LETTER

Digital Image Stabilization Based on Correction for Basic Reference Frame Jitter

Yuefei ZHANG^{†a)}, Mei XIE[†], *Nonmembers*, and Ling MAO[†], *Member*

SUMMARY In this letter, we first study the impact of the basic reference frame jitter on the digital image stabilization. Next, a method for stabilizing the digital image sequence based on the correction for basic reference frame jitter is proposed. The experimental results show that our proposed method can effectively decrease the excessive undefined areas in the stable image sequence resulting from the basic reference frame jitter.

key words: digital image stabilization, compensating motion vector, correction

1. Introduction

Cameras hand-held or mounted on moving vehicles usually suffer from image instability due to unwanted camera motions or jitter. Therefore, digital image stabilization (DIS) has been proposed to produce the stable image sequence by removing the unwanted camera motions (i.e., frame jitter) while retaining the intentional camera motions (such as camera panning) of the acquired image sequence. A DIS system comprises two major parts: global motion estimation (GME) and motion compensation (MC). The purpose of the first part is to estimate the global camera motions through three major steps: 1) extract the features from the current input frame (current frame) and its previous frame (reference frame) of the acquired image sequence; 2) calculate the global motion vector (GMV) of the current frame with respect to the reference frame; 3) repeat step 1 and step 2 on each frame, except the first frame, of the acquired image sequence to produce a series of GMVs. The features extracted from the step 1 of GME can be sub-image block [1], representative point [2], bit-plane [3] and feature points [4]. Next, the global motion can be calculated by feature-based matching (correlation) [2], [3] or solving motion model [4], [5]. The purpose of MC part is to produce the stable image sequence generally through two steps: 1) smooth or low-pass filter the frame positions or global motions to generate the compensating motion vectors (CMVs) for the acquired image sequence; 2) shift each frame of the acquired image sequence to its stable position, according to the CMVs, to produce the stable image sequence. The MC methods can be divided into two kinds: frame position smooth [6]–[8] and global motion smooth [1], [9]. The first kind of MC method

generates the CMVs usually through smoothing the frame positions generated by accumulating the GMVs, while the second one directly smoothes the global motion to generate the CMVs.

The first frame of an unstable image sequence is usually selected as the first reference frame (named basic reference frame (BRF) in this letter) in above DIS methods. Accordingly, a DIS algorithm starts from the BRF which is always assumed to be stable (i.e., the global motion of this frame only contains intentional camera motion such as camera panning). However, it is difficult to keep the BRF stable in actual applications, especially when the video camera is placed on an unstable platform such as plane. As a result, the CMVs generated under this condition will deviate from the real CMVs, which will affect the performances of the traditional DIS methods at the same time. In order to deal with this problem, we first study the relationship between the BRF jitter and the CMVs, and then propose a DIS method based on correction for BRF jitter. It should be emphasized that we preliminarily consider one kind of MC method (frame position smooth) when deriving the relationship between BRF jitter and CMVs. Another kind of MC method (global motion smooth) will be considered in our future study.

2. Proposed Method

2.1 Impact of Basic Reference Frame Jitter on CMVs

In this section, we derive the relationship between the BRF jitter and the CMVs. Figure 1 first illustrates the relationship between the positions of the frames in the original image sequence and those of the frames in the stable image sequence, where the black points denote the positions of two consecutive frames in the original image sequence and the solid line denotes the frame positions in the stable image sequence. These two black points deviate from the solid line due to the unwanted camera motion or frame jitter. We here refer to the unwanted camera motion vector of frame i as $J(i)$ in Fig. 1, so the position of this frame in the original image sequence can be got as Eq. (1):

$$P_o(i) = P_s(i) + J(i) \quad (1)$$

where $P_o(i)$ and $P_s(i)$ are the positions of frame i in the original and the stable image sequences, respectively. Hence, the GMV of frame i with respect to frame $i-1$ in the original image sequence can be calculated as Eq. (2):

Manuscript received January 12, 2010.

Manuscript revised July 1, 2010.

