http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 9 (7): 1390-1396, 2010
ISSN 1812-5638
© 2010 Asian Network for Scientific Information

A Novel Hybrid Coarse-to-Fine Digital Image Stabilization Algorithm

Guanghui Ren, Peng Li and Gangyi Wang
School of Electronics and Information Technology,
Harbin Institute of Technology, Harbin, Heilongjiang 150001, China

Abstract: The gray-scale projection algorithm can detect the motion vector of continuous frames rapidly in digital image stabilization system, but the interference of local moving objects may influence the estimation accuracy. In this study, we present a novel coarse-to-fine image stabilization approach which combined gray-scale projection algorithm with block matching algorithm by analyzing two classical motion estimation algorithms. The proposed method considers both accuracy and efficiency, since it makes good use of low computational complexity of gray-scale projection algorithm and high accuracy of block matching algorithm. The experimental simulations reveal that, expect for its low computational complexity, the proposed algorithm achieves high motion estimation accuracy. Moreover, it can efficiently avoid the interference caused by local moving objects.

Key words: Digital image stabilization, motion estimation, gray-scale projection, block matching, adaptive macro block, local motion

INTRODUCTION

Cameras usually suffer from instability due to unintentional shakes of carrier, when they are held by hand or mounted on an unstable platform. The irregular jitters of camera may lead to many adverse consequences, such as degradation of video sequence, vision fatigue, etc. Therefore, it is necessary to develop image stabilization techniques in order to eliminate the unwanted fluctuation from image sequences (Morimoto and Chellappa, 1998). There are three types of image stabilization techniques proposed in the literature, Mechanical Image Stabilization (MIS), Optical Image Stabilization (OIS) and Digital Image Stabilization (DIS) (Tico *et al*., 2006). Compared with MIS and OIS, DIS is a kind of method using digital image processing to detect the inter-frame motion vectors from an image sequence and obtain a stabilized sequence by compensation and it has the advantages of high accuracy, low computational complexity, low price and small size. Thus, DIS has been applied to both military and civil spheres more and more widely (Wang *et al*., 2009).

In a DIS system, the precise and efficient motion estimation algorithm plays an important role and it occupies the main calculational time and also determines the accuracy and processing velocity of the DIS system. Researchers have addressed various motion estimation algorithms so far, for example, block matching algorithms (BMA) (Luo and Celenk, 2008; Korah and Perinbam, 2006), bit-plane matching (BPM) (Ko *et al*., 1998), Representative Point Matching (RPM) (Hsu *et al*., 2007; Feng-Dong *et al*., 2009) and Gray-scale Projection Algorithm (GPA) (Zhang, *et al*., 2009). However, these approaches have their own limitations. BMA is commonly time-consuming. BPM improves the performance in terms of real-time processing, but it is difficult to choose appropriate bit plane for matching. Besides, RPM is too sensitive to the irregular circumstances, such as moving objects and intentional panning of the camera.

Gray-scale projection algorithm is a sort of feature-based matching algorithm in the sense of statistics. It utilizes the global gray-scale changes of image sequences to acquire the motion vectors between current frame and reference frame, by means of cross-correlation computation of row projection curves and column projection curves. GPA has an advantage over other algorithms in removing translational jitters and also has a good performance in both accuracy and computational complexity. Nevertheless, the traditional GPA requires high contrast. Meanwhile, it adopts row or column projection of entire image, which results in reduction of accuracy in the case of existing local moving objects.

In this study, a hybrid DIS algorithm is proposed by combining GPA and BMA to solve problems that the traditional GPA has. In the algorithm, we employ a coarse-to-fine process to estimate the inter-frame motion vectors. After pre-treating each frame collected by camera, traditional GPA is used to compute the coarse motion

Corresponding Author: Guanghui Ren, School of Electronics and Information Technology, Harbin Institute of Technology, Harbin, Heilongjiang 150001, China

vector of the current frame and reference frame and then the accurate motion vector is calculated by using adaptive BMA to guarantee rich gray-scale information in selected blocks, while preventing interference from local moving objects.

