DIGITAL IMAGE TRANSLATIONAL AND ROTATIONAL MOTION STABILIZATION USING OPTICAL FLOW TECHNIQUE

Jyh-Yeong Chang, Wen-Feng Hu, Mu-Huo Cheng and Bo-Sen Chang
Department of Electrical and Control Engineering
National Chiao Tung University
Taiwan 300, R.O.C.

Abstract– **This paper proposes a new digital image stabilization (DIS) systems based on optical flow technique. Unlike previous DIS systems developed mainly for removing the translational motion disturbance, the proposed system removes not only the translational but also the rotational motion disturbances. A computational scheme that facilitates the local motion vector field to estimate the global disturbing translational and rotational motions is developed. First, the optical flow technique is used to estimate the local motion vector field of the image, yielding the velocity of each pixel in the current image frame. Then, the global translational and rotational motion parameters are determined in terms of the least squares estimation. Finally, these motion vectors are used to generate the counterbalance signals for removing the disturbance motion. Owing to the additional ability of the rotational motion removal, the new DIS system suppresses the undesirable translational and rotational disturbances effectively and thus enhances the DIS performance significantly.**

Keywords: **Digital image stabilization, Motion estimation, Optical flow.**

I. INTRODUCTION

Motion in video images is caused by either the object motion or the camera movement. A digital image stabilization (DIS) system aims to produce a compensated video sequence so that the image motion due to the camera's undesirable shake or jiggles can be removed [1]-[10]. In practice, the recent consumer electronic products, such as digital cameras, camcorders, CCD sensing arrays, and next-generation mobile phone with visual display, etc., would be very needy for an image stabilization loop to remove this inevitable and undesirable fluctuation motion during the image capturing process. As a consequence, it would be worthwhile to have a digital image sequence stabilization scheme that can further stabilize the image sequence for improving the subjective quality of the video sequence obtained. Moreover, an image stabilization algorithm is reported to be beneficial to the coding efficiency of video signals [10].

An electronic image stabilization system for consumer's video cameras has been developed by Oshima *et al.* [1]. Oshima's electronic image stabilization system uses a gyro sensor to measure the angular velocity of undesirable camera motion and then compensates for the image fluctuation by rotating the lens unit in the opposite pitching and yawing directions of the angular velocity. Uomori *et al.* [2], [3] have developed a DIS system for consumer's video camera also. Uomori's system contains a correlation-based matching technique to determine the motion vector using an adaptive motion computation scheme. The motion vector detector in the Uomori's systems is implemented with a chip of small circuit, unlike the mechanical motion detection/correction techniques using gyro sensors [1] or fluid prism [4].

Recent image stabilizer systems in the literature demonstrate the trend of detecting and compensating residue fluctuation motion via image processing and computer vision techniques [5]-[10]. In this trend, the block matching algorithm, a fast searching and matching algorithm commonly used in MPEG coding [12, pp. 101-108], is widely exploited to realize the motion detection of a DIS system [7]-[10]. For example, Paik *et al.* developed an image stabilization system [7] by block matching based scanning area selection. This approach is further improved by exploiting bit-plane matching [8]. The blocking matching approach, however, can only estimate the translational movement. Hence, this approach produced poor performance when the image fluctuation contains both translational and

This research was supported in part by National Science Council, Taiwan, R.O.C. under Grant NSC 89-2218-E-009-044.

Contributed Paper
Manuscript received January 2, 2002

rotational motions.

In this paper, we present a new approach for estimating both the translational and the rotational motions. We apply the optical flow technique to obtain the local velocity vector of each pixel in the current image frame. These velocity vectors are then used to determine the global translational and rotational motions in terms of motion dynamics model and least squares estimation. The resulting system has been tested by computer simulations and proved effective in compensating the translational as well as the rotational motion disturbances.

The rest of the paper is structured as follows. In Sec. II, we employ the zero-order optical flow [12], [13] to obtain a moving flow field of an image sequence. By the use of the least squares estimation, we derive a new scheme that can determine the angular frequency of the rotational motion from the moving flow field. With the estimated angular frequency, a searching procedure is then developed to find the rotational center of the rotational disturbance. The velocity vector at the rotational center determines the image's translational motion. According to the motion parameters estimated, compensated signals are generated to counterbalance the undesirable residual motion in the image sequence. Sec. III contains two simulations to verify the feasibility of the new DIS system. Finally, several concluding remarks are given.

