A NEW CROSS-DIAMOND SEARCH ALGORITHM FOR FAST BLOCK MATCHING MOTION ESTIMATION

Chi-Wai Lam [1], Lai-Man Po [1] and Chun Ho Cheung [1,2]
[1] Department of Electronic Engineering City University of Hong Kong, Hong Kong SAR
[2] Department of Information Technology, Hong Kong Institute of Technology, Hong Kong SAR
cwlam@ee.cityu.edu.hk, eelmpo@cityu.edu.hk, terence@ieee.org

ABSTRACT

In order to fit the small cross-center-biased characteristic of the real world video sequences, an improved version of the well-known cross diamond search algorithm is proposed in this paper. The new algorithm uses a small cross-shaped search patterns in the first two steps to speedup the motion estimation of stationary and quasi-stationary blocks. Experimental results show that this new cross-diamond search algorithm could achieve much higher computational reduction as compared with Diamond Search (DS) and Cross Diamond Search (CDS) while similar prediction accuracy is maintained, and it is especially suitable for videoconferencing sequences. **Keywords**—Motion estimation, new cross-diamond search, cross-center biased characteristic.

1. INTRODUCTION

Motion estimation (ME) is a process to estimate the pels or pixels of the current frame from reference frame(s). Block matching motion estimation or block matching algorithm (BMA), which is temporal redundancy removal technique between 2 or more successive frames, is an integral part for most of the motion-compensated video coding standards. Frames are being divided into regular sized blocks, or so-called macroblocks (MB). Block-matching method is to seek for the best-matched block from the previous frame, usually the first single frame, within a fixed-sized of search window (*w*). Based on a block distortion measure (BDM) or other matching criteria, the displacement of the best-matched block will be described as the motion vector (MV) to the block in the current frame. The best match is usually evaluated by a cost function based on a BDM such as Mean Square Error (MSE), Mean Absolute Error (MAE) or Sum of Absolute Differences (SAD). Due to the high intensive computation required for the full search (FS) method, which performs searching all the candidate blocks within the search window exhaustively. Over the last two decades, many fast motion estimation algorithms has been proposed to give a faster estimation with similar block distortion compared to FS. The most well known fast BMA are the three-step search (3SS) [1], the new three-step search (N3SS) [2], the four-step search (4SS) [3], the diamond search (DS) [4], the cross-diamond search (CDS) [5] and small cross-diamond search [6]. As the characteristic of center-biased motion vector distribution (MVD) which inspired many fast BMA in last decade, more than 80% of the blocks can be regarded as stationary or quasi-stationary blocks, i.e. most of the motion vectors are enclosed in the central 5x5 (blocks) area. This center-based characteristic can even be found in the fast-motion sequences. To exploits this phenomenon, NTSS added 8 center-neighboring blocks and introduced a halfway-stop technique to achieve crucial speedup for stationary and quasi-stationary blocks. 4SS also exploits the center-biased properties of motion vectors distribution by using halfway-stop techniques and smaller square search pattern compared to 3SS. DS was proposed with two novel ideas: a diamond shape searching pattern and unrestricted searching steps. DS is a highly center biased by using a compact diamond search pattern, and the unrestricted searching steps is used for reducing the chances of being trapped by local optima. In this paper, to fit cross-center -biased MVD property of the most real world sequences, a novel fast BMA called new cross-diamond search algorithm (NCDS) is proposed. It uses a small cross-shaped search patterns in the first two steps and speed results in higher the motion estimation of stationary and quasi-stationary blocks. Similar starting pattern can be found in [7]. Experimental simulations show that it can achieve fewer search points over other BMAs, such as DS and CDS, and can obtain the similar distortion performance. The result also shows that it is favorable in videoconferencing sequences. This paper is organized as follows. The second section introduces the Cross Center Biased MVD property. The third section presents the details of the new cross-diamond search algorithm. The fourth section describes the experimental result and the performance evaluation. Some concluding remarks are given in the last section.

2. CROSS CENTER-BIASED MVP DISTRIBUTION

There is a well-known property of image sequences - *The block motion field of a real world image sequence is usually gentle, smooth, and varies slowly.* To demonstrate the property of the global minimum motion vector distribution, by applying FS with spiral block-matching style and MAD as the BDM on the six well-known real-world image sequences which is listed in Table 1, the average motion vector probabilities (MVP) distributions was tabulated in Table 2.