DIS SYSTEM AND MOTION ESTIMATION ALGORITHMS

Structure of DIS system: Typically, a DIS system is composed of three processing parts: the motion estimation unit, the motion decision unit and the motion compensation unit (Lin *et al*., 2009), as described in Fig. 1.

The motion estimation unit is responsible for the estimation of inter-frame global motion parameters that are forwarded to the motion decision unit. The motion decision unit is used to analyze the type of the global motion vector and then extract the unwanted jitters while preserving the deliberate panning motion of the camera. The motion compensation unit accomplishes the stabilization of the video sequence according to the unintentional motion vector generated by the motion decision unit. Among the three parts, the motion estimation unit is the key of the DIS system. So the motion estimation algorithm is supposed to achieve both high accuracy and real-time implementation. Generally, block matching algorithm and gray-scale projection algorithm are widely applied to a DIS system in the part of motion estimation.

Block matching algorithm: The basic idea of block matching algorithm is described as follows. The reference frame is divided into several nonoverlapping Macro Blocks (MB) with the size of M×N. Here, it is assumed that all pixels in each block only have the same translation motion. Suppose that the maximum searching ranges in vertical and horizontal directions are dx and dy, respectively. For each MB in reference frame, a corresponding search window of size (M+2×dx, N+2×dy) is confirmed in current frame. According to certain matching criteria and searching strategy, within the search window, we can find out the best matching block with the optimal Motion Vector (MV) which minimizes the difference between the reference block and the candidate block. The scheme of BMA motion estimation is shown in Fig. 2.

Gray-scale projection algorithm: The principle of gray-scale projection algorithm is to project the two-dimensional image into two independent one-dimensional curves. That is to say, we can acquire the row and the

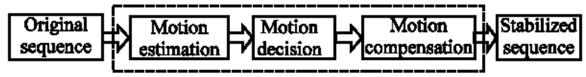

Fig. 1: Block diagram of DIS system

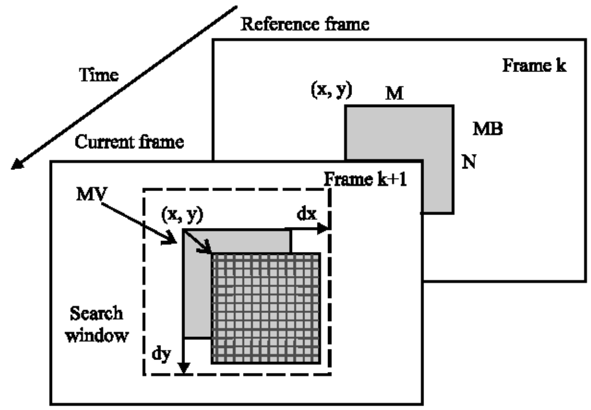

Fig. 2: Scheme of BMA motion estimation

column gray projection curves, which reflect the gray scale distribution of the image. It can be believed that two projections of row and column have little correlation, if there is no rotation motion or the angle of rotation is small. Based on calculating the cross-correlation of row and column projection curves separately, the single peak of the correlation curve is the motion displacement vector of the current frame relative to the reference frame. Gray-scale projection algorithm is mainly divided into two steps: image mapping and cross-correlation computing (Liu *et al*., 2009).

Image mapping: After pre-treating each frame of the input image sequence, its gray-scale information is mapped into row and column projection curves, that is, accumulate the gray scale of all pixels in every row or every column of the image, respectively. Since, the projection process of row is similar to that of column, we just take the column projection for example to illustrate the method of image mapping as follows:

$$G_k(j) = \sum_i G_k(i, j) \tag{1}$$

$$G_{mean} = \frac{\sum_j G_k(j)}{n_c} \tag{2}$$

$$G_{rk}(j) = G_k(j) - G_{mean} \tag{3}$$

where, G_k (j) is the gray-scale accumulation value of column j in the kth frame. $G_k(i, j)$ refers to the gray value of pixel in the position of (i, j) in the kth frame. In (2), G_{mean} is the mean value of the column projection. n_c is the number of column of the image. In Eq. 3, $G_{rk}(j)$ represents projection value of column j in the kth frame after mean normalization.

