^2-1 addition operations. Further, SAD does not produce good results for video inputs with contrast variations. Distortion measures using normalization of pixel intensities have been suggested in [9,10] which are robust to contrast variations in input videos but high computation is still a major bottleneck for real time video applications.

Recently, Nguyen and Tan [5,6], proposed low computation distortion measures namely sum of absolute difference using block sums (SAD_BS) and sum of absolute difference using block variances (SAD_VR) which are based on integral frame features [13] and produced comparable quality results with reduced computation. However, these integral frame based distortion measures have possibility of being trapped in spurious block matching as two or more different blocks in the search window may have same integral frame representation for SAD_BS or SAD_VR.

In this paper, a two level distortion measure criterion has been proposed which further explores multiple minimum points obtained using SAD_BS criterion in its next step to find final minimum point in each iteration of used motion estimation algorithm. Proposed criterion has not been experimentally compared with SAD_VR criterion as variance computation is costlier. This paper is organized as follows. Section 2 gives brief introduction of integral frame feature which is followed by SAD_BS technique in section 3. Section 4 explains the proposed two level block matching criterion and section 5 includes experimental results with computational analysis. Finally, section 6 concludes this paper.

INTEGRAL FRAME

For a given video frame f, its integral frame feature at point (i,j), denoted by $I_f(i,j)$, is defined as the sum of all pixel values that are above and to the left of pixel (i,j) including itself [13], i.e.

$$I_f(i,j) = \sum_{x=0}^{i} \sum_{y=0}^{j} f(x,y) \quad (2)$$

Let $R_f(i,j)$ denotes cumulative row sum of pixel values in frame f, defined as

$$R_f(i,j) = \sum_{x=0}^{i} f(x,j) \quad (3)$$

Assuming $R_f(-1,j) = 0$ and $I_f(i,-1) = 0$, integral frame I_f can be computed by using recursive formulas

$$R_f(i,j) = R_f(i-1,j) + f(i,j) \text{ and } I_f(i,j) = I_f(i,j-1) + R_f(i,j) \quad (4)$$

It is clear from eq. 4 that for frame of size MxN, its integral frame feature can be computed using only 2 MN additions.

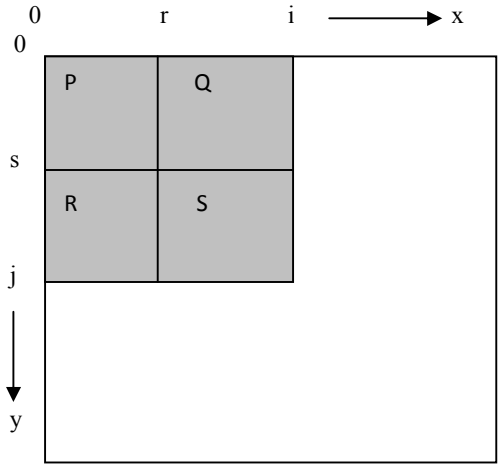

Fig. 2 Computation of block sum for block S using four integral frame values

It can be seen from figure 2 that the block sum of block S in frame f (referred as $BS_f(S)$ can be computed as

$$BS_f(S) = \sum_{x=r+1}^{i} \sum_{y=s+1}^{j} f(x,y) = I_f(i,j) - I_f(r,j) - I_f(i,s) + I_f(r,s) \quad (5)$$

using four corresponding values of the integral frame with only one addition and two subtraction operations.

SAD_BS TECHNIQUE

Authors in [5] have proposed motion estimation algorithm using integral frames to minimize the number of computations. For block matching, each block is partitioned in a number of sub-blocks and block sum of each sub-block is then computed using eq. 5. To find the best matching block in the reference frame, the block sums of all sub-blocks in the current block are compared with their corresponding sub-blocks in the reference frame block. Specifically, the sum of absolute differences

between the corresponding block sums (SAD_BS) are used for block matching, which is given as -

$$SAD\_BS = \sum_{i} | BS_{fc}(S_i) - BS_{fr}(S_i) | \quad (6)$$

where $BS_{fc}(S_i)$ and $BS_{fr}(S_i)$ denote the block sum of i^{th} sub-block from the current and reference block respectively.