[†]The authors are with School of Electronic Engineering, University of Electronic Science and Technology of China, Chengdu, China.

a) E-mail: yuefeicuit@yahoo.cn

DOI: 10.1587/transinf.E93.D.3149

Fig. 1 The positions of the frames in original and stable image sequences. Black points denote the positions of two consecutive frames in the original image sequence; solid line denotes those in the stable image sequence.

$$\begin{aligned}
 GMV(i) &= P_o(i) - P_o(i-1) \\
 &= [P_s(i) + J(i)] - [P_s(i-1) + J(i-1)] \\
 &= [P_s(i) - P_s(i-1)] + [J(i) - J(i-1)] \\
 &= GMV_1(i) + [J(i) - J(i-1)] \quad (2)
 \end{aligned}$$

where $GMV_1(i)$ is the global motion vector of frame i with respect to frame $i-1$, in the stable image sequence. For a given image sequence with $n+1$ frames whose first frame is considered as the BRF, the GMVs from Eq. (2) are accumulated to generate the frame position signal (FPS) which represents the frame positions of the given image sequence, as follows:

$$\begin{aligned}
 FPS(i) &= \sum_{t=1}^i GMV(t) = \sum_{t=1}^i GMV_1(t) + \sum_{t=1}^i [J(t) - J(t-1)] \\
 (i &= 1, 2, \dots, n) \quad (3)
 \end{aligned}$$

where $FPS(i)$ is the position of frame i in the given image sequence. The last term of Eq. (3) can be expanded as Eq. (4).

$$\begin{aligned}
 \sum_{t=1}^i [J(t) - J(t-1)] &= J(1) - J(0) + J(2) - J(1) + \dots \\
 &\dots + J(i) - J(i-1) = J(i) - J(0) \quad (4)
 \end{aligned}$$

From Eq. (3) and Eq. (4), we get

$$\begin{aligned}
 FPS(i) &= \sum_{t=1}^i GMV_1(t) + J(i) - J(0) \\
 &= SFPS(i) + J(i) - J(0) \quad (i = 1, 2, \dots, n) \quad (5)
 \end{aligned}$$

where $SFPS(i)$ is the position of frame i in the stable image sequence, because $GMV_1(i)$ is the global motion vector of the frame i with respect to the frame $i-1$, in the stable image sequence. Smoothing or low-pass filtering the FPS, according to the frame position smooth method [6], [8], can generate the frame positions in the stable image sequence (i.e., the $SPFS$ in Eq. (5)). Next, the CMV for frame i can be generated by subtracting $SFPS(i)$ from $FPS(i)$ as follows:

$$\begin{aligned}
 CMV(i) &= FPS(i) - SFPS(i) = J(i) - J(0) \\
 (i &= 1, 2, \dots, n) \quad (6)
 \end{aligned}$$

where $J(i)$ and $J(0)$ are the unwanted camera motion vectors

Fig. 2 The impact of basic reference frame jitter on stable image sequences. (a) is a sample frame of the stable image sequence obtained by using the real CMV; (b) is a sample frame of the stable image sequence obtained by using the distorted CMV.

of frame i and frame 0 (i.e., the BRF) in the given image sequence, respectively.

If the BRF is unstable (i.e., $J(0) \neq 0$), the CMV for frame i will not only contain the unwanted camera motion vector but also contain that of the BRF. This means that the excessive part of Eq. (6): $J(0)$ distorts the real CMV: $J(i)$, since the real CMV for frame i should only contain the unwanted camera motion vector of this frame. As a result, the positions of all frames in the stable image sequence compensated by using the distorted CMVs will deviate from the real stable positions. Moreover, the excessive compensating on frame position will also increase the undefined area of the frames in the stable image sequence, as the black area illustrated in Fig. 2. Figure 2 (a) and (b) are the sample frames of the stable image sequences compensated by using real CMVs and distorted CMVs, respectively.

2.2 DIS Method Based on Correction for Basic Reference Frame Jitter

The proposed DIS method first estimates the unwanted camera motion vector of the BRF, and then eliminates the excessive parts from the distorted CMVs according to the estimated result.

To do this, the CMVs for m consecutive frames in the original image sequence are averaged as Eq. (7):

$$\text{meanc} = \frac{1}{m} \sum_{i=1}^m CMV(i) = \frac{1}{m} \sum_{i=1}^m J(i) - \frac{1}{m} \sum_{i=1}^m J(0) \quad (7)$$

where $CMV(i)$ is the result of Eq. (6). The first term on the right of Eq. (7) can be neglected for a given m . This is because that the judgment of discriminating the intentional and the unwanted camera motions is performed on the basis of smoothness. Intentional camera motion such as camera panning represents exclusively regular movement, while unwanted camera motion or frame jitter is random of nature. Accordingly in frequency spectrum, regular movement comprises low-frequency component, while frame jitter constitutes the high-frequency component. Consequently, we assume that unwanted camera motions or frame jitters are additive high-frequency noise of zero-mean on the intentional camera motions. Hence, we get:

$$\text{meanc} = -\frac{1}{m} \sum_{i=1}^m J(0) = -J(0) \quad (8)$$

From Eq. (8), we can derive the unwanted camera motion

vector of the BRF as shown in Eq. (9):

$$J(0) = -\text{mean}c \quad (9)$$

Next, correcting the distorted CMVs according to Eq. (6) and Eq. (9) can produce the real CMVs for the original image sequence as follows:

$$\begin{aligned} \text{CMV}'(i) &= \text{CMV}(i) + J(0) = \text{CMV}(i) - \text{mean}c = J(i) \\ (i &= 1, 2, \dots, n) \end{aligned} \quad (10)$$

where $\text{CMV}'(i)$ is the corrected CMV for frame i . Note that $\text{CMV}'(i)$ only contains the unwanted camera motion of frame i . The final stable image sequence can be produced by shifting the frames of the original image sequence to their stable positions according to the corrected CMVs.

It should be noted that the GME part of the proposed DIS method mainly follows the sub-image matching based method [1] since our study mainly concentrates on the correction for CMVs. The whole procedure of our proposed method is described as follows:

- Step 1: Estimate the GMVs of the original image sequence.
- Step 2: Accumulate the GMVs to yield the FPS in Eq. (3).
- Step 3: Smooth (Kalman filter is applied here) the FPS to yield the $SFPS$ in Eq. (5).
- Step 4: Subtract the $SFPS$ from the FPS to yield the CMVs.
- Step 5: Select the CMVs for m consecutive frames of the original image sequence and average them to yield the unwanted camera motion vector of the BRF as Eq. (8).
- Step 6: Correct all the CMVs by using the motion vector obtained in the step 5.
- Step 7: Shift the frames of the original image sequence to their stable positions, according to the corrected CMVs, to yield the final stable image sequence.

3. Experimental Results and Discussion

In this section, the performance of the proposed method is evaluated and compared with two traditional DIS methods: Kalman filter (KF) based method [6] and Low-pass filter (LPF) based method [8]. To do this, four real image sequences: IS1, IS2, IS3, IS4 captured by a video camera under different conditions are used for testing as shown in the first column of Fig. 3. The global motions of IS1 and IS2 only contain the unwanted camera motion and those of IS3 and IS4 contain both the unwanted camera motion and the intentional camera motion (horizontally panning). Figure 4 shows the relationship between the m of Eq. (8) and the $\text{mean}c$ both in horizontal direction and vertical. We can observe that the estimated result of each image sequence converges to a stable value when m is above 40. Hence, the m is set to 40 experimentally. It should be noted that the BRF jitters of IS1, IS3 and IS4 happen in horizontal direction while that of IS2 happens in vertical.

Figure 3 shows the sample frames in the stable image sequences obtained by applying different DIS methods on

Fig. 3 The sub-figures in the 1st column are the sample frames of the testing image sequences; the sub-figures in the 2nd, 3rd and 4th columns are the sample frames of the KF method result, the LPF method result and the proposed method result respectively.

Fig. 4 The relationship between m and $\text{mean}c$. The left subfigure shows the results in horizontal direction and the right one in vertical. The 'o', '+', '□', and '♦' denote the results of IS1, IS2, IS3 and IS4, respectively.

IS1, IS2, IS3 and IS4. The sub-figures in the 2nd, 3rd and 4th columns are the sample frames of the KF method result, the LPF method result and the proposed method result, respectively. We can observe that the results of the proposed method have smaller undefined areas than those of another two DIS methods.

In order to get a credible result, we subtract the undefined areas (UAs) of the proposed method result from that of a traditional DIS method result as Eq. (11):

$$\begin{cases} \text{UA}(i) = W \cdot H - (W - |\text{CMV}_X(i)|)(H - |\text{CMV}_Y(i)|) \\ \text{DUA}(i) = \text{UA}_T(i) - \text{UA}_P(i) \end{cases} \quad (i = 1, 2, \dots, n) \quad (11)$$

where W and H are the width and the height of the frames in the testing sequences respectively, CMV_X and CMV_Y are the CMVs in horizontal and vertical directions respectively, $\text{UA}_T(i)$ and $\text{UA}_P(i)$ are the UAs of frame i of the traditional method result and that of the proposed method result respectively. Figure 5 shows the undefined area differences. We can observe that the DUAs for most frames are positive, which means that the results of our method have smaller UA than those of the other two methods. Nevertheless, there are some negative values for each image sequence. For IS1-IS4 whose BRF jitters happen in horizontal direction or vertical, such results come up on the frame whose unwanted jitter not

