Highly precise and fast digital image stabilization technique based on the control grid interpolation

Jin-Hyung Kim, Ju-Hun Nam
Department of Electronics Engineering, Korea University
lifespring@korea.ac.kr

Jong-Nak Seok, Jeongwoo Han
DANAM SYSTEMS INC.

ABSTRACT

In this paper, we propose a highly precise and fast digital image stabilization technique based on the control grid interpolation. To obtain more stable video sequence than the one from other existing DIS techniques, the small instability should be removed in as small accuracy with sub-pixel. Experimental results show that the proposed digital image stabilizer gives considerable improvement in the sense of computational complexity and the performance of stabilizing compared to conventional DIS techniques.

KEYWORDS

Digital image stabilization, high precision, irregular condition, control grid interpolation.

INTRODUCTION

Cameras are usually hand-held or mounted on the mobile vehicles in various applications such as video coding [1], video surveillance [2]. In such situations, unintentional video vibration causes some degradation in the visual quality and captured images look unstable. The image stabilization (IS) technique is being used to overcome this problem. According to the literatures, there are three kinds of approaches, i.e., optical IS (OIS) [3], electronic IS (EIS) [4], and digital IS (DIS) [5].

OIS employs gyro-sensors or fluid prism to detect the undesirable motion of camera, and then compensate the image fluctuation, occurred by the motion, mechanically in opposite direction. It can handle large motion, but it is typically imprecise and tends to be bulky and heavy. EIS detects the camera motion by checking the changes between even and odd frames, and then compensates the unwanted motion by shifting the scanning area of CCD in opposite direction. EIS is more precise and physically compact compared to OIS. However it needs larger CCD plane than OIS or DIS, which can degrade the image resolution. In addition, it is usually less flexible than DIS.

Instead of mechanical equipments and CCD-based technologies, DIS uses digital image processing techniques to remove the unintentional motion while keeping the intentional one. In general, DIS consists of two stages, motion estimation and compensation. At the motion estimation, local motion vectors (LMVs) are obtained by using differential approaches such as optical flow [6], block-based [7], and feature-based [8] techniques. And then the global motion vector (GMV) is predicted based on the reliable LMVs. Finally the intentional motion vector is obtained by accumulating or smoothing the GMVs from previous frames and current frame [9]. In this smoothing process we may use FIR, IIR, Kalman filter, and

so on. At the compensation, the unintentional motion vector is calculated by subtracting the intentional motion from the GMV, and the unstable image can be compensated by the motion compensation.

Most of the researches have been focused on DIS because of its high flexibility, potentially large market, and low cost. Diverse algorithms are developed in many years. These conventional DIS algorithms have, however, one-pixel accuracy at most. DIS with accuracy of sub-pixel is required for some applications demanding high precision. Furthermore, conventional DIS techniques for the purpose of removing unwanted translational and rotational disturbance need much computational complexity.

PROPOSED DIGITAL IMGAGE STABILIZATION

In this section, we introduce the proposed DIS algorithm. Fig. 1 shows the flowchart of the proposed algorithm and the details are given in the following subsections.