II. A DIGITAL IMAGE STABILIZER USING OPTICAL FLOW

The new image stabilization system will first estimate the global motion between consecutive frames in a video sequence and then compensate for the giggle motion detected. To this end, the system consists of a motion vector detection module which detects input image movement and extracts the image fluctuation component caused from image capturing fluctuation. We then generate a compensated video signal so that the image's undesirable shake or jiggle can be counterbalanced. The configuration of the DIS system is shown in Fig. 1 and its details are illustrated in the following .

A. The Optical Flow Algorithm

This sub-section briefly highlights the optical flow algorithm used in this paper. Denote the image brightness at the point (x, y) in the image plane at time t by $E(x,y,t)$. In notation, let $u = \frac{dx}{dt}$, $v = \frac{dy}{dt}$, $E_x = \frac{\partial E}{\partial x}$, $E_y = \frac{\partial E}{\partial y}$, and $E_t = \frac{\partial E}{\partial t}$. The zero-order optical flow computation minimizes a weighted sum of errors of the pixel-to-pixel variation in the velocity field. Hence the total error to be minimized is

$$\varepsilon^2 = \iint (\alpha^2 \varepsilon_c^2 + \varepsilon_b^2) dx dy, \tag{1}$$

where

$$\varepsilon_b = E_x u + E_y v + E_t \tag{2}$$

denotes the sum of the errors for the rate of change of image brightness;

$$\varepsilon_c^2 = \left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial v}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial y}\right)^2 \tag{3}$$

represents the measure of the departure from smoothness in the velocity flow, and α^2 is a weighting factor. Using the calculus of variation and an approximation to the Laplacian operator, we obtain

$$\begin{aligned}(\alpha^2 + E_x^2 + E_y^2)u \\ = (\alpha^2 + E_y^2)\bar{u} - E_x E_y \bar{v} - E_x E_t,\end{aligned} \tag{4}$$

and

$$\begin{aligned}(\alpha^2 + E_x^2 + E_y^2)v \\ = (\alpha^2 + E_x^2)\bar{v} - E_x E_y \bar{u} - E_y E_t.\end{aligned} \tag{5}$$

where $\bar{u}$ and $\bar{v}$ are the local average of the velocities. In order to reduce the computational complexity, the Gauss-Seidel method is commonly applied to solve the above equations and yields the following iterative algorithm

$$\begin{aligned}u^{n+1} = \bar{u}^n \\ - E_x \left[E_x \bar{u}^n + E_y \bar{v}^n + E_t\right] / (\alpha^2 + E_x^2 + E_y^2),\end{aligned} \tag{6}$$

and

$$\begin{aligned}v^{n+1} = \bar{v}^n \\ - E_y \left[E_x \bar{u}^n + E_y \bar{v}^n + E_t\right] / (\alpha^2 + E_x^2 + E_y^2).\end{aligned} \tag{7}$$

The output of optical flow estimation method [12]-[14] is the velocity field $V(x,y) = (u(x,y), v(x,y))$ of each pixel at (x, y). Note that these estimated velocities may vary pixel to pixel because optical flow technique is rather noise sensitive. A new motion extraction algorithm which can extract the correct motion vectors from the local varying motion vectors is derived below.

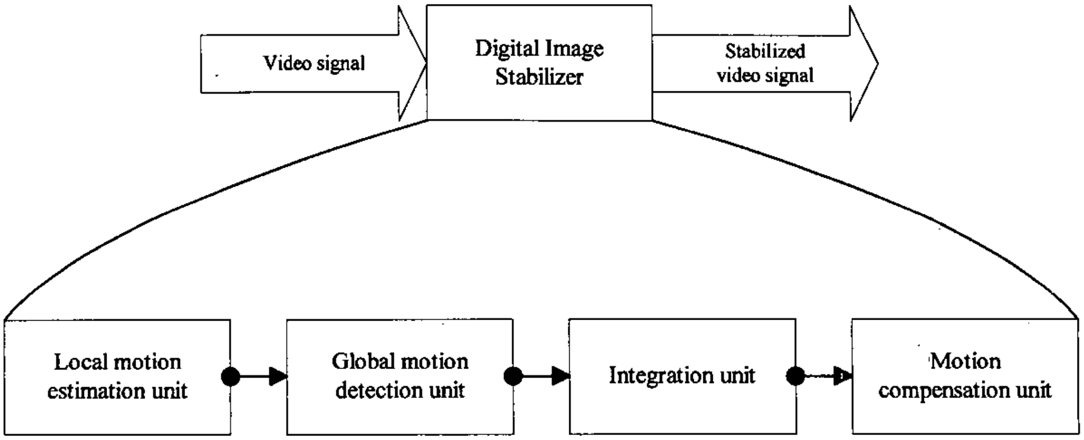


Fig. 1. Basic structure of the DIS system.