Fig. 5 The difference between the undefined area of the traditional DIS method result and the proposed method result. The sub-figures from the top to the bottom are the results for IS1, IS2, IS3, IS4, respectively. The ‘—’ denote the difference between the undefined area of KF result and that of the proposed method result; the ‘...’ denote the difference between the undefined area of LPF result and that of the proposed method result.

Table 1 The undefined area per frame (*MUAPF*) of the stable image sequence.

Testing sequences	DIS methods		
	KF	LPF	Proposed method
IS1	3181.8	3021.7	2916.2
IS2	7459.7	7291.2	6034.5
IS3	8801.5	6881.6	5378.9
IS4	5350.8	5297.2	4989.0

only has the same direction as the BRF jitter but also has bigger magnitude than half of the BRF jitter. However, the image sequences that we focus on have sharp BRF jitters, in which case the BRF jitters have much more big magnitudes than the unwanted jitters of most frames. Thus, the DUAs for most frames are positive. In addition, Table 1 gives the mean undefined area per frame (*MUAPF*) of the stable image sequence which is calculated as Eq. (12):

$$MUAPF = \frac{1}{N} \sum_{i=1}^N UA(i) \quad (12)$$

where N is the frame number of the stable image sequence. From above results, we can conclude that the proposed method can reduce the impact of the BRF jitter on digital image stabilization.

4. Conclusions

In this letter, we propose a digital image stabilization method based on correction for BRF jitter. This method estimates the unwanted camera motion vector of the BRF (i.e., the first reference frame) first of all and then corrects the CMVs to eliminate the impact of the BRF jitter on the stable image sequence. The experimental results show that the proposed DIS method can efficiently reduce the excessive undefined area resulting from the BRF jitter and works well for two major kinds of image sequences. Moreover, the relationship between the CMVs and the BRF jitter is derived only based on one popular MC method: frame position smooth. In the further study, we will consider another popular MC method: global motion smooth.

References

- [1] S. Erturk, “Digital image stabilization with sub-image phase correlation based global motion estimation,” *IEEE Trans. Consum. Electron.*, vol.49, no.4, pp.1320–1325, 2003.
- [2] S.-C. Hsu, S.-F. Liang, K.-W. Fan, and C.-T. Lin, “A robust in-car digital image stabilization technique,” *IEEE Trans. Syst., Man Cybern. C, Appl. Rev.*, vol.37, no.2, pp.234–247, 2007.
- [3] S.-J. Ko, S.-H. Lee, S.-W. Jeon, and E.-S. Kang, “Fast digital image stabilizer based on gray-coded bit-plane matching,” *IEEE Trans. Consum. Electron.*, vol.45, no.3, pp.598–603, 1999.
- [4] C.-Y. Chung and H.H. Chen, “Feature-based full-frame image stabilization,” pp.100–106, *Inst. of Elec. and Elec. Eng. Computer Society, Taichung, Taiwan*, 2007.
- [5] L. Xu and X. Lin, “Digital image stabilization based on circular block matching,” *IEEE Trans. Consum. Electron.*, vol.52, no.2, pp.566–574, 2006.
- [6] C. Wang, J.-H. Kim, K.-Y. Byun, J. Ni, and S.-J. Ko, “Robust digital image stabilization using the Kalman filter,” *IEEE Trans. Consum. Electron.*, vol.55, no.1, pp.6–14, 2009.
- [7] S. Erturk, “Image sequence stabilisation: Motion vector integration (mvi) versus frame position smoothing (fps),” *Image and Signal Processing and Analysis*, 2001. ISPA2001. Proc. 2nd International Symposium, pp.266–271, 2001.
- [8] S. Erturk and T.J. Dennis, “Image sequence stabilisation based on DFT filtering,” *IEE Proc. Vision, Image and Signal Processing*, vol.147, no.2, pp.95–102, 2000.
- [9] S.H. Yang, F.M. Jheng, and Y.C. Cheng, “Two-dimensional adaptive image stabilisation,” *Electron. Lett.*, vol.43, no.8, pp.446–447, 2007.