World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:1, No:6, 2007

New Enhanced Hexagon-Based Search Using Point-Oriented Inner Search for Fast Block Motion Estimation

Lai-Man Po, Chi-Wang Ting and Ka-Ho Ng

Open Science Index, Computer and Information Engineering Vol:1, No:6, 2007 publications.waset.org/3472/pdf

***Abstract*—**Recently, an enhanced hexagon-based search (EHS) algorithm was proposed to speedup the original hexagon-based search (HS) by exploiting the group-distortion information of some evaluated points. In this paper, a second version of the EHS is proposed with a new point-oriented inner search technique which can further speedup the HS in both large and small motion environments. Experimental results show that the enhanced hexagon-based search version-2 (EHS2) is faster than the HS up to 34% with negligible PSNR degradation.

***Keywords*—**Inner search, fast motion estimation, block-matching, hexagon search

I. Introduction

MOTION estimation (ME) is an integral part for most motion-compensated video coding standards. It is a process to remove the temporal redundancy between successive frames by referring current pixels to previously decoded frames. However, in order to find the best reference, the conventional exhaustive method, Full Search (FS), would introduce very high computational complexity. Over the last two decades, numerous fast motion estimation algorithms have been proposed to tackle this problem. Among them, the most popular class is the Block-Matching Algorithm (BMA) using a fixed set of search patterns which makes use of a very successful assumption – unimodal error surface. It means the matching error is monotonically decreased towards the global minimum.

In the 1980s, many fast BMAs were developed based on this assumption and some well-known examples are three-step search (3SS) [1], 2D-logarithmic search [2], and conjugate directional search [3]. They all can achieve substantial computational reduction but with a drawback of modest estimation accuracy degradation. This is mainly due to whole matching error surface is not monotonically decreasing towards the global minimum and there exists a lot of local minimums. These algorithms are, therefore, easily trapped into a local minimum with use of uniformly distributed search points in the early stage. In the early 1990s, experimental results [4, 5] showed that the block motion field of real world image sequence is usually gentle, smooth, and varies slowly. It results in a center-biased global minimum motion vector distribution instead of a uniform distribution. This implies that the chance to find the global minimum is much higher within the center 4x4 region of the search window. To make use of this characteristic, center-biased BMAs were then proposed using smaller searching patterns with search points much nearer the center to speedup the estimation time while the average prediction accuracy is also improved especially for the slow motion sequences. Well-known examples of this category are new three-step search (N3SS) [4], four-step search (4SS) [5], block-based gradient descent search (BBGDS) [6], diamond search (DS) [7], cross diamond search (CDS) [8], hexagon-based search (HS) [9], etc.

Recently, the enhanced hexagon-based search (EHS) algorithm [10] using a 6-side-based fast inner search was proposed to further speedup the original HS algorithm. In EHS, the searching algorithm is divided into two parts: (1) a low-resolution search which maximizes the coverage of the searching area and locates a small region where the motion vector would lies on; and (2) a high-resolution search or so-called inner search which finds the best motion vector inside the small region. In addition, the two major ideas of EHS algorithm raised are

(i) Further speedup can be achieved on saving search points for inner search; and

(ii) The inner search speedup can be made using a local unimodal error surface assumption (LUESA) by checking a portion of the inner search point errors.

It is because the final inner search minimum should have much higher chance with smallest sum of distortion for the search points surrounding it. This is statistically strong in a small region with the localized area around the global minimum. Based on the LUESA, the EHS only checks a portion of the inner search points that are nearer to the evaluated points with smaller distortions, which results in saving more than half of the inner checking points. However, we find the 6-side-based method is quite irregular by inspecting the distance between the inner points and evaluated points, and such a structure would result to lower prediction efficiency. In this paper, the 6-side-based fast inner search for EHS is first discussed and hence some principles of grouping the evaluated points are deduced, which favor our algorithm to reduce the necessary checking points. After that, an analysis of the motion vector

Manuscript received January 5, 2006. The work described in this paper was substantially supported by a grant from City University of Hong Kong, Hong Kong SAR, China. [Project No.7001699].

All the authors (L. M. Po, C. W. Ting and K. H. Ng) are with the Department of Electronic Engineering, City University of Hong Kong, Kowloon, Hong Kong SAR, China (e-mails: eelmpo@cityu.edu.hk, cwting@ee.cityu.edu.hk, khng@ee.cityu.edu.hk).

