Scientia Iranica B (2015) 22(6), 2134–2141

SCIENTIA IRANICA

Sharif University of Technology
Scientia Iranica
Transactions B: Mechanical Engineering
www.scientiairanica.com

Anti-shake system for digital dynamic images

S.J. Chang* **and W.H. Cai**

Department of Mechanical Engineering, National Yunlin University of Science and Technology, 123 University Road, Section 3, Douliou, Yunlin 64002, Taiwan, R.O.C.

Received 22 May 2014; received in revised form 16 April 2015; accepted 22 June 2015

KEYWORDS
Anti-shake;
Dynamic image;
Tremor;
Stabilization.

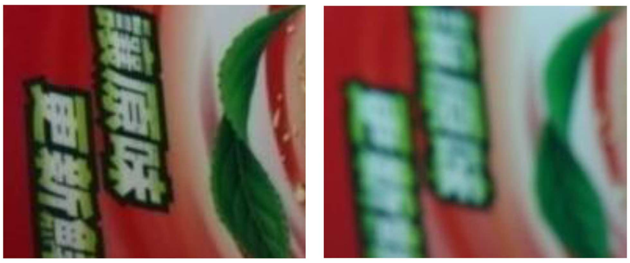
Abstract. This paper presents an anti-shake system to improve the quality of digital dynamic image sequences. The achievement of this study is to develop a digital dynamic image stabilization system using basic image processes and a signal-smoothing algorithm. The dynamic image stabilization method includes three steps: motion vector detection, motion vector smoothing, and image sequence reconstruction. The first step is to estimate the motion vector of dynamic image sequences using the three-step method. The second step is to separate involuntary motion vectors from motion vectors by the Fourier linear combiner algorithm. The last step is to compensate the unstable dynamic image sequences according to the involuntary motion vector and to build stable dynamic image sequence. The performance of the smoothing algorithm is also discussed in this paper. The Smooth Index (SI) is used to evaluate the performance of the anti-shake system presented in this paper. From the results of the experiments under some different conditions, the SI of the digital dynamic image sequences can reach 2.224 through the anti-shake system presented in this paper.

© 2015 Sharif University of Technology. All rights reserved.

1. Introduction

In past years, tiny sized digital cameras were developed vigorously. Digital cameras have become a basic element in most portable devices, such as mobile phones and tablet computers. People take pictures or dynamic videos and share them immediately on Facebook or Twitter in their daily life. Usually, people use portable devices to take pictures and shoot videos single handedly since the size of portable devices tends towards miniaturization. Hand tremors always occur when people take pictures and shoot videos, which causes their low quality. Therefore, shooting a stable video is a basic request when people use a built-in cameras of portable device [1,2]. Moreover, when people use the camera in a dim environment without sufficient light, the shutter release time will automatically be extended to increase exposure, and hand tremors will cause the photos to blur more easily if the shutter release time is too long. Thus, people always get blurry pictures because their hands shake involuntarily during the time interval of shutter release. Figure 1 shows a clear picture taken in an environment with sufficient light and a blurry picture taken in an environment without sufficient light. The hand tremor problem is even more serious when people shoot a dynamic video. The problem is getting worse while the resolution of image sensors and the magnified optical zooming of the camera are getting better. It also costs a lot to apply optical anti-shake technologies in thin portable devices because it is usually difficult to miniaturize the volume of optical anti-shake modules. Thus, developing an anti-shake algorithm for thin portable products to enhance video quality is very important. In this paper, some image processing technologies will be applied to estimate the motion vectors of dynamic image sequences. Then, the involuntary motion signals will

*. *Corresponding author. Tel.: 886-5-5342601 Ext. 4119;*
Fax: 886-5-5342062
E-mail address: changjye@yuntech.edu.tw (S.J. Chang)



(a) (b)

Figure 1. The pictures taken in the environments: (a) With sufficient light; and (b) without sufficient light.

be separated from the motion vector by the signal separation algorithm. Finally, the dynamic image will be reconstructed by compensating the involuntary shaking of each frame to enhance the quality of the dynamic image.