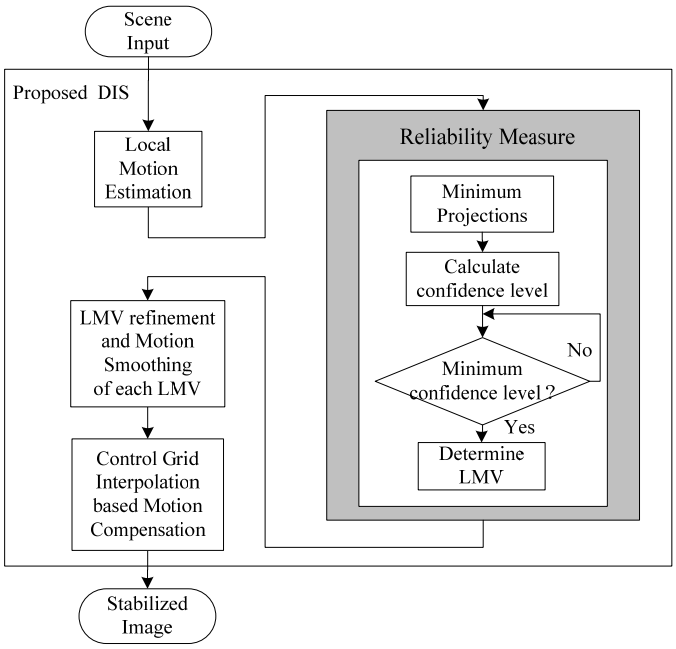


Fig. 1. Flowchart of the proposed DIS algorithm.

LMV Estimation with Sub-pixel Accuracy

The pre-stabilized image is divided into four sub-images, from each of which four macro blocks (MBs) are selected as shown in Fig. 2. We employ the block matching algorithm to estimate the LMV from each MB [10].

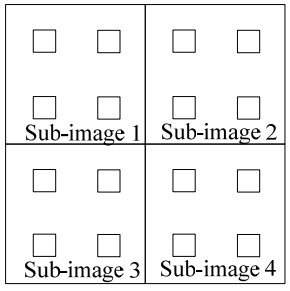


Fig. 2. MB selection for LMV estimation.

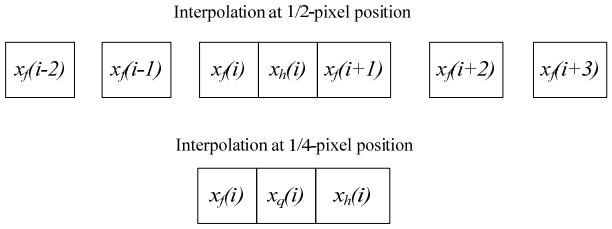


Fig. 3. Interpolation of sub-pixel positions.

Next, to obtain sub-pixel accuracy image is first up-sampled using bilinear interpolation expressed in (1),

$$x_h(i) = (x_f(i-2) - 5x_f(i-1) + 20x_f(i) + 20x_f(i+1) - 5x_f(i+2) + x_f(i+3))/32, \quad (1)$$

Where $x_h(i)$ is the interpolated value and $x_f(i)$ denotes the original pixel value at the position *i*. A graphical representation of the pixel locations used in the interpolation process is shown in Fig.3. The 1/4 pixel value $x_q(i)$ is then calculated using the full and 1/2 pixels as follows,

$$x_q(i) = (x_f(i) + x_h(i))/2, \quad (2)$$

As a result, for example 81 pixels are generated by interpolation from original 3×3 block. Therefore, we may obtain sub-pixel level of the matching precision.

For the fast estimation of subtle motion, we applied a logarithm search method [11] as shown in Fig. 4, where the circle, triangle, and square represent the positions of the full pixel, 1/2 pixel, and 1/4 pixel, respectively.

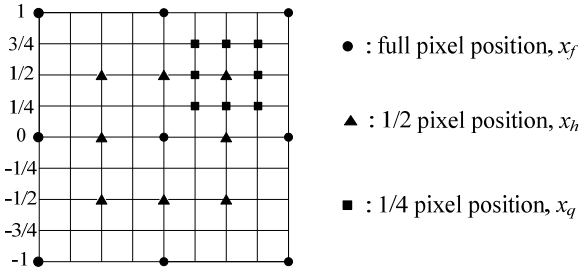


Fig. 4. Logarithm search method using interpolated image.

Reliability Measure

Smooth area	Repeated patterns	Complex texture

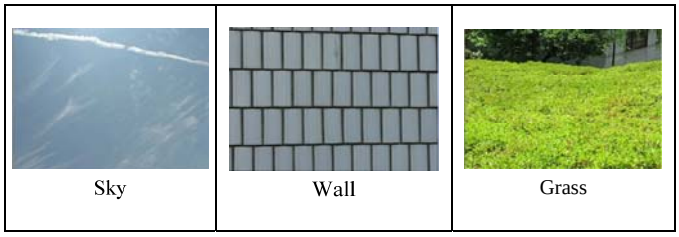


Fig. 5. Sample images with irregular condition.