Calculation of cross-correlation: The cross-correlation of the projection curves of the current frame and the reference frame image is calculated, respectively. Then we can obtain two cross-correlation curves. According to the valley values of the two curves, the translation motion between the two frames can be determined. The column cross-correlation is defined as:

$$C(w) = \sum_{j=1}^{cl}[G_{k+1}(j+w-1) - G_k(m+j)]^2, (1 \le w \le 2m+1) \quad (4)$$

where $G_{k+1}(j)$ is the gray-scale projection of column j in the (k+1)th frame and $G_k(j)$ is gray-scale projection of column j in the kth frame. The length of the selected columns is indicated as cl. m is the one-side searching scope in relation to the reference frame. When $w = w_{min}$, the column projection value C(w) reaches the minimum value. Meanwhile, the corresponding horizontal motion vector δ_x between the two frames is given as:

$$\delta_x = (j+w_{min}-1)-(m+j) = w_{min}-m-1 \quad (5)$$

in Eq. 5, when δ_x is positive, it means that the current frame moves $|\delta_x|$ pixels to the right relative to the reference frame. When δ_x is negative, it indicates that the current frame moves $|\delta_x|$ pixels to the left relative to the reference frame. Similarly, the motion vector δ_y in vertical direction between the two frames can be expressed by following equation:

$$\delta_y = (m+j) - (j+w_{min}-1) = m+1-w_{min} \quad (6)$$

in Eq. 6, when δ_y is positive, it means that the current frame moves up $|\delta_y|$ pixels relative to the reference frame. When δ_y is negative, it indicates that the current frame moves down $|\delta_y|$ pixels relative to the reference frame.

THE PROPOSED HYBRID ALGORITHM

The conventional GPA aims to work out the projections of the total image. If the image contains mobile targets in the foreground or moving objects in the background, these conditions result in inter-frame differences, which reduce the accuracy of the projection algorithm. Some researchers adopted local sub-image projection method to increase the accuracy. The common strategy is to choose four fixed sub-images in the corner of the image and the local motion vector for each sub-image is derived from GPA or BMA. With these local motion vectors, the global motion vector of the two images can be determined (Erturk, 2003).

However, the fixed partition approach can not cover all situations. Because the choice of sub-images, which are selected in four corners of the image, is based on the assumption that the moving objects are usually in the center of the scene. In addition, the matching accuracy will decline, if the gray-scale projection curve of the selected sub-image does not verify obviously due to a small amount of gray-scale information. Hence, the new proposed algorithm combining GPA with BMA is adopted to balance the performance. It goes through a coarse-to-fine process to inherit the low computational complexity of GPA and the high accuracy of BMA. The strategy of choosing MBs should ensure abundant gray-scale information in the selected blocks as well as preventing interference from moving objects to improve the accuracy of motion estimation. The flow chart of the proposed algorithm is shown as Fig. 3.

Coarse motion estimation: Firstly, the traditional GPA is employed to project the global current frame and reference frame images by rows and columns. Next calculate the cross-correlation values of row and column projection curves, then find the valley amplitude values of the two cross-correlation curves with which the corresponding inter-frame motion vector can be assigned as the coarse motion vector.