It can be seen from eq. 6 that the performance of SAD_BS based block matching algorithm depends on how the block is partitioned into sub-blocks. Authors have carried out a number of experiments and finally opted for four symmetrical sub-block division as a tradeoff between quality and computation.

PROPOSED TWO LEVEL BLOCK MATCHING CRITERION

Block sum based matching algorithm using integral frame drastically reduces computational cost with respect to conventional SAD (Eq. 1) but it has possibility of being trapped in spurious block matching as different blocks may have same integral frame feature representation. To minimize this drawback, a two level matching criterion is being proposed in this section which uses two or more minimum distortion points obtained using SAD_BS technique to further explore them in its next step using conventional MAE to find out the final minimum distortion point. The algorithm is explained using the flow chart in figure 3.

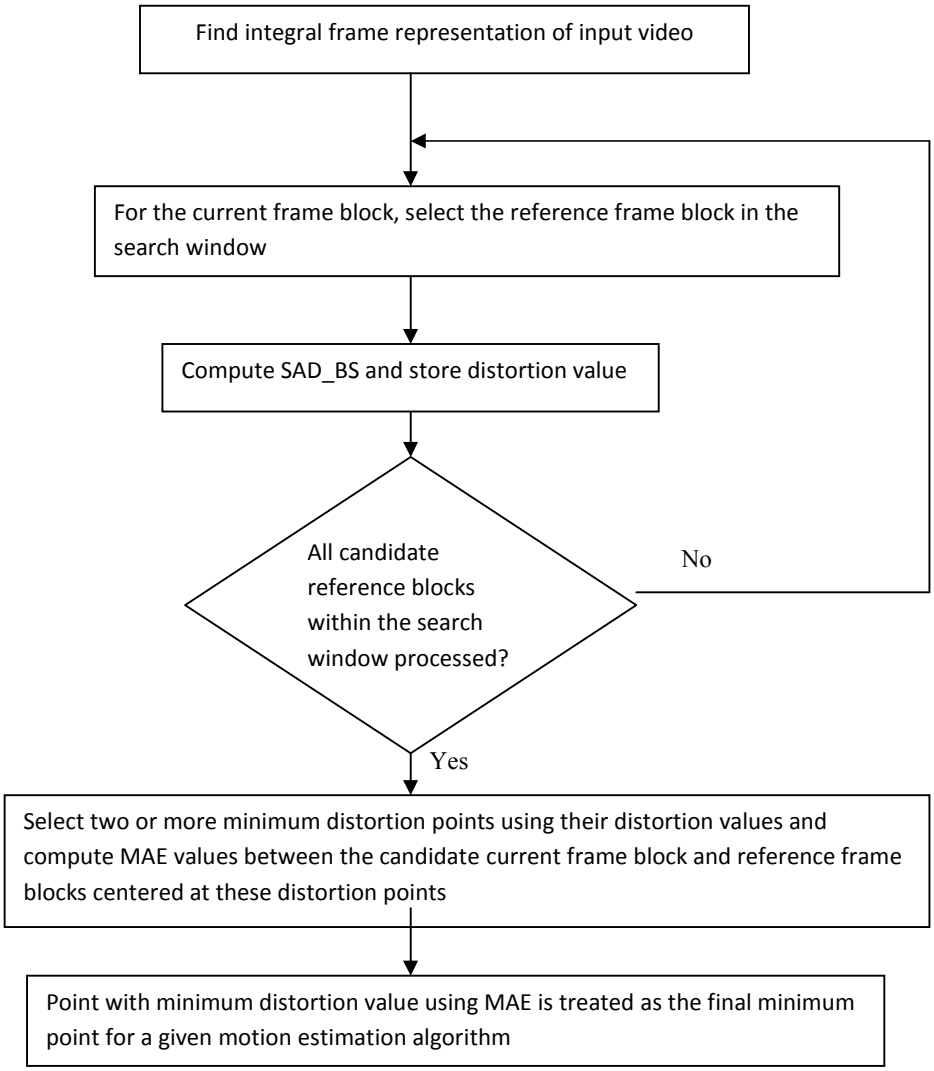

Fig. 3 Flow chart of proposed two level block matching criterion

As shown in figure 3, the proposed algorithm first explores minimum distortion points using SAD_BS technique suggested by authors in [5] but to reduce chances of spurious matching, obtained minimum point is further explored using MAE. To add accuracy in matching, additional minimum points (in the order of their values) are also explored at that stage. Out of these points, final minimum distortion point is obtained using MAE distortion measure. Accuracy of the proposed algorithm is governed by the number of minimum points selected after applying SAD_BS criterion. In this paper, simulation results have been performed using one and two additional

points only as increasing the number of additional points will increase computation cost and results are found to be comparable with pure MAE distortion measure criterion.