B. The Rotational Angular Frequency Extraction

After obtaining the local motion field of the image, we are now in a position to estimate this image fluctuation motion parameters. Since the sampling interval between the image frame sequence is normally short, we can assume that the fluctuation motion in each image frame can be approximated by a rotational motion about a rotational center, combined with a translational motion of the whole image plane. This assumption enables us to derive a new scheme that can detect the rotational angular frequency, rotational center, and translational velocities of each image frame. First we consider the case that the image frame sequence is *purely rotated* about a particular rotational center (x_0, y_0) by an angular velocity ω. For an arbitrary point A, located at (x, y), its rotational velocity vector $V_r = (u_r(x,y), v_r(x,y))$ at (x, y) of the rotational motion about (x_0, y_0), as shown in Fig. 2., can be decomposed into

$$\begin{aligned} u_r(x,y) &= -|V_r|\sin\theta \\ &= -\omega r(\sin\theta) \\ &= -\omega(y-y_0), \end{aligned} \tag{8}$$

and

$$\begin{aligned} v_r(x,y) &= |V_r|\cos\theta \\ &= \omega r(\cos\theta) \\ &= \omega(x-x_0). \end{aligned} \tag{9}$$

Then,

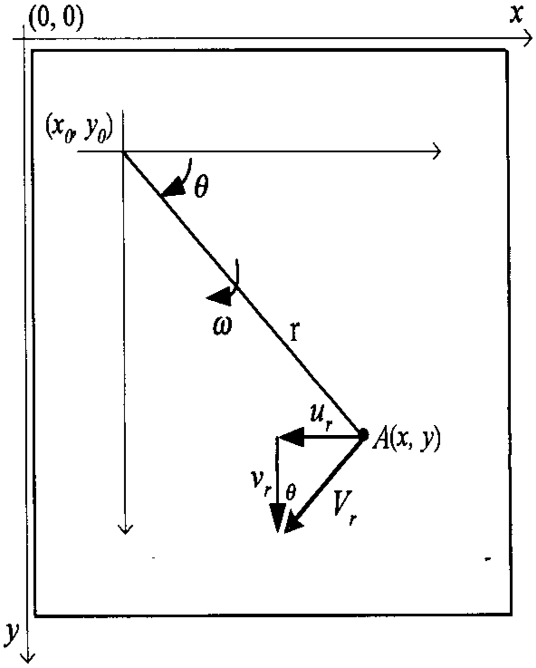


Fig. 2. The analysis of rotational motion at arbitrary position (x, y) in the image plane.

$$(u_r, v_r) = (-\omega(y-y_0), \omega(x-x_0)). \tag{10}$$

Now, let us address the general case of containing *both translational and rotational* motion. Let the coordinate frame at the rotational center have a translational motion, u' and v', in the horizontal and vertical direction, respectively. Then the velocity $(u(x,y), v(x,y))$ at any point (x, y), including translational and rotational angular frequency ω as well, will be given by

$$u(x,y) = u' - \omega(y - y_0), \tag{11}$$

and

$$v(x,y) = v' + \omega(x - x_0). \tag{12}$$

These two equations are best described by Figs. 3(a) and 3(b), which include the plots in the $x - v$ plane and $y - u$ plane, respectively. As evident from these two equations, each equation specifies one line in the figure, respectively. These two lines respectively have ω and $-\omega$ as their slopes and v' and u' as their intercepts. Since the motion vector estimation computation is inevitably corrupted by noise, the obtained velocity vectors of pixels should distribute in scatter over these two lines, as shown in Figs. 3(a) and 3(b). For any two points (x_{k1}, y_{k1}) and (x_{k2}, y_{k2}) in the image frame, they should satisfy Eqs. (11) and (12) and their velocity differences $\Delta u = u(x_{k1}, y_{k1}) - u(x_{k2}, y_{k2})$ and $\Delta v = v(x_{k1}, y_{k1}) - v(x_{k2}, y_{k2})$ take the following forms

$$\Delta u = -\omega \Delta y \tag{13}$$

and

$$\Delta v = \omega \Delta x, \tag{14}$$

where $\Delta x = x_{k1} - x_{k2}$ and $\Delta y = y_{k1} - y_{k2}$. Rewrite these two equations as

$$\begin{bmatrix} \Delta u \\ \Delta v \end{bmatrix} = \begin{bmatrix} -\Delta y \\ \Delta x \end{bmatrix} \omega. \tag{15}$$