Frame Format (Numbers of frames)	*Sequences*
CIF (352x288, 80 frames)	Miss America Sales Claire
SIF (352x240, 80 frames)	Tennis Garden Football

TABLE I: VIDEO SEQUENCES USED FOR ANALYSIS

The CIF sequences can be regarded as low motion (video conference) video, including "Miss America", "Salesman", and Claire". These sequences is relatively gentle, smooth, and with low-motion content. Whereas another three SIF video sequences "football", "garden" and tennis" are relatively with high motion content. Zooming, fast movement object, and panning can be

Probablities (%) at corresponding checking-point within the search window

r_Ver \ r_Hor	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7
7	0.0624	0.02862	0.04252	0.02662	0.04	0.0543	0.03968	0.37227	0.03588	0.0441	0.02432	0.01885	0.03295	0.01768	0.0483
6	0.01337	0.02327	0.00622	0.01483	0.0122	0.02285	0.04115	0.11553	0.05008	0.01517	0.01527	0.00968	0.01053	0.0201	0.01315
5	0.00643	0.00652	0.00517	0.00358	0.01188	0.0239	0.0161	0.11638	0.05063	0.01348	0.00517	0.0061	0.00412	0.0039	0.00642
4	0.03283	0.05345	0.0181	0.0342	0.05188	0.07123	0.11227	0.18973	0.18613	0.04883	0.02842	0.02683	0.01853	0.03798	0.01915
3	0.05125	0.02778	0.04967	0.03033	0.09733	0.25547	0.34103	0.77127	0.35733	0.17993	0.06618	0.02578	0.03895	0.02012	0.04662
2	0.01537	0.02578	0.00643	0.02452	0.06072	0.25075	0.69938	**0.624**	0.41003	0.1612	0.03422	0.02737	0.01748	0.01957	0.02347
1	0.0281	0.01832	0.02442	0.05295	0.1287	0.41457	**0.738**	**4.216**	**0.623**	0.2538	0.08102	0.04705	0.03682	0.0283	0.04032
0	0.1614	0.36468	0.09355	0.22695	0.39983	**0.814**	**4.378**	**45.49**	**3.44**	**0.516**	0.30798	0.22548	0.11332	0.25085	0.131
-1	0.02843	0.00758	0.02525	0.04832	0.11068	0.39058	**0.757**	**11.49**	**0.664**	0.34932	0.10743	0.05103	0.03093	0.01305	0.02915
-2	0.01505	0.03313	0.01212	0.04733	0.07207	0.18257	0.38333	**3.776**	0.52453	0.24832	0.05852	0.03872	0.01843	0.02008	0.01598
-3	0.04513	0.02115	0.04695	0.0324	0.09555	0.25347	0.43328	**2.95938**	0.23842	0.23212	0.05493	0.03313	0.03808	0.02758	0.04778
-4	0.01483	0.03935	0.0118	0.02958	0.05093	0.07218	0.45172	0.67857	0.1452	0.10112	0.05567	0.03378	0.0181	0.04398	0.01347
-5	0.00832	0.00547	0.00768	0.0098	0.01338	0.02935	0.64438	0.57385	0.05008	0.04937	0.01917	0.01935	0.01273	0.00812	0.01022
-6	0.0239	0.02063	0.01453	0.0159	0.01537	0.022	0.17688	0.29398	0.03218	0.03693	0.02747	0.02462	0.01525	0.02598	0.01105
-7	0.06018	0.01463	0.036	0.01557	0.04535	0.04715	0.0524	0.35742	0.06335	0.06788	0.06135	0.06377	0.05252	0.02537	0.0684

TABLE II: Average distribution measured at distance r using 6 CIF/SIF sequences for $|w| = 7$

found in these 3 sequences. From observing the motion vector probabilities distributions on different sequences, we found that most real-world sequences has the center biased MV distribution characteristic (over 80% MVP of the blocks having motion vectors within central 5×5 grid or radius-$r = \pm 2$), instead of an uniform distribution. The result also shows that the cross-center MV distribution is more dominant within this radius. For instance, in Fig.1, 71.76% of the motion vectors are found located in the central 2×2 area, i.e., A+B+D or $r = \pm 1$. And there is about 68.98% of motion vector are located in A+B or the cross-center area. To look at 3×3 area (A+B+C+D+E), total MVP is 81.75% and the cross-center probabilities within this area (A+B+C) has accounted for 74.71%. Inspired from such a highly cross-center based distribution, the searching pattern of BMA can be matched the cross-center shape to minimize the searching point while maintain a similar distortion error.