2. Image stabilization system

The anti-shake system, also known as the image stabilization system, is a desirable feature for digital cameras and portable video cameras to obtain stable pictures and videos. Most previous approaches focused on using either motion vector estimation [3-6] or frame position smoothing [7-15]. Some methods will cause time delay or phase change to the system. In this paper, an anti-shake algorithm was proposed to reduce the hand tremor effect using the Fourier series, which is based on a dynamic sinusoidal model. The system consists of three different stages: motion vector estimation, motion vector separation (extract the tremor signal in the motion vector), and compensation of image sequences to reduce the hand tremor effect. A block diagram of the anti-shake system presented in this paper is shown in Figure 2.

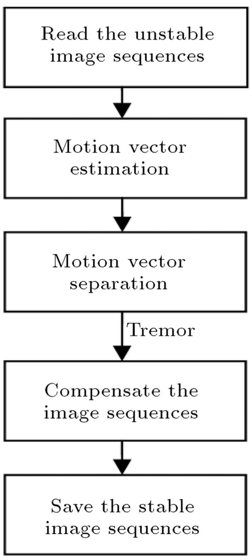


Figure 2. Block diagram of the anti-shake system.

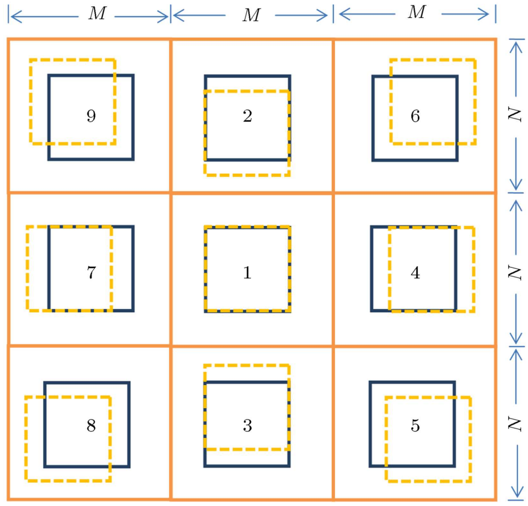


Figure 3. Search mode of motion estimation process.

2.1. Motion vector detection

The search mode of motion vector estimation is shown in Figure 3. Every frame will be compared with the previous frame to estimate motion between these two frames. For example, a $M \times N$ image in the center region of the previous frame in the image is considered as a reference image. Different regions in the current frame are compared with the reference image, and the most similar part in the current frame to the reference image will be found after comparison. The motion vector between the current and previous frame can be obtained according to the most similar region in the current frame. In the search process, the Direction of Minimum Distortion (DMD) is used to quantify the similarity between the current frame and the previous frame. The DMD, defined according to the sum of absolute difference (SAD), is shown as Eq. (1):

$$\mathrm{DMD}[I(t+1), I(t)] = \sum_{x=1}^{m} \sum_{y=1}^{n} |I_{x,y}(t+1) - I_{x,y}(t)| \,. \tag{1}$$

$I(t)$ is the tth frame of the video, x and y are the coordinates of the pixel of the image, and m and n are the sizes of the image [16]. In every comparing step, nine DMD values will be obtained by the comparing processes of nine different regions. The minimum DMD value region is considered as the most similar image to the center region in the previous frame. Thus, the motion vector between the current and previous frame will be obtained.

In Figure 3, the dashed and solid squares represent the compared images and the referenced images, respectively. In the comparing process, the nine compared images are defined by pixels of the offset in the x- and y-directions from the center region in the current frame. The center region in the current frame is also one of compared images. In every step