Any pair of pixels should satisfy Eq. (15), which is in a form of $A\omega = b$. A least squares solution is usually adopted to minimize the discrepancy from Eq. (15) for the pixel pairs in the image. In this way, the least square estimate of ω can be obtained by $\hat{\omega} = (A^T A)^{-1} A^T b$.

C. *The Rotational Center and Translational Velocity Extraction*

The angular frequency ω of rotational disturbance motion, as discussed above, is estimated in terms of least squares method. It follows from Eqs. (11) and (12) that the translational velocity components in the $x-$direction, u', and $y-$direction, v', can be obtained if the rotational center (x_0, y_0) is known, as given by

$$u' = u(x_0, y_0), \tag{16}$$

and

$$v' = v(x_0, y_0). \tag{17}$$

In the following, we will develop an approach to find the rotational center (x_0, y_0) of the disturbance motion. For a given rotational center (x_0, y_0), the velocity components, $u(x_i, y_i)$ and $v(x_i, y_i)$ at point (x_i, y_i) can be estimated by

$$\hat{u}(x_i, y_i) = u(x_0, y_0) - \hat{\omega}(y_i - y_0), \tag{18}$$

and

$$\hat{v}(x_i, y_i) = v(x_0, y_0) + \hat{\omega}(x_i - x_0). \tag{19}$$

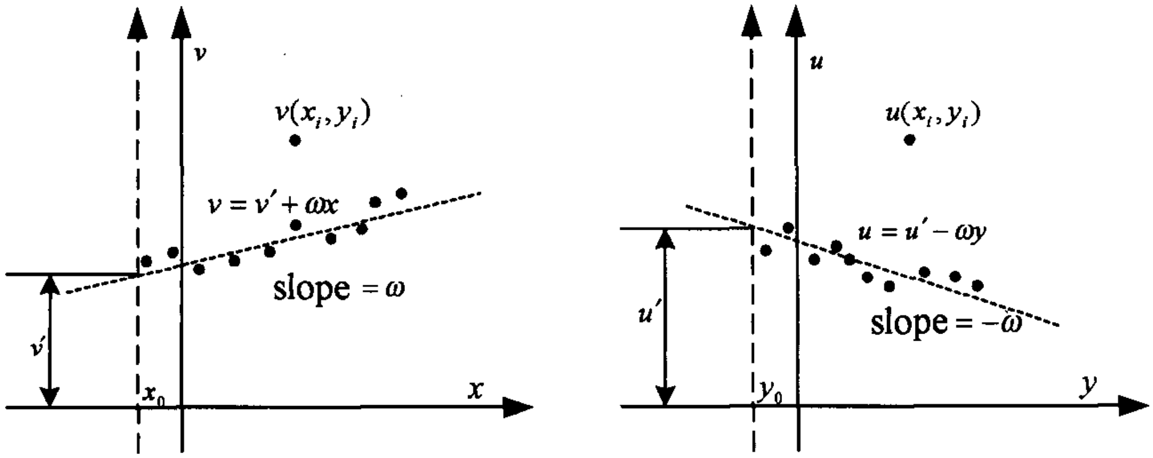
We then define the error function e as

$$e = \sum_{x_0 - p \le x_i \le x_0 + p} \sum_{y_0 - p \le y_i \le y_0 + p} \{[\hat{u}(x_i, y_i) - u(x_i, y_i)]^2 + [\hat{v}(x_i, y_i) - v(x_i, y_i)]^2\}, \tag{20}$$