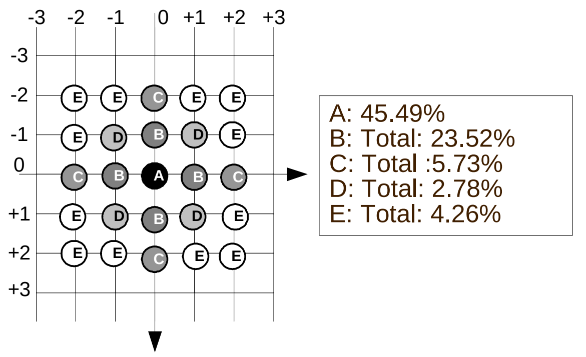


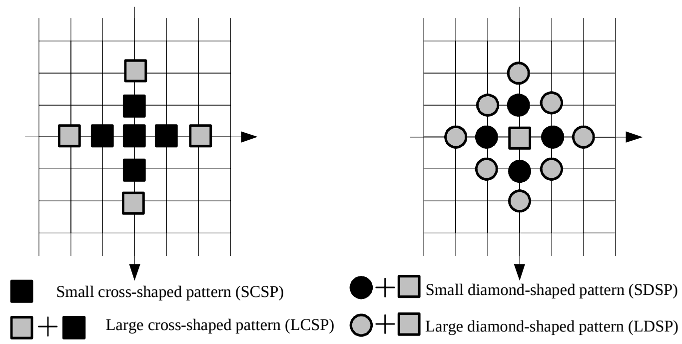
Fig. 1. The MVD within 3×3 area

3. NEW CROSS DIAMOND SEARCH (NCDS) ALGORITHM

In this section we first to describe search pattern used in the algorithm, and later the search path strategy will be explained.

A. Search Patterns

The search-point configuration used in the NCDS is divided in 2 different shapes: Cross-shaped pattern and diamond-shaped pattern. Fig.2 (a) show the small cross-shaped pattern (SCSP) and the large cross-shaped pattern (LCSP). The same search pattern from DS: small diamond-shaped pattern (SDSP) and large diamond-shaped pattern (LDSP) are shown in Fig.2 (b).



(a) (b)

Fig. 2 Search Patterns used in the new-cross-diamond search

B. The NCDS algirhtm

From the stimulation result on those six well-known sequences, we found that nearly 70% blocks that can be regarded as stationary ($r = 0$) or quasi-stationary blocks ($r = 1$). By having this highly cross center-biased property in most real world sequences, we take the small cross-shaped patterns as first two steps to the NCDS. The main difference between NCDS and cross diamond search (CDS) is that the first 2 steps of NCDS are keeping a small cross shaped pattern which is saving the number of search point for stationary or quasi-stationary blocks. The details and the analysis of the algorithm are given below:

- ***Step 1*** *(Starting - Small Cross Shape Pattern SCSP):* A minimum BDM is found from the 5 search points of the SCSP (Small Cross-Shaped Pattern) [Fig.1 (a)] located at the center of search window. If the minimum BDM point occurs at the center of the SCSP (0,0), the search stops (First Step Stop); otherwise, go to Step 2.
- ***Step 2*** *(SCSP):* With the vertex (minimum BDM point) in the first SCSP as the center, a new SCSP is formed. If the minimum BDM point occurs at the center of this SCSP, the search stops (Second Step Stop); otherwise go to Step 3.
- ***Step 3*** *(Guiding - Large Cross Shape Pattern - LCSP):* The three unchecked outermost search points of the central - LCSP

are checked, in which the step is trying to guide the possible correct direction for the subsequent steps. And then go to Step 4.