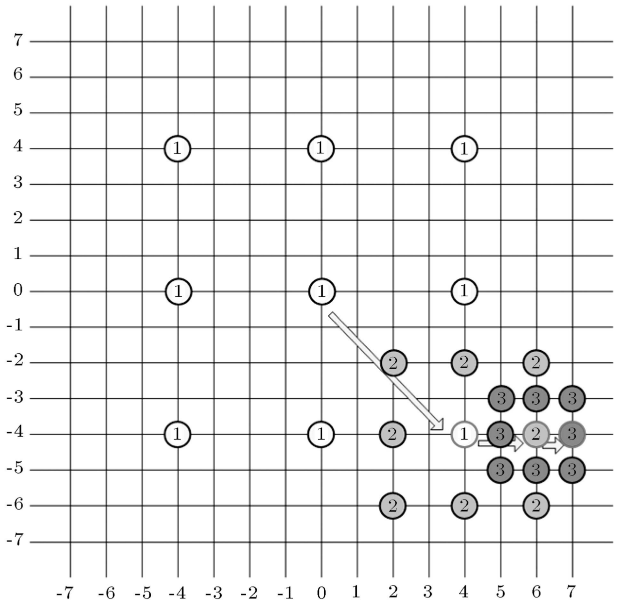


Figure 4. Example of the three step searching method.

of the comparing process, each compared image should be compared with the referenced image, as shown in Figure 3. Figure 4 shows an example of the three step searching method [17]. In the first searching step, the offset is set as 4 pixels. If the 5th region was the matched region with minimum DMD value in the first searching step, then the 5th region would be set as the new center region of the compared image of the second searching step. In the second searching step, the offset is set as 2 pixels. If the 4th region was the matched region with minimum DMD value in the second searching step, then the 4th region would be set as the new center region of the compared image of the third searching step. In the third searching step, the offset is set as 1 pixel. If the 4th region was the matched position with minimum DMD value in the third searching step, the searching process is finished and the motion vector of (7, -4) between the current and referenced image is obtained. By repeating the searching process and finding all motion vectors between two neighboring frames, the motion vector sequences can be obtained.

2.2. Motion vector smoothing

Measurement of hand motion can be described as the following equation, which is a combination of voluntary hand motion, involuntary hand tremor, and noise:

$$S_k = v_k + t_k + n_k, \tag{2}$$

where S_k is the measured signal, v_k is the desired or voluntary signal, t_k is the tremor or involuntary motion signal, and n_k is the noise due to measurement and other errors.

It is assumed that all three terms, v, t and n, are uncorrelated with each other. The problem is how

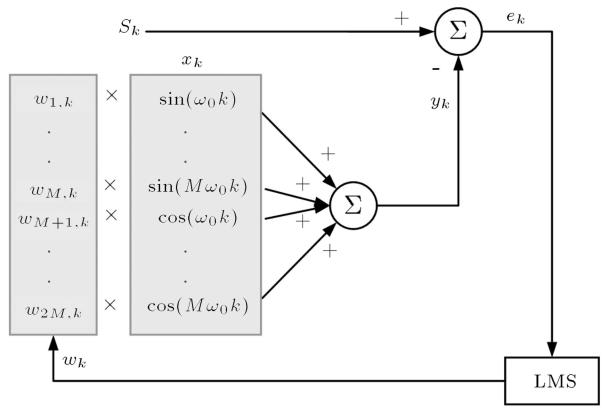


Figure 5. The block diagram of the Fourier linear combiner.

to separate v or t from the equation. The majority of smoothing algorithms are devoted to voluntary and involuntary motion separation using the Fourier linear combiner. It is very similar to the standard adaptive filter. This algorithm uses a group of sine and cosine functions to estimate voluntary motion and involuntary tremor. The assumption is that a hand tremor can be described as a linear combination of sine and cosine functions at different frequencies. The block diagram of the Fourier linear combiner is shown in Figure 5.

The input of the system, the linear combination of sine and cosine functions, is independent of the input signal, S_k. The assumption is that voluntary hand motion, v_k, is independent of involuntary hand tremor, t_k. The weights, w_k, are the Fourier coefficients. The x_k are listed as Eq. (3):

$$x_{r,k} = \begin{cases} \sin(r\omega_0 k), & 1 \le r \le M \\ \cos[(r-M)\omega_0 k], & M+1 \le r \le 2M \end{cases} \tag{3}$$

The y_k are listed as Eq. (4):

$$y_{r,k} = \sum_{r=1}^{M} [w_{r,k}\sin(r\omega_0 k) + w_{r+M,k}\cos(r\omega_0 k)], \tag{4}$$

where M is number of sine or cosine equations, r is the equation index, and ω_0 is the fundamental frequency of the system.