where $\hat{u}(x_i, y_i)$ and $\hat{v}(x_i, y_i)$ are calculated from Eqs. (18) and (19), and $u(x_i, y_i)$ and $v(x_i, y_i)$ are obtained from the results of optical flow computation. In a searching basis, a point that can produce a smallest e measure over pixels in a certain region centered at this point would be the rotational center of this image frame. For instance, in this contribution a window of size $(2p + 1) \times (2p + 1)$, $p = 7$, is exploited as the region for such computation purpose. To locate the rotational center of the first image frame, whole image pixels have to be searched. Owing to the continuity constraint imposed by object inertia, the search region, after the first image frame, can be reduced to a neighborhood region centered around the previous frame rotational center. In this study, the searching block is chosen to be a 21×21 window centered at the previous frame rotational center estimated. With the rotational center being located by Eq. (20), the translational velocities u' and v' can be obtained from Eqs. (16) and (17). Using the estimated rotational center (x_0, y_0) and the translational velocities u', and v', together with $\hat{\omega}$ estimated previously, we can generate the counterbalance signal to stabilize the fluctuation in the current image frame.

III. EXPERIMENTAL SIMULATIONS

In order to confirm the validity of the proposed algorithm, computer simulations were conducted to stabilize a disturbed "Airplane" image sequence as described below. The image frame sequence is of



(a) (b)

Fig. 3. Motion vectors with disturbing noise in (a) $x - v$ and (b) $y - u$ plane.

size 200 × 200 pixels and contains 80 frames in length. The video sequence proceeds 16 frames per second, i.e., the sampling period T is 1/16 second. The synthetic disturbing displacements in translational movements t_x and t_y and rotational angle θ are all to be sinusoidal, as given by

$$
\begin{aligned}
t_x &= 4\sin(2\pi 0.5kT), \\
t_y &= 4\sin(2\pi 0.5kT),
\end{aligned}
$$

and

$$\theta = 0.04\sin(2\pi 0.5kT),$$

and the rotational center is located at (60, 60) in the first image frame. In the equations above, the units of t_x and t_y and θ are in pixel and radian, respectively.

A. Simulation 1

We apply the optical flow-based stabilization algorithm to the disturbed "Airplane" image sequence and set the maximal number of iteration in optical flow computation to be 64. To get rid of the boundary effect and reduce the computation complexity, we choose the center 100 × 100 pixels of the frame as the active block for optical flow operation. In our study, the estimate of the rotational center of the first image frame is involved with a full-plane search, whereas locating the rotation after the second image frame involves only a search region of ±10 pixels horizontally and vertically, i.e., a window of size 21 × 21, whose center is the rotational center of the previous frame. Two consecutive image frames after the optical flow computation are shown in Fig. 4, in which the white line vectors denote the velocity vectors at pixels of interests. After the 80 video frames were processed by the proposed image stabilizer, the results are shown in Figs. 5 and 6. In Fig. 5, the dotted lines denote the imposed velocity disturbance signals and the solid lines denote the corresponding velocities computed, which are to be compensated by the image stabilizer. Figs. 6(a) and 6(b), respectively, show the estimated x and y positions of the rotational center, in which the solid lines denote the true rotational center and the "×" symbols denote the rotational center estimated by the stabilizer. Averaged over these 80 frames, the mean absolute errors (MAEs) of the imposed signals and compensated signals of Figs. 5(a)-(c), i.e., in horizontal, vertical direction, and the rotational angle, are 7.038×10^{-4}, 0.076, and 0.188, respectively. The MAEs in estimating the rotational center, on average over the total 80 image frames, are 1.26 pixels and 1.14 pixels in the $x-$ and $y-$ directions, respectively.

We calculated the gray level differences of each pixel in the original images and its corresponding stabilized images over the 80 frames. In terms of this gray level difference, the mean absolute error (MAE) using the new optical flow-based image stabilization is 0.221, and the mean square error (MSE) is 0.075. By our experiment, the MAE and MSE of