- ***Step 4*** *(Diamond Searching)*: A new Large Diamond Search Pattern LDSP is formed by repositioning the minimum BDM found in previous step as the center of the LDSP. If the new minimum BDM point is at the center of the newly formed LDSP, then go to Step 5 for converging the final solution; otherwise, this step is repeated.
- ***Step 5*** *(Ending – Converging step):* With the minimum BDM point in the previous step as the center, a SDSP (Small Diamond-Shaped Pattern) is formed. Identify the new minimum BDM point from the SDSP, which is the final solution for the motion vector.

C. Analysis of NCDS algorithm

To compare the CDS and the DS, the main improvement of this algorithm is the speed performance (the number of searching point). NCDS reduces the number of search points significantly if there is stationary block or quasi-stationary blocks. To fit the cross-center-biased MV distribution characteristics, it provides more chance to save up the searching points for motion vectors. In Fig 3, it shows 4 typical examples of NCDS and each candidate point is marked with the corresponding step number. Fig.3 (a) and (b) show two halfway-stop examples. The NCDS only takes 5 (first step stop) and 8 (second step stop), whereas the CDS took 9 and 11 search points, and the DS took 13 search points for searching the same block respectively. Another two search paths for $r > 1$ are shown in Fig 3 (c) and (d). Due to the chance of being trapped to local minimum if the algorithm is keeping a SCSP in first and second search step, a guiding step in Step 3 of NCDS is trying to guide a possible correct direction by using a larger search pattern when $r > 1$ (the minimum BDM point is till on the vertex of the second SCSP in step 2). Thus, we employed LCSP step to avoid the algorithm being trapped by local minimum, which will influence the distortion performance. After step 3, the subsequent steps will be exactly the diamond search. This algorithm is suitable for fast and a low bit-rate video conferencing application, which is gentle and small motion.

4. EXPERIMENTAL RESULTS

In our simulations, the BDM is defined to be the mean absolute error (MAE). The block size is at 16 ×16, and the maximum displacement in the search areas is ±7 pixels in both the horizontal and the vertical directions. The simulation is performed with a total of six sequences with different degrees and types of motion content as listed in Table.1. We compared the NCDS against DS, CDS and SCDS using the following test criteria: 1) Average searching point (ASP) – the average number of search point used to find the motion vector; and 2) Average MAE per pixel– This shows the magnitude of distortion per pixel. Table 3 and 4 summarize the experimental results of each search strategy over the test criteria using the 6 tested sequences.

And the speed/MAD improvement in percentage of the NCDS over DS and CDS are tabulated in Table V. By observing the result obtained in Table III, The NCDS takes the smallest average number of search points per block (i.e. the fastest algorithm) among other fast BMA for all of the six test sequences.