If we treat w, x, and y as column vectors, the cost function and update equations are listed as Eqs. (5) and (6), as follows:

$$\text{cost} = (S_k - y_k)^2 = (S_k - w_k^T x_k)^2 = e_k^2, \tag{5}$$

$$w_{k+1} = w_k + 2\mu x_k e_k, \tag{6}$$

where μ is the learning rate in the LMS algorithm to move the cost function toward its local minimum.

At first, w_1 is initialized to zero, and the first x_k and y_k for each k are calculated. Then, the cost

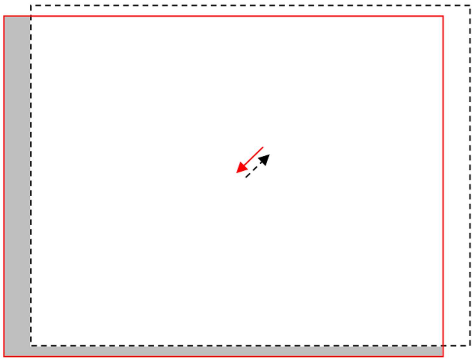


Figure 6. The schematic drawing of the compensation between two frames.

function can be obtained by using input S_k and y_k just calculated. Finally, w_k is updated, according to Eq. (6). w_k can be viewed as the Fourier coefficients. The Fourier linear combiner can effectively estimate and cancel the periodic interference of the known frequency. It is essentially an adaptive notch filter, as $M = 1$, and it will be a multi-notch filter, as $M \geq 1$. The advantage of the Fourier linear combiner is that it can track the amplitude and phase of an unknown oscillating signal.

2.3. Image sequence reconstruction

In the image sequence reconstruction process, the original frame should be compensated according to the involuntary hand tremor vector separated from the Fourier linear combiner. The schematic drawing of the compensation between two frames is shown in Figure 6. The dashed square is the original frame, and the dashed arrow is the involuntary hand tremor vector separated from the Fourier linear combiner. The compensation is to move the original frame along the opposite direction to the involuntary hand tremor vector. The solid arrow is the compensation vector, and the solid square is the compensated frame. The compensation process will cause the invalid image area, shown as the gray area in Figure 6. Each compensation is different in quantity and direction, since the involuntary hand tremor vector of each frame is also different. Therefore, the margin of each frame should be cut based on the involuntary hand tremor vector in all directions after the compensation process.

The final step in the image stabilization system is to reconstruct the image sequences and save them as a new video file. Figure 7 shows the schematic drawing of the image sequence reconstruction. The three step method estimates the motion vector of the image sequence. The motion vector consists of voluntary motion, involuntary tremor, and other noise due to measurement and other errors. Then, the Fourier linear combiner is used to separate the involuntary tremor from the motion vector. Each frame can be compensated based on the involuntary tremor vector.

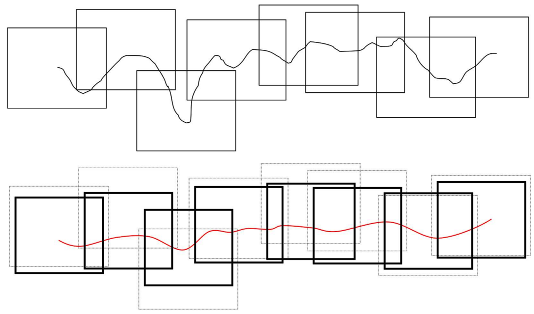


Figure 7. The schematic drawing of the image sequences reconstruction.

The stable/smooth image sequences can finally be obtained by the above processes.

3. Experiments

The effectiveness of the proposed method is confirmed by some experiments on real persons. From the experimental results, it was found that the proposed system provides robust motion compensation of image sequences under various conditions. When people take a video shoot using a portable device, the desired or voluntary motion is usually in the lower frequency range, and the involuntary motion is usually in the higher frequency range. The noise due to measurement and other errors are also in the higher frequency range, as usual. However, the Fourier linear combiner can be divided into two parts: the lower frequency terms and the higher frequency terms. Voluntary motion can be described by the lower frequency terms, and the involuntary hand tremor vector can be described by higher frequency terms.