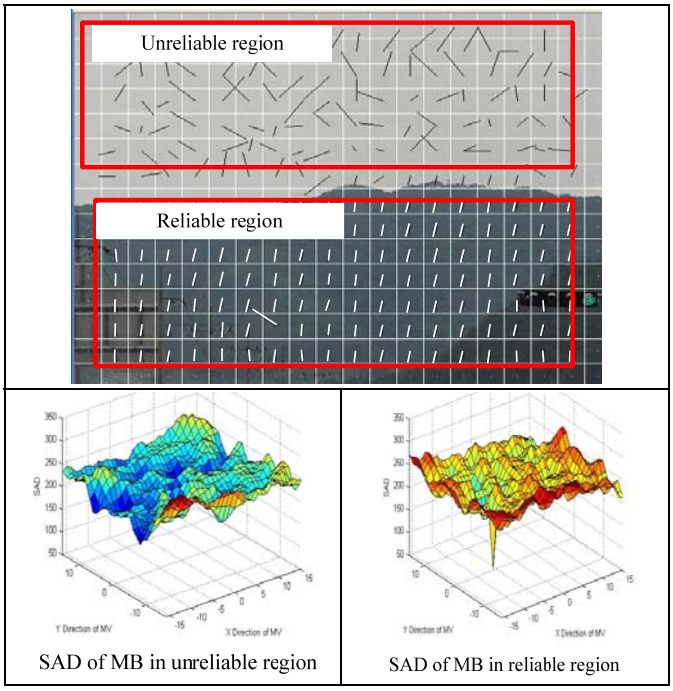


Fig. 6. LMVs and distribution of SAD values in reliable and unreliable region.

Fig. 5 shows the irregular conditions that generally exist in the smooth area, repeated patterns, and complicated texture such as the sky, wall, and grass, respectively. The irregular conditions in image can lead to some unreliable LMVs. In the unreliable region, it is difficult to find the minimum sum of absolute difference (SAD) and thus fails to estimate the LMV correctly as shown in Fig. 6. The lines which are indicated in the unreliable and reliable regions represent the LMVs of each MB. The LMVs in the unreliable region are wrong-estimated and the ones in the reliable region are correctly estimated.

To detect the irregular conditions, the inverse triangle method is employed [12]. Its main idea comes from the fact that the high reliable SAD curve for determining the LMV has one sharp and obvious peak as shown in Fig. 6. Therefore, among four LMVs in each sub-image, the LMV which has most steep SAD curve is selected as the reliable one.

LMV Refinement and Motion Smoothing

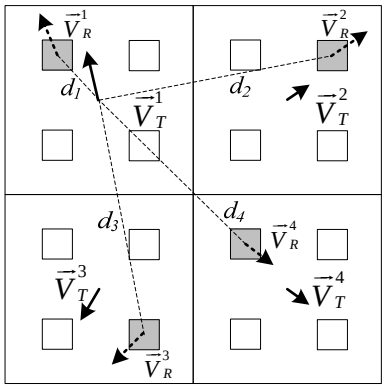


Fig. 7. LMV refinement.

The LMV refinement is illustrated in Fig. 7. In order to set the control grid points for motion compensation, the selected reliable LMVs are refined to the equivalent ones located at the center of each sub-image,

$$\vec{V}_T^x(n) = \sum_{i=1}^{4} w_i \cdot \vec{V}_R^i(n) \quad \text{for } x = \{1,2,3,4\}, \tag{3}$$

where $\vec{V}_R^x(n)$ and $\vec{V}_T^x(n)$ denote the selected and refined LMVs in each sub-image of the *n*th frame, respectively. The weighted factor w_i is defined as follows,

$$w_i = \frac{1}{3}\left(1 - \frac{d_i}{\sum_{r=1}^{4} d_r}\right) \quad \text{for } i = \{1,2,3,4\}, \tag{4}$$

where d_i is the distance between the position of the selected reliable LMVs and the one of the center position at each sub-image.