EXPERIMENTAL RESULTS AND ANALYSIS

Performance evaluation of the proposed two level block matching criterion using integral frames has been carried out with respect to the conventional sum of absolute difference (SAD) and sum of absolute difference using block sum (SAD_BS) in terms of two parameters -- average PSNR (dB) and execution time (seconds). Full search technique has been used as motion estimation algorithm and nine video sequences with different features have been taken as input. Search window size used for motion estimation is ±15 and size of each block is 16x16. Further, four sub-blocks representation of a block has been used for comparison purposes. All simulation work has been carried out in MATLAB 7.4.

Table 1 shows the performance comparison of proposed criterion in terms of PSNR for frame distance one and it can be seen that it has gain of nearly 5% and 10% over SAD_BS for one and two additional minimum points respectively whereas table 2 shows the same for frame distance two and the proposed criterion has slightly more PSNR gain of nearly 9% and 12% over SAD_BS for these additional points respectively. Moreover, PSNR for the proposed criterion is almost same to that of SAD in these tables.

Table 1: Performance comparison in terms of PSNR (dB) for frame distance 1

Video inputs	SAD	SAD_BS	Two level matching with one additional point	Two level matching with two additional points
Cactus comb (240x352)	24.55	21.32	22.95	23.98
Carphone (144x176)	24.61	21.89	23.88	23.88
Flower (240x352)	20.01	16.68	18.89	19.86
Miss America (144x176)	36.91	32.44	35.67	36.69
Tsunami (240x320)	26.73	22.43	25.25	25.25
Trojan lift (240x320)	20.58	18.97	19.81	20.58
India pride (240x320)	21.11	19.07	20.28	20.97
Susie (240x352)	39.27	36.69	37.95	39.25
Airbus (288x352)	31.16	26.06	29.49	30.88

Table 2: Performance comparison in terms of PSNR (dB) for frame distance 2

Video inputs	SAD	SAD_BS	Two level matching with one additional point	Two level matching with two additional points
Cactus comb (240x352)	21.81	19.96	20.87	20.87
Carphone (144x176)	24.75	21.96	23.99	24.75
Flower (240x352)	15.15	13.83	14.76	15.05
Miss America (144x176)	34.81	29.54	32.97	33.78
Tsunami (240x320)	24.82	21.19	23.76	24.82
Trojan lift (240x320)	18.98	17.75	18.98	18.98
India pride (240x320)	20.09	18.51	19.63	19.97
Susie (240x352)	37.05	34.83	36.61	37.05
Airbus (288x352)	35.09	30.12	33.86	34.68

Table 3: Performance comparison in terms of execution time (sec.) for frame distance 1

Video inputs	SAD	SAD_BS	Two level matching with one additional point	Two level matching with two additional points
Cactus comb (240x352)	15.920	4.013	4.049	4.121
Carphone (144x176)	4.326	1.057	1.066	1.098
Flower (240x352)	17.703	4.736	4.910	5.212
Miss America (144x176)	4.825	1.153	1.183	1.206
Tsunami (240x320)	16.123	4.116	4.219	4.324
Trojan lift (240x320)	14.779	4.748	4.759	4.806
India pride (240x320)	14.800	3.610	3.640	3.702
Susie (240x352)	16.180	4.095	4.106	4.147
Airbus (288x352)	19.892	5.020	5.094	5.216

In table 3 and table 4, execution time (sec.) for these block matching criterions has been shown for frame distance 1 and 2 respectively. Values have been shown up to three decimal places for more accurate analysis. It can be seen that there is marginal increase in execution time for the proposed two level matching criterion than SAD_BS whereas this execution time is quite high for SAD technique.