Open Science Index, Computer and Information Engineering Vol:1, No:6, 2007 publications.waset.org/3472/pdf

distribution in the inner search area is discussed. The simulation results and comparison of the enhanced and original algorithms are presented at the end with a short conclusion.

II. Enhanced Hexagonal Search

The efforts of 3SS, 4SS, DS, and HS are spent on reducing the number of searching points of the coarse search using different patterns such as square, diamond, and hexagon. However, exhaustive search is still often applied to the inner small region. This is because people commonly have a misconception that the long-range coarse search should be much heavier than the short-range inner search in terms of computational complexity. However, this is not always true. By the continuous improvement of search patterns and strategies, the number of search points required for the coarse search has been dramatically decreased. The EHS is an improved algorithm of the original HS using the 6-side-based fast inner search technique. The 6-side-based fast inner search is a group-oriented method. That means it first divides the evaluated coarse search points into a number of groups and accordingly assigns the inner points to different candidate groups. Fig. 1 (a) shows the groups for the large hexagonal search pattern (group 1 to group 6) and the eight inner points ('*a*' to '*h*'). This method separates the hexagon into six groups by the sides, and so each group consists of two evaluated points. A group distortion is calculated by summing up all the individual distortions in the group. By comparing the group distortions, this method only checks the points near to the minimum distortion group. The corresponding groups and confined regions are demonstrated in Fig. 1 (b). Two examples are shown in Fig. 1 (c) and (d), if group 4 has the minimum distortion, then point '*b*' and '*g*' will be checked, or if group 2 has the minimum, then point '*e*', '*a*', and '*f*' will be checked. While the above grouping method seems sensible, we find it cannot optimize the prediction result for each inner point. Instead of using such a group-oriented method, we believe a point-oriented method should yield better prediction accuracy and lower complexity. That means each inner point is considered as an individual. A dedicated group is formed by selecting tailor-made neighbors for an inner point. To find out the principles of determining these points, a few analyses are conducted with respect to distortion, distance, and correlation in the next section.

III. Point-Oriented Inner Search Strategy

A. Locally Unimodal Error Surface

The locally unimodal error surface assumption is an essential basis for inner search techniques. Within a localized region, if the block-matching error is smoothly increased in a monotonic way apart from the global minimum point, then the distortion of any other points in this region can easily be approximated to its neighbors by their separation distance. In other words, we can find out an approximate value of the distortion for a particular point without calculating its actual block distortion measure (BDM). Here we first define a metric for this purpose as

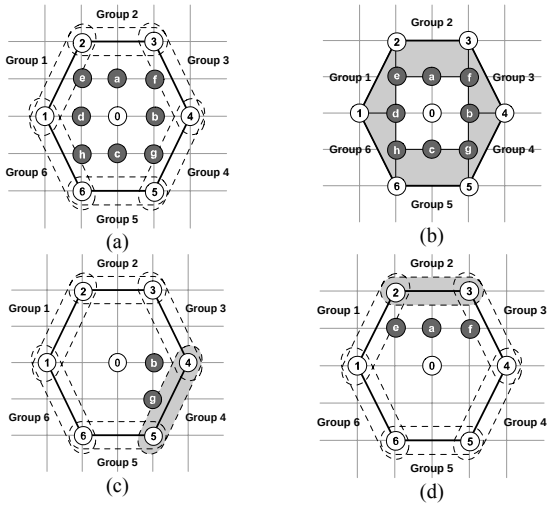


Fig. 1 Configurations and examples of the 6-side-based inner search; (a) groups of evaluated points; (b) defining shaded region for each group; (c) group 4 wins and checks point '*b*' and '*g*'; (d) group 2 wins and checks point '*a*', '*e*' and '*f*'

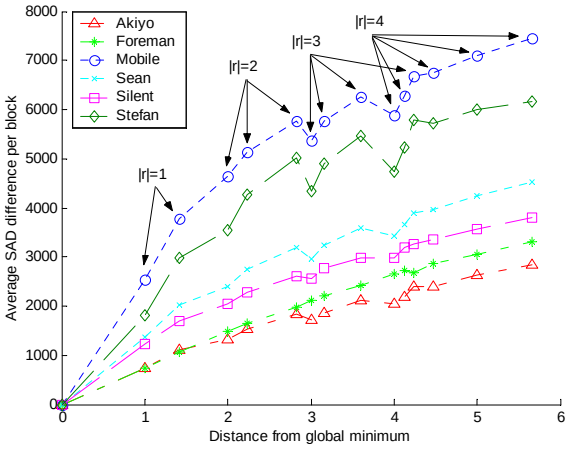


Fig. 2 Analysis of the relation between block SAD and distance

$$\overline{D}(d) = \sum_{i=1}^{N} \frac{D_i(d)}{N} = \sum_{i=1}^{N} \frac{SAD_i(d) - SAD(0)}{N} \quad (1)$$