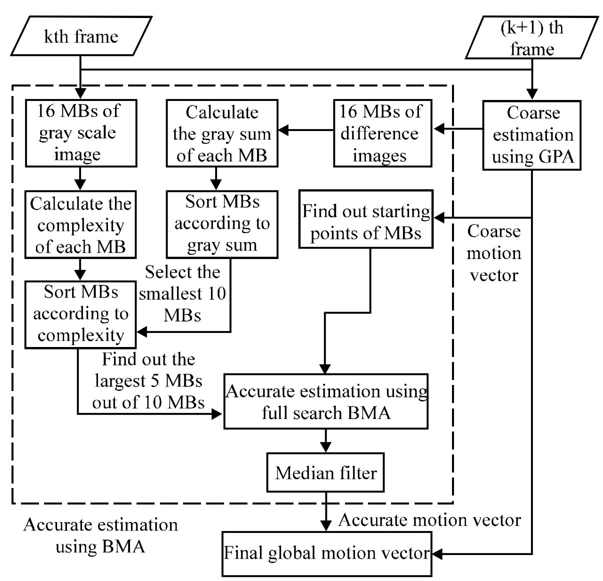

Fig. 3: Flow chart of the proposed algorithm

Accurate motion estimation: At first, taking both the assumption of motion consistency and enough gray-scale information in each block into consideration, the difference image of the reference image and the compensated image, which is gained by shifting the current image in corresponding reverse direction of the coarse motion vector, is divided into 16 MBs with size 64×64 pixels after cutting off the searching length p = 16 pixels around the difference image. After that, compute the gray-scale summation of all pixels in each MB and find out the smallest 10 MBs out of the 16 MBs with consideration of computational complexity and estimation accuracy, meanwhile write down the coordinates of the MBs' starting points.

Secondly, the reference image is also divided into 16 MBs with size 64×64 pixels according to the same method mentioned above. Then the complexity of feature of every block is evaluated, that is, pick out the MBs with a great amount of gray-scale changes as candidate blocks. The chief reason for this choice is that the candidate blocks which include different edge information or have obvious gray-scale changes inside them, will contribute greatly to raise the accuracy of matching. Otherwise, if the two candidate MBs are quite similar to each other, the motion vector may have larger matching error.

There are many kinds of methods to describe feature complexity of MBs, mean variance, gray-scale span of histogram and mean gradient, for instance. Based on an overall consideration of the complexity and accuracy of the algorithm, mean variance is chosen to describe the complexity of feature. Variance is used to measure the uniformity of gray-scale distribution. For a MB, the larger is the variance, the less uniform becomes the gray-scale distribution.

Assume that the entire image is divided into k MBs with size M×N, $f_k(i, j)$ is the gray-scale value at location (i, j) in the kth frame. First, the mean values of MBs are calculated in Eq. 7:

$$u_k = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N} f_k(i, j) \quad (7)$$

Then, calculate the mean variances of the corresponding 10 MBs selected before according to the formula Eq. 8:

$$\bar{\sigma}_k = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left[f_k(i, j) - u_k\right]^2 \quad (8)$$

Find out the largest 5 MBs out of the 10 candidate MBs in terms of mean variance and record the coordinates of the 5 MBs' starting points. And then, through shifting the starting point coordinates of the candidate 5 MBs to the opposite direction of the coarse motion estimation, we can obtain the starting point coordinates of the candidate 5 MBs for accurate motion estimation in current image. Next, 5 local motion vectors of the corresponding candidate MBs are obtained by using full search BMA. Then, the simple median filter is applied to the 5 local motion vectors to find the accurate estimation vector by calculating the mean value of the 3 middle motion vectors. Lastly, the final global motion vector is achieved by adding the coarse motion vector and the accurate motion vector together.

SIMULATION RESULTS AND DISCUSSION

Experimental results of the proposed algorithm are presented and compared with existing algorithms in this section. The real experimental video sequence with the CIF size at 25 frames per second is used. So the resolution of each frame is 352×288×8 bits. The video sequence mainly includes translational jitters as well as moving objects in the foreground. For gaining more obvious motion vector, we choose the first and fifth frame of the video sequence as the reference and current image, respectively. Figure 4a is the reference image while Fig. 4b is the current image. To achieve the coarse-to-fine motion estimation of the two frames, the row and column projections of the two entire frames are calculated firstly. Fig. 4c and e show the column and row projection curves of the reference image, respectively and the column and row projection curves of the current image are given as Fig. 4d and f. After projection, the cross-correlations of the row and column projection curves of the two images are calculated separately. Then, two cross-correlation curves can be obtained, as shown in Fig. 4g and h. According to the corresponding minimum value of two curves, the motion vector of coarse estimation between two frames is (2, 6).

Next, the accurate motion estimation is accomplished by using full search BMA. The size of each MB is 64×64 pixels and the searching scope is ±16. In Fig. 5a, the results of dividing the difference image into 16 MBs are displayed. After computation and sorting, the smallest 10 MBs with regards to gray-scale summation are shown in Fig. 5b and the largest 5 MBs out of 10 candidate MBs in terms of mean variance are given in Fig. 5c.