Each $\vec{V}_T^x(n)$ includes both the unwanted motion and the intentional one. To extract the undesirable motion, the intentional motion $\vec{V}_I^x(n)$ in each sub-image is firstly estimated using the IIR filter, which is processed as follows,

$$\vec{V}_I^x(n) = \alpha \cdot \vec{V}_I^x(n-1) + (1-\alpha) \cdot \vec{V}_T^x(n), \quad 0 \le \alpha \le 1, \tag{5}$$

where α is the damping factor. Finally, the unintentional LMV $\vec{V}^x(n)$ can be calculated as follows,

$$\vec{V}^x(n) = \vec{V}_T^x(n) - \vec{V}_I^x(n), \quad 0 \le \alpha \le 1, \tag{6}$$

These four unintentional LMVs are then used as the control grid to compensate the unintentional motion.

Motion Compensation based on Control-Grid Interpolation

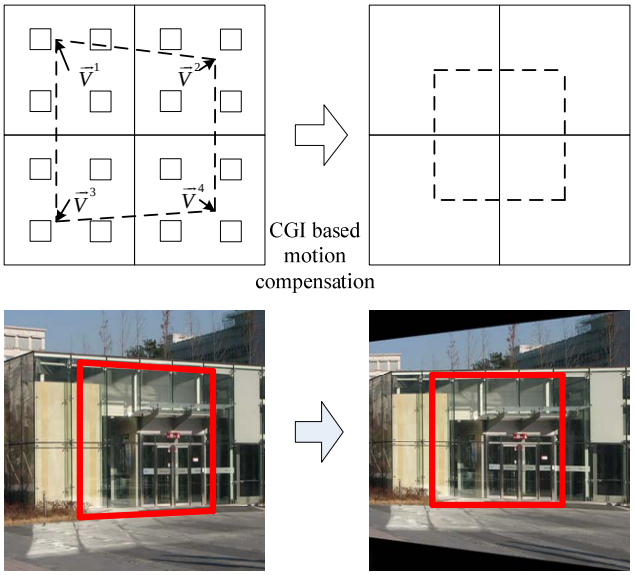


Fig. 8. CGI based motion compensation.

The four $\vec{V}^x$ are taken as the control-grid and then the control grid interpolation (CGI) is employed to obtain the compensated image, as shown in Fig. 8. The CGI is defined as follows,

$$\hat{I}(n,\vec{s}) = \tilde{I}\left(n,\ \vec{s} - \sum_{x=1}^{4} \xi^x(\vec{s}) \cdot \vec{V}^x\right), \qquad \sum_{x=1}^{4} \xi^x(\vec{s}) = 1 \quad \text{for } \vec{s} \in A_{i,j}, \tag{7}$$

where $\hat{I}_{(n,\vec{s})}$ and $\tilde{I}_{(n,\vec{s})}$ denote the pixel intensity at spatial position $\vec{s} = (i, j)^t$ in the block $A_{i,j}$ of the nth to-be-predicted frame and of the current frame, respectively. The weighted factor $\xi^x(\vec{s})$ is similarly calculated with (4) except that d_i is the distance between the position of each pixel and that of the center position at each sub-image.

EXPERIMENTAL RESULTS

In this section, we verify the stabilizing performance of the proposed DIS. In the simulation, we test the video sequeces captured by the OIS-installed camera mounted on the helicopter. The sample images are illustrated in Fig. 9. These video sequences have no intentional translation and rotation disturbance so that we can fairly compare the performance of the proposed system and the OIS-based one. The experimental results for the “Island” are taken as the example for performance assessment, and the other two samples, “Road” and “Golf course” also have similar results.