For (1), $\overline{D}(d)$ is the mean value of the SAD (Sum of Absolute Difference) differences between the global minimum and any other points with a separation distance *d*. Accordingly, $D_i(d)$ is the SAD difference of the sample *i* from the global minimum, SAD*(d)* is the SAD of a point with distance *d* from the global minimum, and *N* is the total number of samples at the distance. The function is visualized and plotted against the distance in Fig. 2. Since the set of testing sequences has similar motion contents from class to class, only six representative sequences of CIF format are picked for the analysis. The first 100 frames, totally 39,600 blocks, are used to build the statistic. In Fig. 2, we can find that the SAD difference gently increases with the distance, but not monotonically on the whole. For the high motion sequences, "Mobile" and "Stefan", there are two obvious valleys at *d*=3 and *d*=4, and they are referred to the local minima of a motion search. The main reason for the drops is that the SAD is normally lower for a vertical or horizontal displacement rather than a diagonal displacement. However,

Open Science Index, Computer and Information Engineering Vol:1, No:6, 2007 publications.waset.org/3472/pdf

for the localized inner search region, we may just consider the fore part of the curve (i.e. $|r| \leq 2$), and the SAD difference will then become approximately linearly proportional to the distance. This characteristic is not only valid for the global minimum found by FS, but should also be valid for other minimum points found by suboptimal search method. Therefore, we may generally assume the distortion linearly increases in the targeted inner search area.

B. Correlation and Distance

The analyses reveal that the distance between two points not only affects the distortions, but also influences the correlation of them. An experiment was set up to verify the correlation between any two points with a separation distance *d*. Ten pairs of samples are randomly selected from each macroblock for a particular distance. As we only concern about the correlation within the inner search area, the samples are restricted in a radius |r|=2 region around the global minimum. Let $D''_i(d)$ be the SAD difference for the pair *i*, and $\overline{D''}(d)$ be the mean SAD difference for all the samples. Then the variance is defined as:

$$V(d) = \sum_{i=1}^{N} \frac{(D''_i(d) - \overline{D''}(d))^2}{N} \quad (2)$$

The variance *V(d)* is a measure of how spread out the distribution of the block difference from the mean is. Fig. 3 shows the variance increases linearly along with the distance, and has a sudden drop when *d*=3 due to the same reason mentioned in the last subsection. The higher the variance is, the lower the correlation exists between the pair. As the samples are randomly taken from the 4x4 region, the behavior shown by the figure can be generalized and applied to any position inside. It delivers an important message that the neighbors do not give a significant meaning to their central point when they are separated too far away.

C. Assumptions and Principles for Inner Search

To sum up the analysis results, a point-oriented inner search method requires two basic assumptions. For the localized inner search region,

1. There is only one minimum point, and
2. The distortion increases linearly away from the minimum.

Based on these assumptions, one's distortion can be approximated by the average of its neighbors with the following grouping principles:

1. Neighbors are as more as possible;
2. Neighbors are within shortest distance;
3. Distortion of each neighbor is normalized for calculating group distortion; and
4. Each group has the same size for comparing group distortion.

The first two principles simply emphasize the resulted precision and correlation of a group. The third principle states the varying intra-group weighting that the neighboring correlations decrease with distance, and so the distortion should be normalized by the distance before adding. The last principle requires the uniformity of each group so that inter-group comparisons are fair and under the same conditions. In order to measure the group size, a simple metric called mean internal

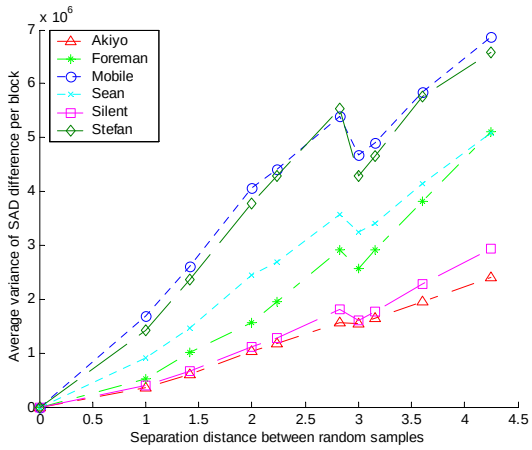


Fig. 3 Variance of block SAD difference for the random sample pairs taken from the 4x4 region around the global minimum