Table 1 shows estimation results of the proposed algorithm. In accurate estimation, the 5 local motion vectors of the corresponding 5 candidate MBs can be

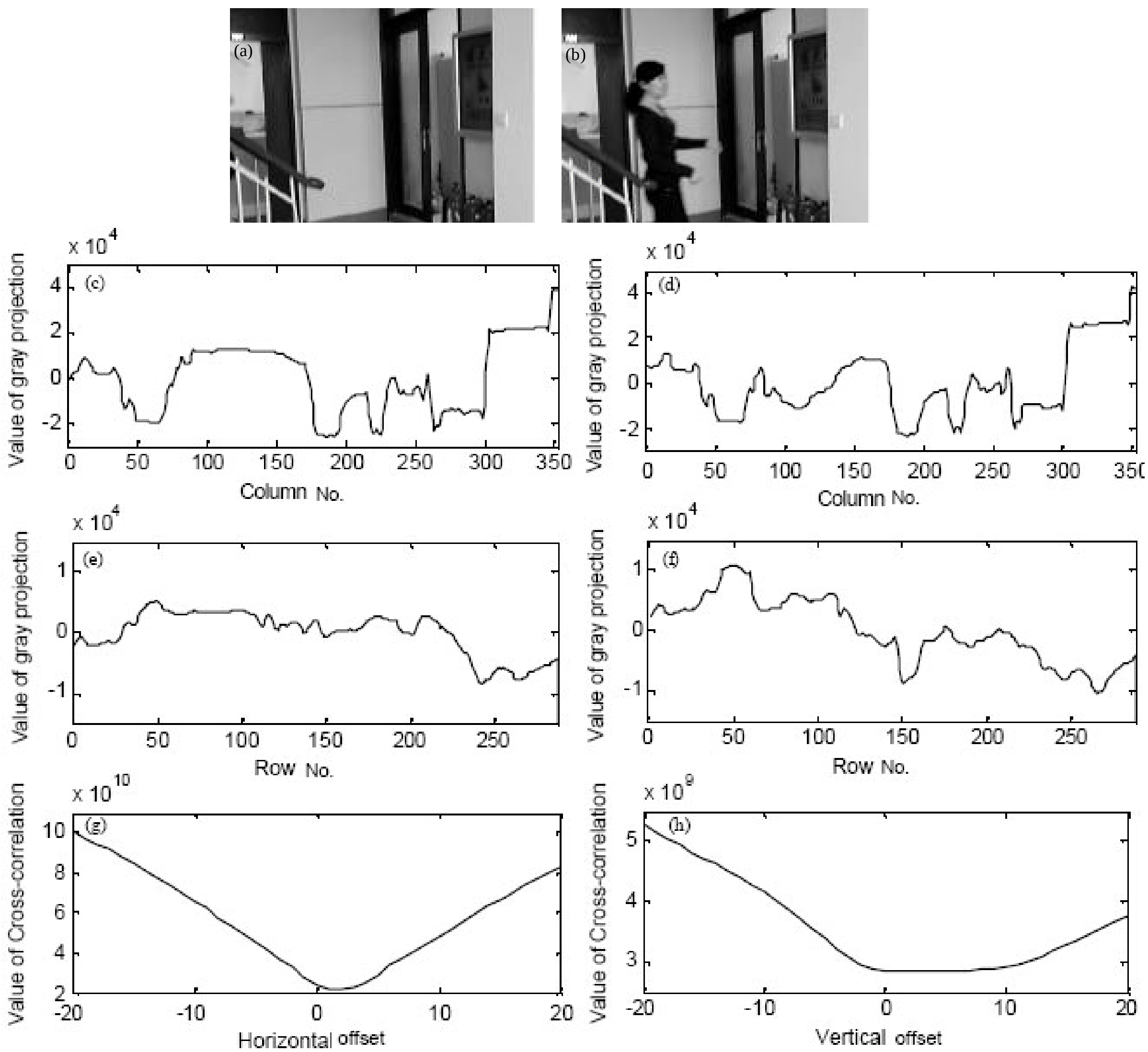

Fig. 4: Reference image and current image with their projection curves: (a) Reference frame, (b) current frame, (c) projection curve of column for reference frame, (d) projection curve of column for current frame, (e) projection curve of row for reference frame (f) projection curve of row for current frame, (g) correlation curve of column for two frame and (h) correlation curve of row for two frames