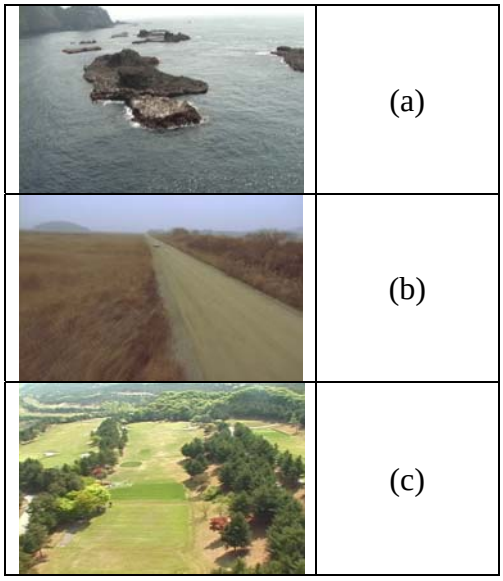
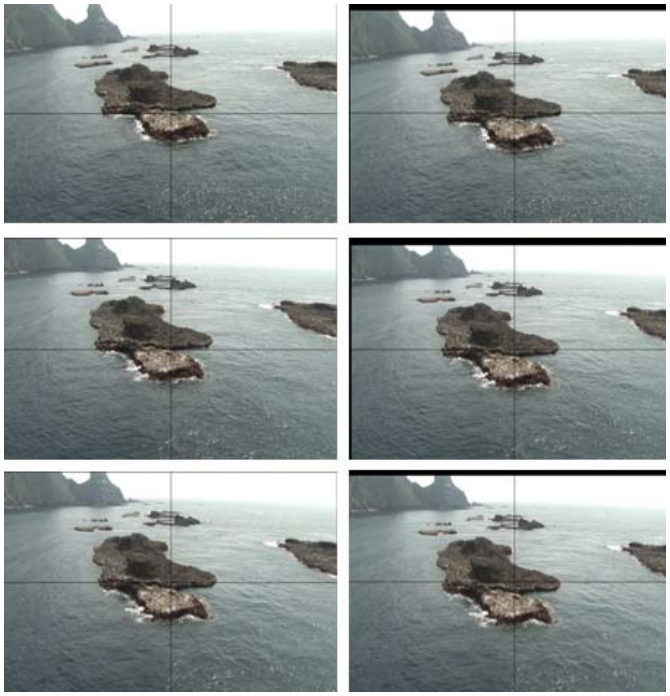


Fig. 9. Sample images.
(a) “Island.” (b) “Road.” (c) “Golf course.”

Fig. 10 (a) and (b) show the images stabilized by OIS and the proposed hybrid IS, respectively. However, we can not observe the difference between (a) and (b) clearly. To show the remaining jitters in the stabilized images, the central area of these two resultant images are enlarged and shown in Fig. 11, where the arrows represent the remaining motion. It is noticed that the proposed hybrid IS has better stablizing performance thant the OIS.



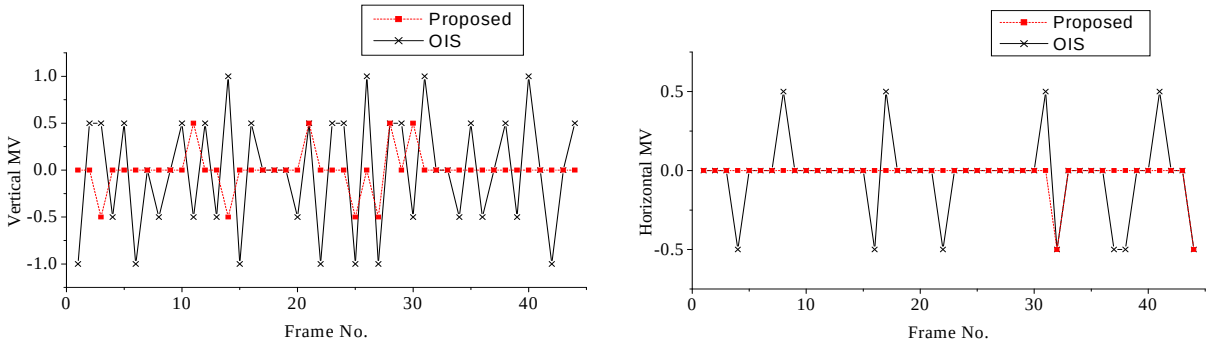
(a) (b)

Fig. 10. The image frames stabilized by
(a) OIS. (b) The proposed DIS.