It has been found that proposed criterion produces good quality results with slightly more execution time. Therefore, one parameter - quality/computation ratio where quality and computation has been measured in terms of PSNR (dB) and execution time (sec.) respectively

has been used for analysis. Table 5 shows the performance of various criterions for this parameter and it can be seen that the proposed two level matching criterion has 25-26% gain over SAD_BS in terms of this ratio. Frame distance of one has been used for this parameter.

Table 4: Performance comparison in terms of execution time (sec.) for frame distance 2

Video inputs	SAD	SAD_BS	Two level matching with one additional point	Two level matching with two additional points
Cactus comb (240x352)	15.915	4.045	4.080	4.117
Carphone (144x176)	4.327	1.005	1.064	1.094
Flower (240x352)	17.899	4.595	4.620	4.641
Miss America (144x176)	4.759	1.187	1.206	1.258
Tsunami (240x320)	16.102	4.074	4.131	4.192
Trojan lift (240x320)	16.030	4.166	4.198	4.217
India pride (240x320)	14.725	3.704	3.732	3.765
Susie (240x352)	16.147	4.530	4.580	4.859
Airbus (288x352)	19.827	5.261	5.311	5.357

Table 5: Performance comparison in terms of quality/computation ratio

Video inputs	SAD	SAD_BS	Two level matching with one additional point	Two level matching with two additional points
Cactus comb (240x352)	1.54	5.31	5.67	5.82
Carphone (144x176)	5.69	20.71	22.40	21.75
Flower (240x352)	1.13	3.52	3.85	3.88
Miss America (144x176)	7.65	28.14	30.15	30.42
Tsunami (240x320)	1.66	5.45	5.98	5.84
Trojan lift (240x320)	1.39	4.00	4.16	4.28
India pride (240x320)	1.43	5.28	5.57	5.66
Susie (240x352)	2.43	8.96	9.24	9.46
Airbus (288x352)	1.97	5.19	5.79	5.92

Further, to study the PSNR and execution time for entire video sequence, plots are shown in figure 4 and figure 5 for frame distance one. These performance parameters have been plotted with respect to frame numbers of susie video sequence and it can be seen that the proposed criterion is producing better PSNR with almost same execution time than SAD_BS criterion. Notations are shown in the corresponding legends of the figure.

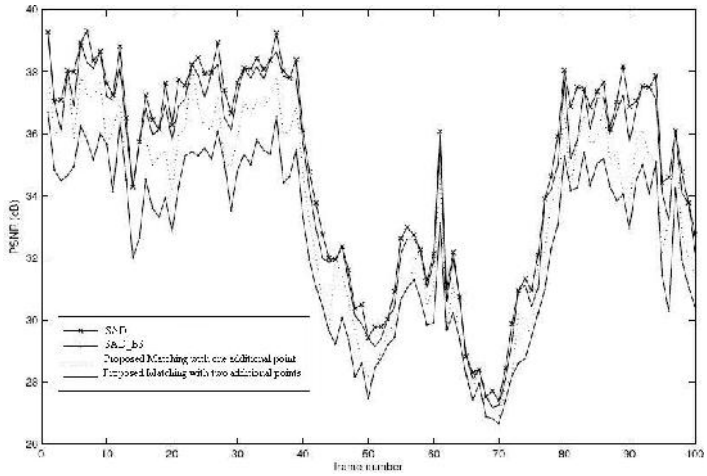

Fig. 4 Performance analysis of PSNR versus frame number for susie video

Authors in [5] have analyzed the computational gain of their SAD_BS technique over conventional SAD for full search algorithm and the number of arithmetic operations per frame required by SAD and SAD_BS measures are $3MN*(2W+1)^2$ and $2MN + 21*(2W+1)^2*(MN/16^2)$ respectively for MxN frame with search area size of ±W. Proposed two level matching criterion will use $k*3*16^2$ additional operations over SAD_BS for k additional points using full search. It is because MAE (used at next level) for each 16x16 block will use 3 arithmetic operations - one subtraction, one addition and one absolute difference.

CONCLUSION

In this manuscript, a two level block matching criterion for motion estimation has been suggested which is an extension of SAD_BS technique based on integral frames. The proposed criterion uses multiple minimum points obtained by SAD_BS technique and it further explores

them using conventional MAE to find the final minimum distortion point. A significant gain in PSNR has been found with marginal increase in computation. Moreover, in terms of quality/computation ratio, proposed criterion has an average gain of 25-26% over SAD_BS technique.