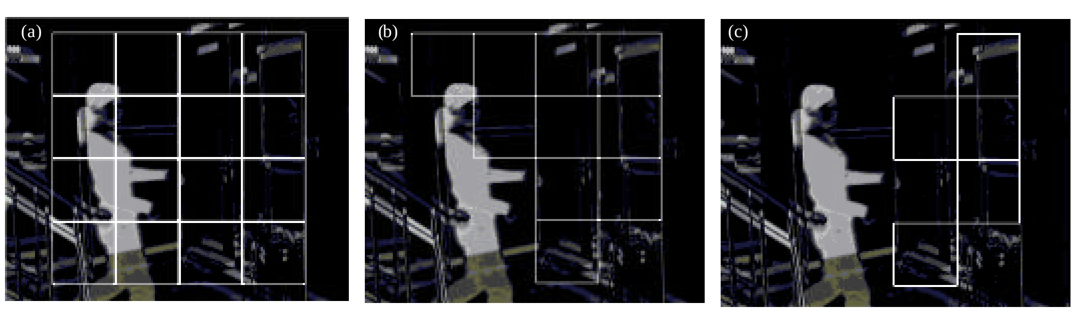

Fig. 5: Selection results of MBs for accurate estimation, (a) MBs of difference image (b) 10 MBs of the smallest gray-scale summation and (c) 5 MBs of the largest complexity

Table 1: Estimation results of the proposed algorithm

Searching initial position	Motion vectors of accurate estimation	Motion vectors of coarse estimation	Final motion vectors	PSNR (dB)
(16, 240)	(0, -5)	(2, 6)	(2, 1)	15.369
(80, 240)	(0, -8)	(2, 6)	(2, -2)	15.626
(80, 176)	(0, -9)	(2, 6)	(2, -3)	15.476
(142, 240)	(0, -8)	(2, 6)	(2, -2)	15.626
(206, 176)	(0, -8)	(2, 6)	(2, -2)	15.626

Table 2: Comparison of different motion estimation algorithms

Motion estimation algorithms	PSNR (dB)	Computation (8 bits)
Reference frame and current frame	13.842	NA
Full search block matching algorithm	15.836	3345408
Gray-scale projection algorithm	14.552	426624
Proposed algorithm	15.626	478080

achieved by full search BMA. After filtering the 5 local motion vectors with median filter and calculating the mean value of the three vectors sorted in the middle of the five, we can obtain the accurate motion vector (0, -8) and add the coarse motion vector and the accurate motion vector together. In this way, we can get the final global motion vector (2, -2).

Accuracy and efficiency are two measurement standards of DIS algorithms. Peak signal to noise ratio (PSNR) is usually used to evaluate the accuracy of DIS algorithms and effects of eliminating unintentional jitters and it can also be used to measure the similarity of two images. The higher is the PSNR between two images, the larger is the accuracy of the DIS system. When two images are completely coincident with each other, PSNR gets its maximum value. PSNR is commonly defined by the following equation:

$$PSNR = 10\log_{10}\left(\frac{x_{max}^2}{MSE}\right) \tag{9}$$

$$MSE = \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}(f_k(m,n) - f_{k-1}(m,n))^2 \tag{10}$$

where x_{max} represents the peak gray-scale value of the image (for gray images with a maximum intensity of 255). In (10), MSE is the mean square error between the current frame and reference frame and f_k (m, n) is the gray-scale value of kth frame at location (m,n) as well as f_{k-1} (m, n) is the gray-scale value of (k-1)th frame at location (m, n).