3.1. Performance index

In this study, the Smooth Index (SI) is used to quantify the performance of the anti-shake system for the dynamic image. The smooth index is based on the Smooth Degree (SD). It is an index to evaluate the performance when smoothing a rough signal. The definition of a smooth degree is shown in Eq. (7) [18]:

$$\mathrm{SD} = \frac{1}{N-1} \sum_{t=2}^{N} \left|\mathrm{MV}(t) - \mathrm{MV}(t-1)\right|, \tag{7}$$

where $\mathrm{MV}(t)$ is the motion vector of the tth frame, $\mathrm{MV}(t-1)$ is the motion vector of the $(t-1)$th frame, and N is the number of frames of the image sequence.

$(N-1)$ motion vectors can be estimated from a N-frames image sequence because one motion vector estimation needs two frames: current and previous frames. The SD in Eq. (7) is defined by the average of the absolute value of the motion vectors estimated from the image sequences. The rougher the original motion vector sequence is, the greater the SD will be. After motion vector smoothing, the motion vector sequence will become smoother, and the SD will be smaller. The smooth index is defined in Eq. (8). A greater SI means better smoothing performance:

$$\mathrm{SI} = \frac{\mathrm{SD}_{\text{before smoothing}}}{\mathrm{SD}_{\text{after smoothing}}}. \tag{8}$$

3.2. Parameters design

Some parameters of the Fourier linear combiner need to be determined before implementation of the anti-shake system. In this study, the number of sine or cosine equations (M), the fundamental frequency of the system (ω_0), and the learning rate in the LMS algorithm (μ) are chosen as the important parameters to affect the performance of the motion vector sequence smoothing. The three factors and their levels are listed in Table 1.

Table 1 is a full factorial experiment array including 32 experiments with all combinations of M, μ, and ω_0. The levels of the three factors are 2, 4, and 4, respectively. In every experiment, 7 different video files are smoothed through the smoothing algorithm and the SI values are also calculated. The experimental results are all listed in Table 2. Seeing the experimental results in Table 2, optimal smoothing performance occurs in the 30th experiment, since the SI value is the larger, the better. The SI value is 2.224 with the following factor combination: $M = 3$, $\mu = 0.13$, and $\omega_0 = 4$ Hz. The original experimental results (SD) of the 30th experiment are listed in Table 3. The SD values are reduced almost 53% through the smoothing algorithm.

3.3. Anti-shake experiments

Several real video shooting experiments are also carried out using the anti-shake system in this study. The parameters of the smoothing algorithm are set as the optimal parameter combination with $M = 3$, $\mu = 0.13$, and $\omega_0 = 4$ Hz. The video shoots are taken by two different persons in indoor and outdoor environments. Therefore, the performance of the smoothing algorithm can be tested with different hand-tremors and in different shooting environments. Figures 8 and 9 are

Table 1. List of factors and levels.

	M	μ	ω_0
1	2	0.10	2
2	2	0.10	4
3	2	0.10	6
4	2	0.10	8
5	2	0.11	2
6	2	0.11	4
7	2	0.11	6
8	2	0.11	8
9	2	0.12	2
10	2	0.12	4
11	2	0.12	6
12	2	0.12	8
13	2	0.13	2
14	2	0.13	4
15	2	0.13	6
16	2	0.13	8
17	3	0.10	2
18	3	0.10	4
19	3	0.10	6
20	3	0.10	8
21	3	0.11	2
22	3	0.11	4
23	3	0.11	6
24	3	0.11	8
25	3	0.12	2
26	3	0.12	4
27	3	0.12	6
28	3	0.12	8
29	3	0.13	2
30	3	0.13	4
31	3	0.13	6
32	3	0.13	8

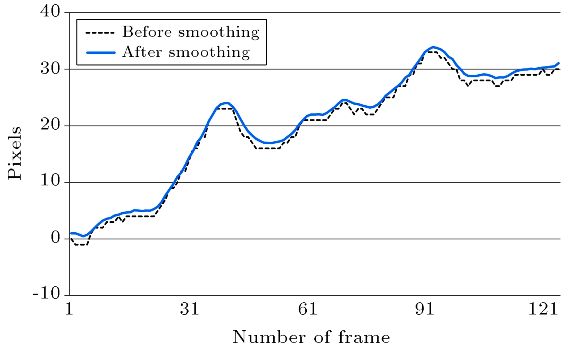


Figure 8. The motion vectors diagrams of the video taken by the first person in the indoor environment.