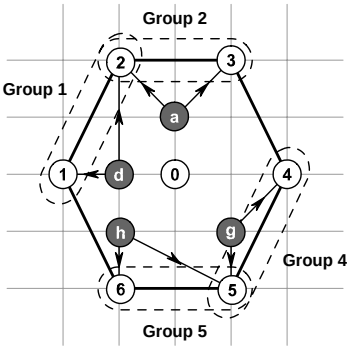


Fig. 4 Four values of mean internal distance are raised by 6-side-based fast inner search

distance (MID) is proposed as

$$MID = \sum_{i=1}^{N} \frac{d_i}{N} \quad (3)$$

For each group where the inner point is positioned in the center, the MID is defined as the mean of the distances d_i from the central point (inner point) to each neighbor (evaluated point), and *N* is the number of neighbors in the group.

Up to now, we have inspected the 6-side-based fast inner search and our statistical analyses. Two different philosophies for the fast inner search – group-oriented and point-oriented are brought out. Their major difference is that the former approach does not make a dedicated group for each inner point while the latter one does specifically assign one single group for each. Owing to this reason, some shortcomings of the 6-side-based method are observed. First, it has relatively large grouped regions, which contain up to three inner points and require more subsequent effort to converge. Besides, if the inner points are considered individually, we find its internal structure is quite irregular that there are four different values of MID as shown in Fig. 4.

IV. Point-Oriented Enhanced Hexagon-Based Search

Based on the principles proposed in the previous section, a more regular and efficient grouping method can be defined for the hexagon-based search pattern. The new grouping of the

Open Science Index, Computer and Information Engineering Vol:1, No:6, 2007 publications.waset.org/3472/pdf

evaluated points is based on minimizing the MID for each inner point, as shown in Fig. 5 (a). Since the hexagon shape is not entirely symmetric, we unavoidably have two different values of MID inside the hexagon. For fair comparisons between groups, they are further classified into two sets by the MID – Set-1: $\{a, c, e, f, g, h\}$ and Set-2: $\{b, d\}$. Within the same set, the inner points have the same MID to the others so that the evaluation can be made under equivalent correlation. These points are either surrounded by three or two nearest evaluated points, for instance, point '*a*' is grouped with points 0, 2, 3 and point '*b*' is grouped with point 0, 4. Similar to the 6-side-based inner search, a group distortion is used as a reference to predict the minimum location, but the evaluated distortions are normalized before adding up. The normalized group distortion (NGD) is expressed as follow.

$$NGD = \sum_{i=1}^{N} \frac{SAD_i}{d_i} = \sum_{i=1}^{N} \frac{SAD_i}{\sqrt{(x_i - x)^2 + (y_i - y)^2}} \qquad (4)$$

The NGD is the sum of all the distortions (SAD is used here) of the neighbors divided by the distance *d*. Accordingly, (x_i,y_i) and (x,y) are the coordinates of neighbor *i* and the inner point respectively, and *N* is the total number of grouped neighbors.

Within each set (set-1 and Set-2), the NGDs of the different groups are compared. Each set will find one inner point with the smallest NGD. Finally the 2 inner points will be searched. This is depicted by the example in Fig. 5(b). The two selected points have a much higher chance to be the final minimum point as compared with the other six points. This grouping has higher utilization of the neighboring correlations than the 6-side-based grouping and only requires 2 additional search points constantly. The overhead caused by the group distortion computations and comparisons is actually quite similar to that of the 6-side-based inner search, and should be negligible. With the use of the proposed point-oriented fast inner search, the hexagon-based search can achieve faster convergence, and so it is named enhanced hexagonal-based search using point-oriented inner search (EHS version -2). The algorithm is summarized in three steps below:

Step 1: Set the minimum distortion point to the center of the search area (0, 0)

Step 2: A minimum distortion point is found from the 7 checking points of the hexagon with the center at the previous minimum distortion point. If the new minimum distortion occurs at the center of the hexagon, go to Step 3; otherwise this step is repeated again.

Step 3: Compute the normalized group distortions of the hexagon and find out the minimum NGDs for Set-1 and Set-2. Based on the minimum NGDs, compute the distortions of the two additional search points and then identify the new minimum distortion point, which is the final motion vector.