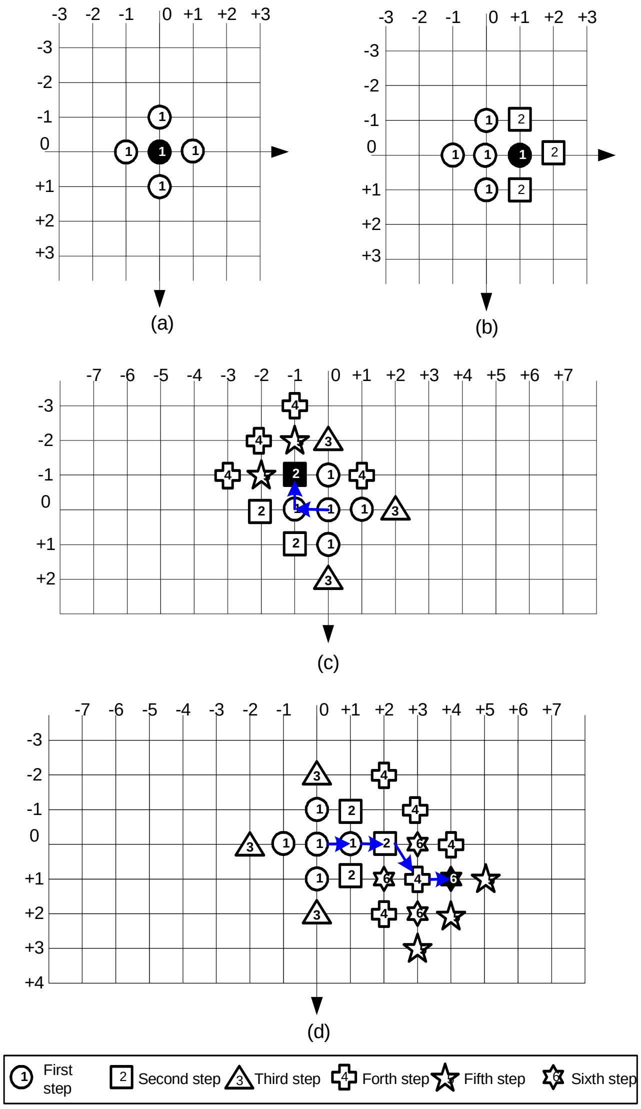


Fig. 3.Examples of the NCDS: (a) first-step-stop with MV(0,0). (b) Second-step-stopwith MV(1,0). (c) and (d)an unrestricted search path for MV(-1,-1) and (4,1) respectively.

Moreover, the average search points per block with NCDS < SCDS< CDS < DS < N3SS < 3SS < FS is observed. Among the video conferencing sequence, such as “Miss America”, “Sales”, and “Claire”, the proposed NCDS obtains at least 46% of speed improvement, even in vigorous motion content like “Football”, the speed up ratio can being achieved up to 42% and the worst is 18% on Tennis sequence. The trade off of the block distortion for faster speed is tabulate in Table IV to compare the difference of average MAE per pixel from FS. The result shows that the NCDS gives nearly the same MAE performance compared to CDS and even better compared to DS on those Videoconferencing sequences (around 0.02% to 0.1% of the quality degradation to CDS). For high motion content, the NCDS introduce a slight degradation (around 0.5% to 2.5% of the degradation). Furthermore, Fig. 4 plots the ASP and MAE per pixel frame-by-frame using the sequence “Miss America”. From the graph, the NCDS uses the smallest number of ASP and the MAE performance is reported very close CDS and also better compared to DS.

Average searching point ASP

	FS	3SS	4SS	N3SS	DS	CDS	SCDS	NCDS
Tennis	202.1	23.20	18.65	20.67	16.25	15.38	13.9	**13.2772**
Garden	202.1	23.24	18.80	21.38	16.84	15.09	14.87	**13.4562**
Football	202.1	23.06	16.69	17.65	13.67	10.96	8.24	**7.9022**
MissA	202.1	23.46	18.319	19.99	16.36	11.75	10.75	**8.7445**
Claire	202.1	23.22	15.924	16.19	12.4	8.92	5.38	**5.1753**
Sales	202.1	23.21	16.206	16.94	13.02	9.5	6.99	**6.1272**

TABLE III: THE AVERAGE NUMBER OF SEARCHING POINTS OF FS,3SS, 4SS, N3SS, DS, CDS, SCDS, AND NCDS OVER THE SIX SEQUENCES

Difference of average MAE per pixel from FS

	3SS	4SS	N3SS	DS	CDS	SCDS	NCDS
Tennis	1.0374	0.4383	0.488	0.2415	0.2935	0.3584	**0.4434**
Garden	0.9845	0.6502	0.1568	0.2337	0.1906	0.2056	**0.2342**
Football	0.2436	0.1683	0.1034	0.1452	0.1709	0.19	**0.204**
MissA	0.1169	0.1165	0.0253	0.1021	0.0352	0.0371	**0.0391**
Claire	0.0038	0.0035	0.001	0.0014	0.0029	0.0033	**0.0033**
Sales	0.0521	0.044	0.0081	0.0423	0.0094	0.01	**0.01**

TABLE IV: DIFFERENCES OF AVERAGE MAE PER PIXEL FROM FS OF FS,3SS, 4SS, N3SS, DS, CDS, SCDS, AND NCDS OVER THE SIX SEQUENCES

	NCDS over DS		**NCDS over CDS**	
	SIR (%)	MAE	SIR	MAE
Tennis	18.28058	3.442634	13.66332	2.554707
Garden	20.07626	0.005714	10.83708	0.498263
Football	42.18847	0.885594	27.92331	0.497677
MissA	46.54102	-2.71493	25.59519	0.168067
Claire	58.2802	0.17321	41.96858	0.028868
Sales	52.93539	-1.12371	35.47191	0.020874