(a) (b)

Fig. 11. The enlarged resultant image frames.
(a) OIS. (b) The proposed DIS.

Furthermore, Fig. 12 gives the remaining horizontal and vertical motions of the sample images stabilized by the OIS and the proposed IS. As shown in this figure, the remaining motions for the proposed IS is almost zero whereas the ones for the OIS is around or under one pixel. This shows that the remaining jitters in the images pre-stabilized by OIS are almost removed by the proposed DIS algorithm.



(a) (b)

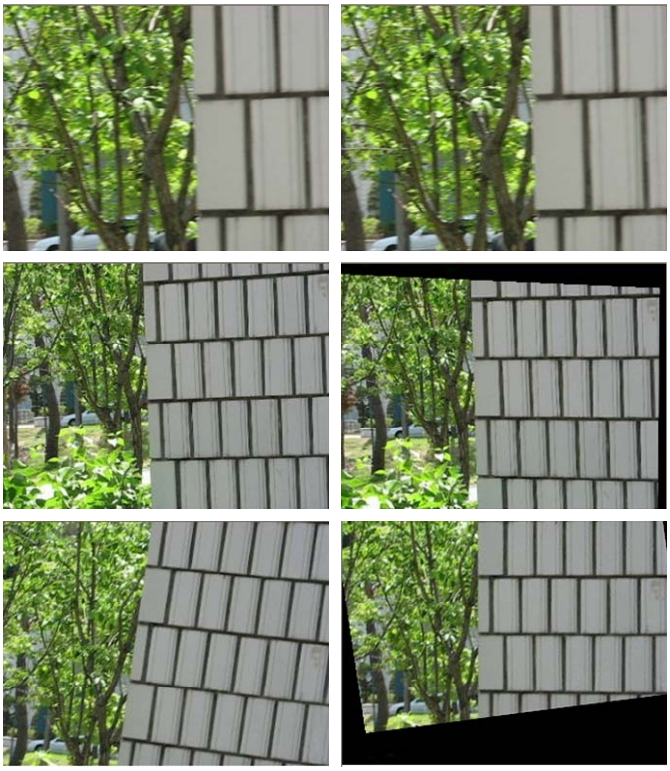
Fig. 12. The remaining motions of stabilized images.
(a) Horizontal motion. (b) Vertical motion.

Moreover, we compare the root mean square error (RMSE) of the proposed system and the OIS-based one [13] to show the performance of robustness against irregular conditions in image. Table I summarizes the experimental results under three irregular conditions, i.e., lack of features, low SNR, and repeated patterns. The given values in table I are the average RMSE for various video sequences shown in Fig. 5. It's observed that the proposed DIS has significant performance improvement in the sense of RMSE against some conventional IS.

TABLE I
RMSE VALUE OF TEST SEQUENCE (30 FRAMES)

Ill conditions of sample sequences	OIS	DIS [9]	Proposed
Sky (lack of features)	1.53	3.77	0.40
Grass (low SNR)	5.68	4.55	0.90
Wall (repeated patterns)	22.07	18.16	9.43

The proposed DIS can also remove the rotational disturbance in the images. To demonstrate this, the video sequence "Wall" with both small translation and rotation motions is further tested in our simulation. The results are illustrated in Fig. 13. Fig. 13 (a) shows the 1st frame (reference frame), 140th frame, and 256th frame of the sample sequence, and Fig. 13 (b) illustrates the stabilized image using the proposed algorithm. It is seen that the proposed CGI based compensation method can remove both the translational and rotational disturbances even though there are irregular conditions such as repeated pattern and complex area in the image.