Besides the new inner search, an early termination method for the inner search is also proposed. It is reasonable to assume if the current sub-optimal distortion is already small enough, it is not necessary to spend further effort to search another optimum which simply does not reflect any difference on the visual quality. Based on our study on the eight inner checking points, around 35% to 96% motion vectors are located in the position of point 0. This locally center-biased characteristic is especially obvious for the sequences containing large amount of static background, e.g. Akiyo, Sean, and Silent. This method terminates the inner search if the current minimum distortion (point 0) is smaller than a threshold. To maintain the prediction accuracy, a relatively low threshold value of 384 is selected for a MAE block distortion measure. The threshold is set to keep the picture quality unchanged, but further reduces the minimum number of inner checking points to zero.

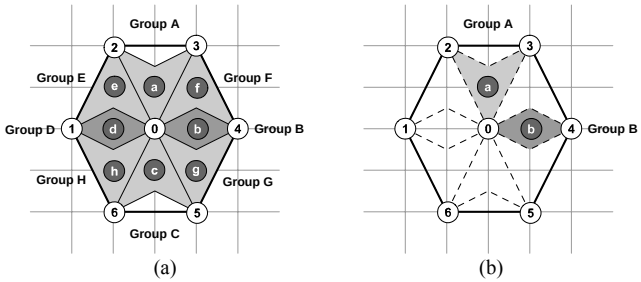


Fig. 5 (a) The new grouping for the hexagonal search pattern based on minimizing mean internal distance (MID). The eight inner points are separated into 2 sets by the MID value – Set-1: {a, c, e, f, g, h} and Set-2: {b, d}. (b) Point 'a' and 'b' are searched finally, because group A and group B have the smallest normalized group distortion in the corresponding set

V. Experimental Results

To demonstrate the performance of the proposed algorithms, the EHS, EHS2 (with fast inner search only) and EHS2+ (with fast inner search + early termination) are compared with the original HS. The simulation is performed on seven representative CIF sequences from three difference classes of MPEG-4; ClassA: Akiyo, Sean; ClassB: Foreman, News, Silent; ClassC: Mobile, Stefan, which contain various real world motions and objectively justify our method. The simulation settings are: 100 frames, 16x16 block size, ±16 search window, and MAE block distortion measure.

Experimental results are tabulated by two testing criteria – average PSNR per frame and average number of search point per block. To compare the visual quality, Table I shows that the PSNR changes of EHS2 and EHS2+ against HS are around –0.25dB (Mobile) to 0.36dB (Foreman), while that of EHS is around –0.42dB (Mobile) to 0.33 (Foreman). The maximum difference between EHS2 and EHS2+ is 0.008dB (News) that the early termination does not ruin the picture quality at all. The above data reveals that a negligible degradation may be caused by the fast inner search algorithms, while sometimes a slightly increase of PSNR may also be shown due to the original HS only checks a small cross (4 points) around the center. In general, our fast inner search is more accurate and has a smaller impact on picture quality as compared with the 6-side-based fast inner search.

To show the speed performance, we have two different comparisons. The number of search points reflects the absolute

TABLE I
AVERAGE PSNR PER FRAME FOR DIFFERENT ALGORITHMS ON SEVEN SEQUENCES

PSNR	Akiyo	Sean	Foreman	News	Silent	Mobile	Stefan
HS	42.556	38.858	32.340	37.038	35.714	23.633	23.717
EHS	42.314	38.736	32.666	36.997	35.747	23.217	23.724
EHS2	42.310	38.747	32.700	36.991	35.755	23.386	23.722
EHS2+	42.310	38.746	32.698	36.982	35.755	23.386	23.722

TABLE II
AVERAGE SEARCH POINTS PER BLOCK FOR DIFFERENT ALGORITHMS ON SEVEN SEQUENCES

Pts	Akiyo	Sean	Foreman	News	Silent	Mobile	Stefan
HS	10.346	10.569	12.659	10.588	10.958	10.704	13.758
EHS	9.373	9.612	11.346	9.499	9.751	9.459	12.353
EHS2	8.720	8.944	11.007	8.964	9.336	9.075	12.110
EHS2+	6.793	7.221	10.744	7.146	7.742	8.999	11.603

TABLE III
SPEED IMPROVEMENT RATE FOR THE THREE ALGORITHMS OVER THE ORIGINAL HEXAGON SEARCH.