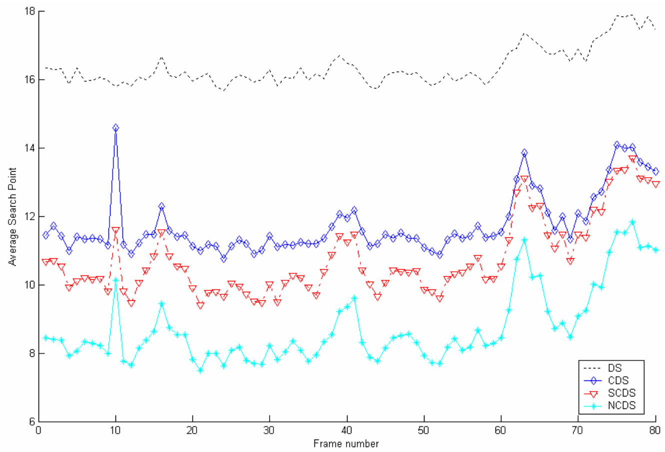
TABLE V: AVERAGE SPEED IMPROVEMENT RATIO AND AVERAGE MAE CHANGED PERCENTAGE OF THE NCDS OVER DS AND CDS

5. CONCLUSION

Based on the cross-center biased motion vector distribution of the real world video sequences, a novel new cross-diamond search algorithm (NCDS) is proposed. Simulation results showed that NCDS is the fastest algorithm among the testing BMA while providing similar prediction accuracy. It is especially suitable for videoconferencing sequences.

ACKNOWLEDGMENT

The work described in this paper was substantially supported by a grant from City University of Hong Kong, Hong Kong SAR, China. [Project No.7001385]



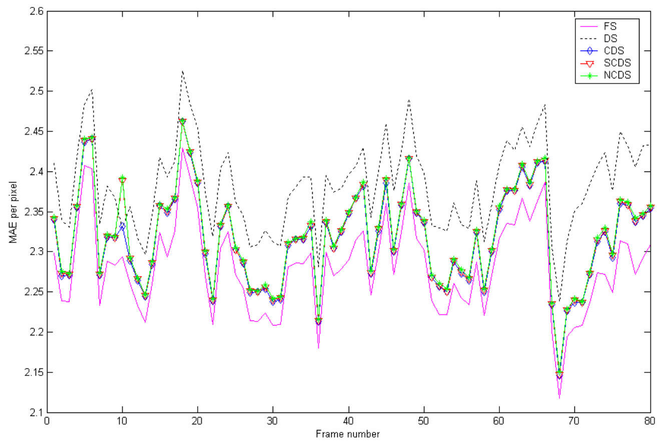


Fig.4. Performace comparison of FS,DS, CDS, SCDS and NCDS: ASP and MAE per pixel frame-by-frame using the sequence "Miss America"

REFERENCES

[1] T. Koga, K. Iinuma, A. hirano, Y. Iijima, and T. Ishiguro, "Motion compensated interframe coding for video conferencing", in Proc. Nat. Telecommun. Conf., New Orleans, L.A., Nov.-Dec. 1981, pp. G5.3.1-G5.3.5.

[2] R. Li, B. Zeng, and M. L. Liou, "A new three-step search algorithm for block motion estimation", IEEE Trans. Circuits Syst. Video Technol., vol. 4, pp. 438-443, Aug 1994.

[3] L. M. Po and W. C. Ma, "A novel four-step search algorithm for fast block motion estimation", IEEE Trans. Circuits Syst. Video Technol., vol. 6, pp. 313-317, Jun 1996.

[4] J. Y. Tham, S. Ranganath, M. Ranganath, and A. A. Kassim, "A novel unrestricted center-biased diamond search algorithm for block motion estimation", IEEE Trans, Circuits Syst. Video Technol., vol. 8, no. 4, pp. 369-377, Aug 1998.

[5] C. H. Cheung, and L. M. Po, "A Novel Cross-Diamond Search Algorithm for Fast Block Motion Estimation", IEEE Trans, Circuits Syst. Video Technol., vol. 12, no. 12, Dec 2002.

[6] C. H. Cheung and L. M. Po, "A novel small-cross-diamond search algorithm for fast video coding and videoconferencing applications", in Proc. IEEE ICIP, Sept. 2002.

[7] Yao Nie; Kai-Kuang Ma, "Adaptive rood pattern search for fast block-matching motion estimation", IEEE trans. Image Processing, vol. 11, pp.1442-1449, Dec 2002.