(a) (b)

Fig. 13. Test results of sample sequence "wall".
(a) Original. (b) Stabilized images.

CONCLUSION

In this paper, we propose highly precise and fast digital image stabilizer. The proposed DIS algorithm is implemented in FPGA and DSP based system to achieve real time processing, where the FPGA is used to pre-process the input image and the DSP to execute the proposed DIS algorithm with sub-pixel accuracy.

The proposed DIS algorithm utilizes the block matching algorithm with sub-pixel accuracy to estimate

LMVs, then select the reliable LMVs to refine the selected LMVs using inverse triangle method, and finally compensates the unintentional motion with the CGI technique. Simulation results show that the proposed DIS achieves considerable performance improvement against some conventional IS.

REFERENCES

[1] A. Engelsberg and G. Schmidt, "A comparative review of digital image stabilising algorithms for mobile video communications," *IEEE Trans. Consumer Electron.,* vol. 45, no. 3, pp. 592-597, Aug. 1999.

[2] L. Marcenaro, G. Vernazza, and C. S. Regazzoni, "Image stabilization algorithms for video-surveillance applications," in *Proc. of IEEE International Conference on Image Processing,* vol. 1, Oct. 2001, pp.349-352.

[3] K. Sato, S. Ishizuka, A. Nikami, and M. Sato, "Control techniques for optical image stabilizing system," *IEEE Trans. Consumer Electron.*, vol. 39, no. 3, pp. 461-466, Aug. 1993.

[4] T. Kinugasa, N. Yamamoto, H. Komatsu, S. Takase, and T. Imaide, "Electronic image stabilizer for video camera use," *IEEE Trans. Consumer Electron.,* vol. 36, no. 3, pp. 520-525, Aug. 1990.

[5] S.-J. Ko, S.-H. Lee, and K.-H. Lee, "Digital image stabilizing algorithms based on bit-plane matching," *IEEE Trans. Consumer Electron.*, vol. 44, no. 3, pp. 617-622, Aug. 1998.

[6] J.-Y. Chang, W.-F. Hu, M.-H. Cheng, and B.-S. Chang, "Digital image translational and rotational motion stabilization using optical flow technique," *IEEE Trans. Consumer Electron.,* vol. 48, no. 1, pp. 108-115, Feb. 2002.

[7] F. Vella, A. Castorina, M. Mancuso, and G. Messina, "Digital image stabilization by adaptive block motion vectors filtering," *IEEE Trans. Consumer Electron.,* vol. 48, no. 3, pp. 796-800, Aug. 2002.

[8] Y.-M. Liang, H.-R. Tyan, S.-L. Chang, H.-Y. M. Liao, and S.-W. Chen, "Video stabilization for a camcorder mounted on a moving vehicle," *IEEE Trans. Vehicular Technology*, vol. 53, no. 6, pp. 1636-1648, Nov. 2004.

[9] S.-H. Yang and F.-M. Jheng, "An adaptive image stabilization technique," in *Proc. IEEE International Conf. on Systems, Man, and Cybernetics,* vol. 3, Oct. 2006, pp. 1968–1973.

[10] Y.-W Huang, S.-Y. Chien, B.-Y. Hsieh, and L.-G. Chen, "Global elimination algorithm and architecture design for fast block matching motion estimation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 14, no. 6, pp. 898-907, Jun. 2004.

[11] X. Jing and L.-P. Chau, "An efficient three-step search algorithm for block motion estimation," *IEEE Trans. Multimedia*, vol. 6, pp. 435–438, Jun. 2004.

[12] S.-C. Hsu, S.-F. Liang, and C.-T. Lin, "A robust digital image stabilization technique based on inverse triangle method and background detection," *IEEE Trans. Consumer Electron.*, vol. 51, pp. 335–345, May 2005.

[13] A. A. Yeni and S. Erturk, "Fast digital image stabilization using one bit transform based sub-image motion estimation," *IEEE Trans. Consumer Electron.*, vol. 51, pp. 917–921, Aug. 2005.