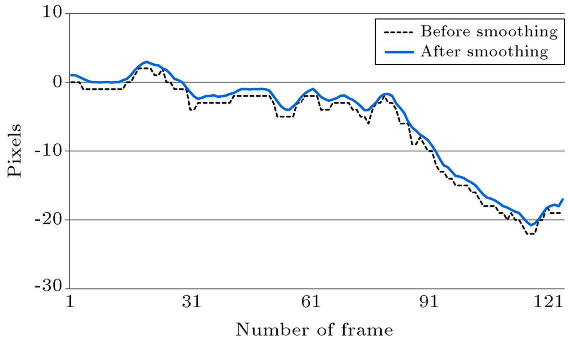


Figure 9. The motion vectors diagrams of the video taken by the first person in the outdoor environment.

Table 2. Experimental results of parameters design (SI).

Exp.	Video 1	Video 2	Video 3	Video 4	Video 5	Video 6	Video 7	Ave. SI
1	1.779	1.671	1.827	1.948	1.845	1.94	1.945	**1.851**
2	1.861	1.882	1.906	1.744	1.892	1.747	1.725	**1.822**
3	1.883	1.855	1.879	1.723	1.845	1.925	1.965	**1.868**
4	1.911	1.961	2.042	1.924	2.027	2.017	2.069	**1.993**
5	2.036	2.055	2.094	1.904	2.057	2.096	2.016	**2.037**
6	2.022	1.967	2.088	2.065	1.973	1.881	2.023	**2.003**
7	1.98	1.888	1.93	1.933	1.901	1.905	1.968	**1.929**
8	1.955	1.927	1.891	1.858	1.836	2.084	2.022	**1.939**
9	1.991	2.047	2.033	2.039	2.003	1.949	2.053	**2.016**
10	1.943	1.898	1.964	1.896	2.082	1.986	1.98	**1.964**
11	1.926	1.967	1.865	1.934	2.022	2.099	1.972	**1.969**
12	1.968	2.056	1.9	1.908	1.873	1.888	1.927	**1.931**
13	2.013	2.034	2.102	1.994	2.194	2.008	1.986	**2.047**
14	1.964	1.915	1.922	1.806	1.873	1.859	1.766	**1.872**
15	1.854	1.87	1.891	1.957	1.949	1.841	1.889	**1.893**
16	1.944	2.064	2.014	2.04	2.068	2.062	2.009	**2.029**
17	1.989	2.08	1.913	1.962	1.958	1.987	1.865	**1.965**
18	1.985	1.92	1.992	2.028	1.904	2.062	1.868	**1.966**
19	1.875	1.794	1.751	1.929	1.841	1.919	1.874	**1.855**
20	1.814	1.731	1.866	1.696	1.735	1.647	1.772	**1.752**
21	1.951	1.708	1.82	1.955	1.9	1.923	1.837	**1.871**
22	1.918	1.948	2.068	1.89	2.06	1.857	1.879	**1.946**
23	1.888	1.967	1.914	2.059	1.876	2.095	1.941	**1.963**
24	1.954	1.854	1.923	1.737	2	1.75	2.012	**1.89**
25	1.937	1.94	1.964	1.929	1.978	1.924	1.899	**1.939**
26	1.963	1.863	1.947	1.8	2.047	2.002	2.085	**1.958**
27	2.091	1.949	1.957	1.791	2.029	2.045	1.914	**1.968**
28	1.943	1.845	1.829	1.993	1.834	1.933	1.94	**1.902**
29	1.987	2.044	2.135	2.058	1.947	2.242	1.902	**2.045**
30	2.3	2.143	2.26	2.207	2.336	2.096	2.225	**2.224**
31	1.841	2.095	2.04	2.081	2.043	2.151	2.057	**2.044**
32	2.083	1.964	2.068	2.006	1.943	2.107	1.906	**2.011**

Table 3. Experimental results of the 30th experiment.