Table 2 is the comparison of different motion estimation algorithms between the two frames. Through the comparison results, we can see that the PSNRs after processing by all mentioned algorithms are larger than the one before processing, that is, all of the three algorithms have stabilization effect. Moreover, the PSNR of the novel proposed algorithm using coarse-to-fine estimation is observably larger than the PSNR between original images and it is improved a lot than traditional GPA, close to the PSNR of full search BMA. The experimental results indicate that the proposed algorithm in this paper has a higher accuracy. Furthermore, it has a low complexity according to the comparison of computation, much lower than that of full search BMA and but a little higher than that of traditional GPA. Therefore, the proposed hybrid algorithm sacrifices a part of computational complexity for high estimation accuracy. However, it is still suitable for real-time application to video sequence in the DIS system.

CONCLUSIONS

To improve the accuracy of motion estimation in DIS system, in this study, we propose a novel approach by combining gray-scale projection algorithm with block matching algorithm. The new algorithm is a coarse-to-fine process. Firstly, the algorithm utilizes gray-scale projection to obtain coarse motion vector between current frame and reference frame. Next, it uses block matching algorithm for accurate motion estimation. Then, the final motion vector of the two frames is obtained. With regard to the selected strategy of macro blocks used for block matching, the proposed algorithm considers the complexity of macro blocks. Blocks with abundant gray-scale information are selected for block matching. This method not only reduces the mismatching, but also avoids the interference caused by local moving objects. Experimental results demonstrate that the proposed algorithm can simultaneously provide both high accuracy and low computational complexity and can be used for handling video sequence in real-time DIS system.

REFERENCES

Erturk, S., 2003. Digital image stabilization with sub-image phase correlation based global motion estimation. IEEE Trans. Consum. Electron., 49: 1320-1325.

Feng-Dong, C., B.R. Hong and G.D. Liu, 2009. Planar displacement detection with point feature matching. Inform. Technol. J., 8: 383-387.

Hsu, S.C., S.F. Liang, K.W. Fan and C.T. Lin, 2007. A robust in-car digital image stabilization technique. IEEE Trans. Syst. Man Cybern. Part C Appl. Rev., 37: 234-247.

Ko, S.J., S.H. Lee and K.H. Lee, 1998. Digital image stabilizing algorithms based on bit-plane matching. IEEE Trans. Consum. Electron., 44: 617-622.

Korah, R. and J.R.P. Perinbam, 2006. A novel coarse-to-fine search motion estimator. Inform. Technol. J., 5: 1073-1077.

Lin, C.T., C.T. Hong and C.T. Yang, 2009. Real-time digital image stabilization system using modified proportional integrated controller. IEEE Trans. Circuits Syst. Video Technol., 19: 427-431.

Liu, Y., D. Fu, B. Li and Q. Xu, 2009. Real-time global motion vectors estimation based on phase correlation and gray projection algorithm. Proceedings of 2nd International Congress on Image and Signal Processing, Oct. 17-19, Tianjin, China, pp: 1-5.

Luo, Y. and M. Celenk, 2008. A hybrid block-matching approach to motion estimation with adaptive search area. Proceedings of 15th International Conference on Systems, Signals and Image Processing, Jun. 25-28, Bratislava, Slovakia, pp: 85-88.

Morimoto, C. and R. Chellappa, 1998. Evaluation of image stabilization algorithms. Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, May 12-15, Seattle, USA., pp: 2789-2792.

Tico, M., S. Alenius and M. Vehvilainen, 2006. Method of motion estimation for image stabilization. Proceedings of the 2006 IEEE International Conference on Acoustics, Speech and Signal Processing, May 14-19, France, pp: 277-280.

Wang, C., J.H. Kim, K.Y. Byun, J. Ni and S.J. Ko, 2009. Robust digital image stabilization using the Kalman filter. IEEE Trans. Consum. Electron., 55: 6-11.

Zhang, Y., F. Chen, J. Wang, W. Zhang and Q. Zhou, 2009. Implementation of gray projection image stabilization algorithm based on TMS320DM642. Proceedings of 5th International Joint Conference on INC, IMS and IDC, Aug. 25-27, Korea, pp: 1724-1726.