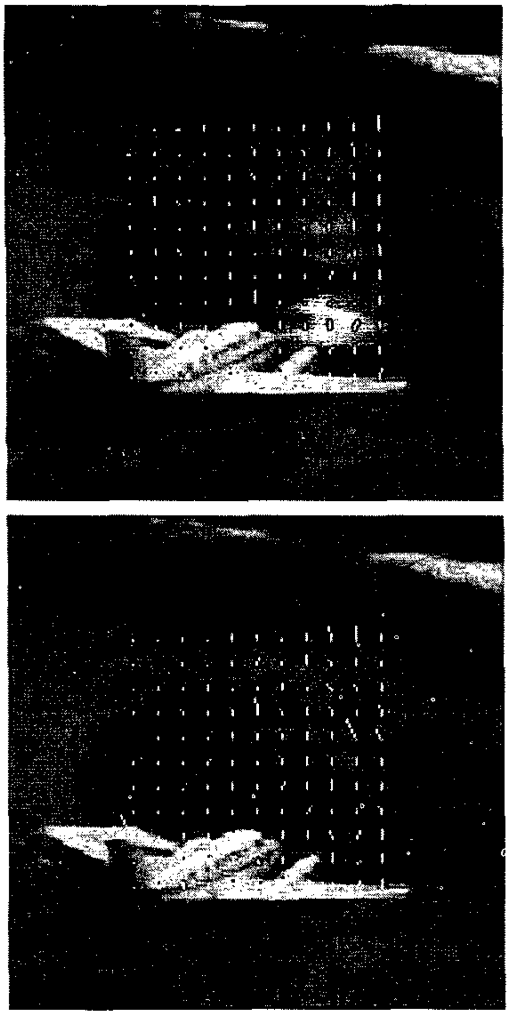


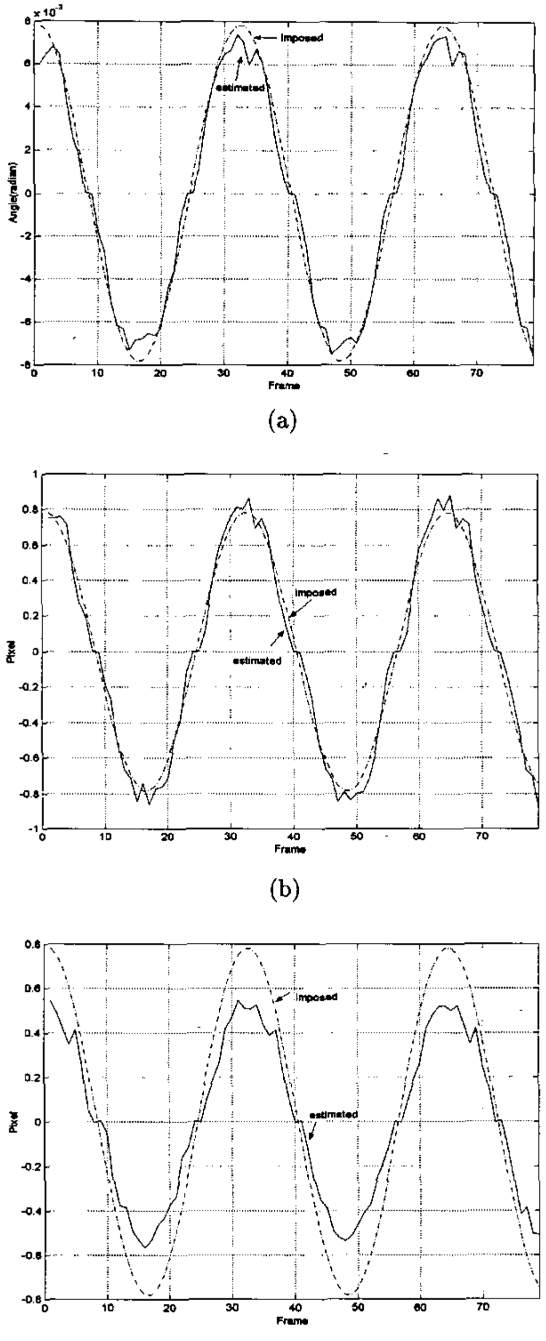
Fig. 4. Two consecutive frames estimated using optical flow method in the image sequence.

the image stabilization system based on the block matching method [7] is 1.989 and 4.624, respectively. The MAE measure of the proposed method yields one order of magnitude improvement; the MSE measure improves significantly with two orders of magnitude in error reduction, in comparison to the state-of-the-art block matching technique.