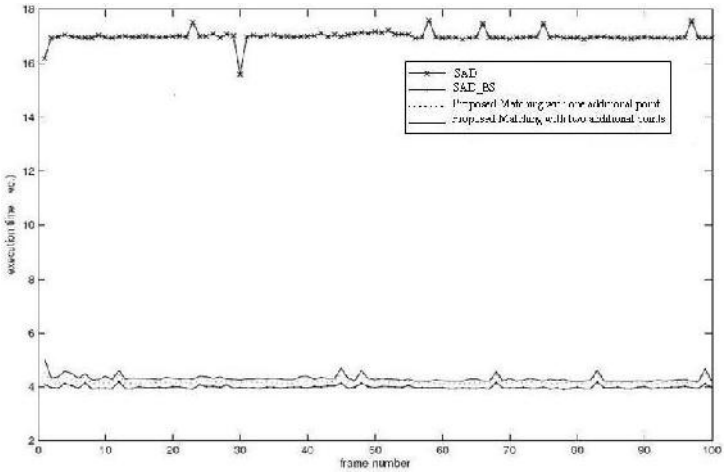

Fig. 5 Performance analysis of execution time versus frame number for susie video

REFERENCES

1. Cheung, C. H. and Po, L.M. (2002) A novel cross diamond search algorithm for fast block motion estimation. IEEE Transactions in circuit systems and video technology. 12, 1168-1177.

2. Jain, J. R. and Jain, A.K. (1981) Displacement measurement and its application in interframe image coding. IEEE Transactions in communication. 29, 1799-1808.

3. Li, R., Zeng, B. and Liou, M. L. (1994) A new three step search algorithm for block motion estimation. IEEE Transactions in circuit systems and video technology. 4, 438-442.

4. Liu, L.K. and Fieg, E. (1996) A block based gradient descent search algorithm for block motion estimation in video coding. IEEE Transactions in circuit systems and video technology. 6, 419-423.

5. Nguyan, V. A. and Tan, Y. P. (2004) Fast block based motion estimation using integral frames. IEEE signal processing letter. 11, 744-747.

6. Nguyan, V. A. and Tan, Y. P. (2006) Efficient block matching motion estimation based on integral frame attributes. IEEE Transactions in circuit systems and video technology. 16, 375-385.

7. Nie, Y. and Ma, K. K. (2002) Adaptive rood pattern search for fast block matching motion estimation. IEEE Transactions in image processing. 11, 1442-1449.

8. Po, L. M. and Ma, W. C. (1996) A novel four step search algorithm for fast block motion estimation. IEEE Transactions in circuit systems and video technology. 6, 313-317.

9. Purwar, R. K., Prakash, N., Rajpal, N. (2008) A block matching criterion for interframe coding of video. IEEE international conference on audio language and image processing held at Sanghai, China. 7-9 July.

10. Purwar, R. K., Prakash, N., Rajpal, N. (2009) A matching criterion for motion compensation in the temporal coding of video signal. Springer Journal of signal image and video processing. DOI: 10.1007/s11760-009-0149-9 (online first).

11. Purwar, R. K., Rajpal, N. and Prakash, N. (2009) A dynamic pattern search for real time block matching motion estimation. International Journal of imaging. 2, 15-22.

12. Sikora, T. (1997) MPEG digital video coding standards. IEEE signal processing magazine. 82-100.

13. Viola, P. and Jones, M. (2004) Robust real time face detection. International Journal of computer vision. 57, 137-154.

14. Zhu, C., Lin, X. and Chau, L. P. (2002) Hexagon based search pattern for fast block motion estimation. IEEE Transactions in circuit systems and video technology. 12, 349-355.

15. Zhu, C., Lin, X., Chau, L. P. and Po, L. M. (2004) Enhanced hexagonal search for fast block motion estimation. IEEE Transactions in circuit systems and video technology. 14, 1210-1214.

16. Zhu, S. and Ma, K. K. (2000) A new diamond search algorithm for block matching motion estimation. IEEE Transactions in circuit systems and video technology. 9, 287-290.