	Video 1	Video 2	Video 3	Video 4	Video 5	Video 6	Video 7
$\mathbf{SD}_{\textbf{before smoothing}}$	11.856	10.739	13.133	11.611	10.589	11.4	12.928
$\mathbf{SD}_{\textbf{after smoothing}}$	5.317	5.011	5.811	5.261	4.533	5.439	5.811

the motion vector diagrams of the videos taken by the first person in the indoor and outdoor environments, respectively. The dashed curve is the motion vectors before smoothing and the solid curve is the motion vectors after smoothing. Figures 10 and 11 are the motion vector diagrams of the videos taken by the second person in indoor and outdoor environments, respectively. From Figures 8 to 11, the motion vector curve after smoothing looks smoother than the motion vector curve before smoothing. It means that the anti-shake system can reduce the effects of different hand tremors in either indoor or outdoor environments.

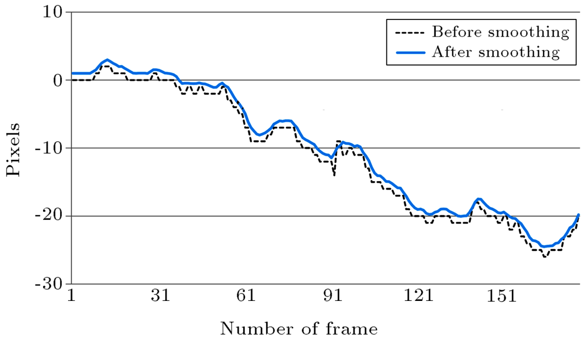


Figure 10. The motion vectors diagrams of the video taken by the second person in the indoor environment.

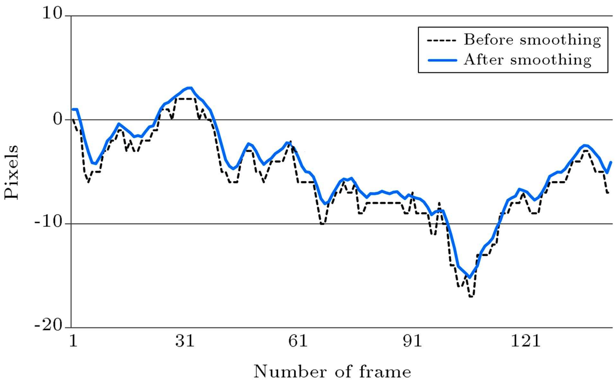


Figure 11. The motion vectors diagrams of the video taken by the second person in the outdoor environment.

Conclusions

In this paper, study of an anti-shake system for dynamic images is presented. The smoothing algorithm includes three stages: motion signal estimation, motion signal smoothing, and tremor compensation, to reduce the hand tremor effect when people are shooting a video. The Fourier linear combiner algorithm is used to separate voluntary hand motion and involuntary hand tremor. Some experiments are carried out to find the optimal combination of the parameters of the Fourier linear combiner. The experimental results show that the motion vector curve after smoothing is much smoother than the motion vector curve before smoothing. In addition, the SI value is used to quantify the performance of the anti-shake system. Using the results of experiments with some video files taken in different situations, the SI value can be greater than 2 through the anti-shake algorithm presented in this paper.

Acknowledgments

The authors would like to thank the Ministry Of Science and Technology (MOST) for financial support of the project (Grant number: MOST99-2218-E-224-002-).

References

1. Battiato, S., Castorina, A., Guarnera, M. and Vivirito, P. "A global enhancement pipeline for low-cost imaging devices", *IEEE Transactions on Consumer Electronics*, **49**(3), pp. 670-675 (2003).

2. Bosco, A., Findlater, K., Battiato, S. and Castorina, A. "Noise reduction filter for full-frame imaging devices", *IEEE Transactions on Consumer Electronics*, **49**(3), pp. 676-682 (2003).