B. Simulation 2

The disturbance signals imposed in this example are given by

$$\begin{aligned} t_x &= 5\sin(2\pi 0.75kT), \\ t_y &= 3\sin(2\pi 0.75kT), \end{aligned}$$

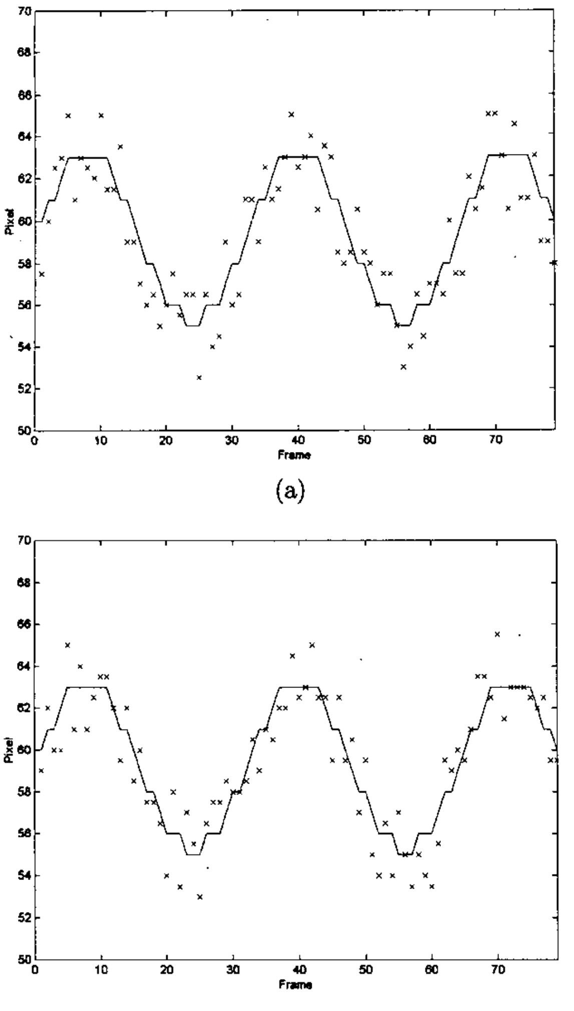


(a)

(b)

(c)

Fig. 5. The compensation signals estimated (a) rotational angular frequency, (b) translational velocity u in the x-direction, and (c)translational velocity v in the y-direction; in comparison with the imposed jiggling velocity signals.



(a)

(b)

Fig. 6. The estimated rotational center (×) and the true rotational center (-) of 80 image frames in (a) x-axis and (b) y-axis.

and

$$\theta = 0.03 \sin(2\pi 0.8kT),$$

and the rotational center is located at the image frame center (100, 100). The results of applying the optical flow based image stabilizer to this example are also successful. The resulting figures are omitted here for brevity, only the comparative performance with respect to blocking matching approach is summarized. Applying the block matching based image stabilization system to the disturbed video sequence of this example produces a MAE of 1.135 and MSE of 1.443. The MAE and MSE using the new optical flow based image stabilizing system are 0.349 and 0.155, respectively. Likewise, we have greatly improved the stabilization accuracy compared to the conventional blocking matching method.

According to the experimental results by changing the imposed sinusoidal amplitude and frequency, our image stabilization system using optical flow technique can remove the translational and rotational disturbance accurately when the amplitude-frequency product is smaller than four. For the first example, the amplitude-frequency product is 2 (4 × 0.5) while this product becomes 4 (5 × 0.8) for the second. Although the product of the second example is in the working boundary of amplitude-frequency product, the imposed disturbance can still be removed by the new stabilization loop. The amplitude-frequency product working bound meets the requirement in mobile video communication applications [10], in which the maximal jiggling frequency of 3 Hz and jiggling amplitude of 1.25 pixel are reported to be necessary.

IV. CONCLUSION

In this paper we have proposed an image stabilization system using optical flow technique. A computational scheme that facilitates the local motion vector field to estimate the global disturbing translational and rotational motions is derived. In addition to suppressing the translational disturbance, the developed DIS system is also capable of removing the rotational disturbance. Hence we have enhanced its performance significantly. Extensive tests have demonstrated that the new image stabilization system is effective in suppressing the undesirable jiggling motion using the rotational center, rotational angular frequency, and translational velocities we have estimated. The proposed DIS system is applicable to the various video consumer products of interests.

REFERENCES

[1] M. Oshima, T. Hayashi, S. Fujioka, and T. Inaji, "VHS camcorder with electronic image stabilizer," *IEEE Trans. on Consumer Electronics*, vol. 35, no. 4, pp. 749-758, Nov. 1989.

[2] K. Uomori, A. Morimura, H. Ishii, T. Sakaguchi, and Y. Kitamura, "Automatic image stabilizing system by full-digital signal processing," *IEEE Trans. on Consumer Electronics*, vol. 36, no. 3, pp. 510-519, Aug. 1990.