SIR%	Akiyo	Sean	Foreman	News	Silent	Mobile	Stefan
EHS	9.41	9.06	10.37	10.29	11.02	11.63	10.21
EHS2	15.72	15.38	13.05	15.35	14.81	15.22	11.98
EHS2+	34.34	31.68	15.13	32.51	29.35	15.93	15.67

speed of different algorithms, and the speed improvement rate (SIR) reflects the speedup percentage relative to the original search. The SIR of method 1 over method 2 is defined by $SIR = (N_2-N_1)/N_2 \times 100\%$ where N_1 is the number of search point used by method 1 and N_2 is that by method 2. From Table II, the EHS2 saves around 1.6 search points constantly, which is a bit less than the theoretical result of 2 search points. It is because of the implementation that additional inner points will be checked if their corresponding groups are out of the frame boundary. For EHS2+, it saves around 1.7 to 3.6 search points. It is clearly that the early termination can efficiently reduce two more search points for most low to medium motion sequences, e.g. Akiyo, News, Silent, etc. Besides, the EHS saves around 1.0 to 1.4 points which are also less than the theoretical result due to the same reason. Finally, Table III shows the speed improvement rate over the original algorithms. The SIR of EHS2 and EHS2+ over HS are around 13% to 16% and 15% to 34% respectively, while that of EHS is around 9% to 11%, which is significantly smaller than that of the proposed methods. The experimental results justify the idea that the new grouping method is more suitable and efficient for the hexagonal search pattern. We also believe that the suggested grouping principles should be generally suitable for other kinds of search patterns such as diamond shape. The superior performance on both accuracy and speed is also demonstrated for the new fast inner search.

VI. CONCLUSION

The local unimodal error surface assumption facilitates the fast inner search as a complement of the low-resolution search to further reduce the overall checking points. By properly grouping the evaluated points, the group distortion can be used to predict the minimum of the inner points. In this paper, new grouping principles are proposed based on an analyzed statistic of the inner area. It is found that the mean internal distance (MID) is a very good measurement for the correlation between the inner points and coarse points. Therefore, the enhanced hexagon-based search version-2 (EHS2) algorithm is proposed using new point-oriented fast inner search techniques. It is shown that the EHS2+ significantly reduces 15% to 34% computations over the original HS. At the same time, a negligible PSNR loss is kept below 0.25dB. The new algorithm also out-performs the EHS on both accuracy and speed improvement rate.

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.,NO, L.A.,* Nov-Dec 1981, pp. G5.3.1-G.5.3.5.

[2] 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.

[3] R. Srinivasan and K. R. Rao, "Predictive coding based on efficient motion estimation," *IEEE Trans. Commun.*, vol. COM-33, pp. 888-896, Aug. 1985.

[4] 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.

[5] 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, June 1996.

[6] L. K. Liu and E. Feig, "A block-based gradient descent search algorithm for block motion estimation in video coding," *IEEE Trans. Circuits Syst. Video Technol*., vol. 6, no. 4, pp. 419-422, Aug 1996.

[7] 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.

[8] 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.

[9] C. Zhu, X. Lin, and L. P. Chau, "Hexagon-based search pattern for fast block motion estimation," *IEEE Trans. Circuits Syst. Video Technol*., vol. 12, pp.349-355, May 2002.

[10] C. Zhu, X. Lin, L. P. Chau, and L. M. Po, "Enhanced hexagonal search for fast block motion estimation," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 14, Issue 10, pp. 1210 - 1214, Oct. 2004.

Lai-Man Po (M'92) received the B.Sc. and Ph.D. degrees, both in electronic engineering, from the City University of Hong Kong (CityU HK) in 1988 and 1991, respectively. Since November 1991, he has been with the Department of Electronic Engineering, CityU HK, and is currently Associate Professor. His research interests are in vector quantization, motion estimation and image/video compression. Dr. Po was awarded First Prize (1988) in the Paper Contest for Students and Non-corporate Members organized by the Institute of Electronics and Radio Engineers of Hong Kong, and a Postgraduate Fellowship of the Sir Edward Youde Memorial Council (1988 to 1991).

Chi-Wang Ting received the B.Eng. and MPhil. degrees, both in computer engineering, from the City University of Hong Kong (CityU HK) in 2001 and 2005, respectively. He is currently a research assistant in department of electronic engineering, CityU HK. His research interests include fast motion estimation and fast mode decision algorithms for H.264/AVC.

Ka-Ho Ng received the B.Eng. degrees in information engineering from the City University of Hong Kong (CityU HK) in 2005, where he is currently pursing the MPhil degree. His research interests include fast motion estimation and fast mode decision algorithms for H.264/AVC.

Open Science Index, Computer and Information Engineering Vol:1, No:6, 2007 publications.waset.org/3472/pdf