3. Uomori, K., Morimura, A., Ishii, H., Sakaguchi, T. and Kitamura, Y. "Automatic image stabilizing system by full-digital signal processing", *IEEE Transactions on Consumer Electronics*, **36**(3), pp. 510-519 (1990).

4. Vella, F., Castorina, A., Mancuso, M. and Messina, G. "Digital image stabilization by adaptive block motion vectors filtering", *IEEE Transactions on Consumer Electronics*, **48**(3), pp. 796-800 (2002).

5. Yuan, D., Liu, M., Yin, J. and Hu, J. "Camera motion estimation through monocular normal flow vectors", *Pattern Recognition Letters*, **48**, pp. 59-64 (2015).

6. Kim, S.W., Yin, S., Yun, K. and Choi, J.Y. "Spatio-temporal weighting in local patches for direct estimation of camera motion in video stabilization", *Computer Vision and Image Understanding*, **118**, pp. 71-83 (2014).

7. Pan, Z. and Ngo, C.W. "Selective object stabilization for home video consumers", *IEEE Transactions on Consumer Electronics*, **51**(4), pp. 1074-1084 (2005).

8. Ertürk, S. "Digital image stabilization with sub-image phase correlation based global motion estimation", *IEEE Transactions on Consumer Electronics*, **49**(4), pp. 1320-1325 (2003).

9. Ertürk, S. "Image sequence stabilisation: Motion vector integration (MVI) versus frame position smoothing (FPS)", *2nd International Symposium on Image and Signal Processing and Analysis*, Pula, Croatia, pp. 266-271 (2001).

10. Ertürk, S. "Real-time digital image stabilization using Kalman filters", *Real-Time Imaging*, **8**(4), pp. 317-328 (2002).

11. Matsushita, Y., Ofek, E., Weina, G., Tang, X. and Tang, X. "Full-frame video stabilization with motion inpainting", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **28**(7), pp. 1150-1163 (2006).

12. Chang, H.C., Lai, S.H. and Lu, K.R. "A robust real-time video stabilization algorithm", *Journal of Visual Communication and Image Representation*, **17**(3), pp. 659-673 (2006).

13. Liang, Y.M., Tyan, H.R., Chang, S.L., Liao, H.Y.M. and Chen, S.W. "Video stabilization for a camcorder mounted on a moving vehicle", *IEEE Transactions on Vehicular Technology*, **53**(6), pp. 1636-1648 (2004).

14. Tsai, T.H. "Design and implementation of efficient video stabilization engine using maximum a posteriori

estimation and motion energy smoothing approach", *IEEE Transactions on Circuits and Systems for Video Technology*, **22**(7), pp. 817-830 (2011).

15. Rohania, B., Yazicioglua, Y., Mutlub, M., Ogucub, O., Akgulb, E. and Saranlib, A. "Lagrangian based mathematical modeling and experimental validation of a planar stabilized platform for mobile systems", *Journal of Computational and Applied Mathematics*, **259**, pp. 955-964 (2014).

16. Jain, J. and Jain, A. "Displacement measurement and its application in interframe image coding", *IEEE Transactions on Communications*, **29**(12), pp. 1799-1808 (1981).

17. Sun, N.N., Chao, F. and Xia, X. "An effective three-step search algorithm for motion estimation", *IEEE International Symposium on IT in Medicine & Education*, Jinan, China, pp. 400-403 (2009).

18 Vaz, C., Kong X. and Thakor, N. "An adaptive estimation of periodic signals using a Fourier linear combiner",*IEEE Transactions on Signal Processing*, **42**(1), pp. 1-10 (1994).

Biographies

Shyang-Jye Chang received BS, MS, and PhD degrees in Mechanical Engineering from National Cheng-Kung University, Taiwan, in 1986, 1988 and 1993, respectively. He is currently Assistant Professor in the National Yunlin University of Science and Technology, Douliou, Yunlin. His research interests include system integration, applications of sensors and actuators, applications of sun energy and 3D printing technology.

Wen-Hao Cai received his MS degree from the National Yunlin University of Science and Technology, Douliou, Yunlin, in 2012, and is currently employed by the Sinon Corporation, Taiwan. His research interests include image processing and design of experiments.