[3] K. Uomori, A. Morimura, H. Ishii, T. Sakaguchi, and Y. Kitamura, "Electronic image stabilisation system for

video cameras and VCRs," *SMPTE Journal* , vol. 101, no. 2, pp. 66-75, 1992.

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

[5] S. Kim, Y. Kim, K. Yim, H. Chung, K. Choi, Y. Kim, and G. Jung, "A fast motion estimator for real-time system," *IEEE Transactions on Consumer Electronics,* vol. 43, no. 1, pp. 24-33, February 1997.

[6] T. Kinugasa, N. Yamamoto, H. Komatsu, S. Takase, and A. Imaide, "Electronic image stabilizer for video camera use," *IEEE Transactions on Consumer Electronics,* vol. 3, no. 3, 520-525, Aug. 1990.

[7] J. K. Paik, Y. C. Park, and D. W. Kim, "An adaptive motion decision system for digital image stabilizer based on edge pattern matching," *IEEE Trans. on Consumer Electronics,* vol. 38, no. 3, pp. 607-615, Aug. 1992.

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

[9] C. Guestrin, F. Cozman, and M. G. Simoes, "Industrial applications of image mosaicing and stabilization," *Knowledge-Based Intelligent Electronic Systems,* vol. 2, pp. 174-183, 1998.

[10] A. Engelsberg and G. Schmidt "A comparative review of digital image stabilising algorithms for mobile video communications," *IEEE Transactions on Consumer Electronics,* vol. 45, no. 3, pp. 591-597, Aug. 1999.

[11] A. M. Tekalp, *Digital Video Processing,* Upper Saddle River, NJ: Prentice-Hall, 1995.

[12] B. K. P. Horn and B. G. Schunck, "Determining optical flow," *Artificial Intell.,* vol. 17, pp. 185-203, 1981.

[13] Y. Q. Shi, and H. Sun, Image and video compression for multimedia engineering: fundamentals, algorithms, and standards, CRC Press, 2000.

[14] A. D. Bimbo and J. L. C. Sanz, "Analysis of optical flow constraints," *IEEE Transactions on Image Processing,* vol. 4, pp. 460-469, April 1995.

Jyh-Yeong Chang received his B.S. degree in control engineering in 1976 and M.S. degree in electronic engineering in 1980, both from National Chiao Tung University, Taiwan. From 1976 to 1978 and 1980 to 1982, he was a research fellow at Chung Shan Institute of Science and Technology (CSIST), Taiwan. He received the Ph.D. degree in electrical engineering from North Carolina State University in 1987.

Since 1987, he has been an Associate Professor in the Department of Electrical and Control Engineering at National Chiao Tung University. His research interests include fuzzy sets and systems, image processing, pattern recognition, and neural network applications.

Wen-Feng Hu was born in Mia-Li County, Taiwan, R.O.C. in 1969. He received the B.S. degree in Mechanical Engineering from National Taiwan University of Science and Technology, Taiwan R.O.C in 1996. He received his M.S. degree in the Department of Electrical and Control Engineering at National Chiao Tung University, Taiwan R.O.C in 1999. Since then, he has been with Mustek Inc., Hsinchu, Taiwan in developing video phone. His current research interests include image processing, fuzzy theory, and digital signal processing.

Mu-Huo Cheng received the B.S. degree in electrical engineering from National Cheng-Kung University, Tainan, Taiwan, R.O.C. in 1980, and the M.S. degree in electronic engineering from National Chiao Tung University, Hsinchu, Taiwan, R.O.C. in 1987. He received his Ph.D. in the Department of Electrical and Computer Engineering at Carnegie Mellon University in 1995. His current research interests include adaptive signal processing, detection, and estimation.

Bo-Sen Chang was born in Taipei County, Taiwan, R.O.C. in 1968. He received his B.S. degree in the Department of Electrical Engineering from National Taiwan University, Taipei, Taiwan, R.O.C in 1998. He received his M.S. degree from Department of Electrical and Control Engineering at National Chiao Tung University, Taiwan R.O.C in 2000. He is now with VIA Tech. Inc., Taipei, Taiwan in developing multimedia software for PC. His current research interests include ASIC design, computer vision, and neural fuzzy hybrid systems.