

RESEARCH

Open Access

Digital video stabilizer by adaptive fuzzy filtering

Mohammad J Tanakian, Mehdi Rezaei* and Farahnaz Mohanna

Abstract

Digital video stabilization (DVS) allows acquiring video sequences without disturbing jerkiness, removing unwanted camera movements. A good DVS should remove the unwanted camera movements while maintains the intentional camera movements. In this article, we propose a novel DVS algorithm that compensates the camera jitters applying an adaptive fuzzy filter on the global motion of video frames. The adaptive fuzzy filter is a simple infinite impulse response filter which is tuned by a fuzzy system adaptively to the camera motion characteristics. The fuzzy system is also tuned during operation according to the amount of camera jitters. The fuzzy system uses two inputs which are quantitative representations of the unwanted and the intentional camera movements. The global motion of video frames is estimated based on the block motion vectors which resulted by video encoder during motion estimation operation. Experimental results indicate a good performance for the proposed algorithm.

Keywords: Adaptive, Digital video stabilizer, Motion estimation, Fuzzy filter, Motion vector, Video coding, Video stabilization

1. Introduction

Digital video stabilization (DVS) techniques have been studied for decades to improve visual quality of image sequences captured by compact and lightweight digital video cameras. When such cameras are hand held or mounted on unstable platforms, the captured video generally looks shaky because of undesired camera motions. Unwanted video vibrations would lead to degraded view experience and also greatly affect the performances of applications such as video encoding [1-4] and video surveillance [5,6]. With recent advances in wireless technology, video stabilization systems are also considered for integration into wireless video communication equipments for the stabilization of acquired sequences before transmission, not only to improve visual quality but also to increase the compression performance [1]. Solutions to the stabilization problem involve either hardware or software to compensate the unwanted camera motion. The hardware-based stabilizers are generally expensive and lack the kind of compactness that is crucial for today's consumer electronic devices [7,8]. On the contrary, a DVS system that is implemented by software can easily be miniaturized and updated. Consequently, DVS

system is suitable for portable digital devices, such as digital camera and mobile phone.

In general, a DVS system consists of two principal units including motion estimation (ME) and motion correction (MC) units. The ME unit estimates a global motion vector (GMV) between every two consecutive frames of the video sequence. Using the GMVs, the MC unit then generates smoothing motion vectors (SMVs) needed to compensate the frame jitters and warp the frames to create a more visual stable image sequence.

According to the motion models being considered, the already proposed global ME techniques for DVS system can roughly be divided into two categories: (1) two-dimensional stabilization techniques which deal with translational jitter only [9-20] and (2) multi-dimensional stabilization techniques which aim at stabilizing more complicated fluctuations in addition to translation [21-25]. Most of the existing algorithms fall into the first category because the translation is the most commonly encountered motion and the complexity of estimating translation parameters is relatively low for real-time stabilization.

Regarding to the ME task of DVS systems, most previous approaches attempt to reduce the computational cost by using fast ME algorithms, e.g., gray-coded bit-plane matching [9], two-bit transform [10], multiplication-free one-bit transform [11], Laplacian two-bit transform [12], and binary image matching of color weight [13]. In

* Correspondence: mehdi.rezaei@ece.usb.ac.ir

Faculty of Electrical & Computer Engineering, University of Sistan and Baluchestan, Zahedan, Iran

another approach, the global ME is limited to small, pre-defined regions [16,17]. Such approaches consider DVS and video encoding separately and attempt to trade the accuracy of motion vectors (MVs) for the computational efficiency; nevertheless they improve the computational efficiency at the expense of degradation in the accuracy in ME and thereafter in MC tasks.

Since both the video encoder and the digital stabilizer of a digital video camera use a ME unit, we can integrate digital stabilizer with video encoder [2,4,26] by making the two modules of a digital video camera share a common local motion vectors (LMVs) estimation process, as shown in Figure 1. The ME task in video encoders usually is implemented on frame blocks by a block matching process to estimate anMV for each block (BMV).

In video frames with smooth or complex texture regions, the estimated BMVs may not be in coincidence with the real motion of the blocks. Although such LMVs are applicable to the local motion compensation task which is executed in the encoder, they cannot be used for the global motion compensation which is executed by the DVS. These LMVs include some noises that degrade the global ME task. In order to remove the noisy LMVs in these regions some algorithms are proposed in [27-30]. The valid BMVs as LMVs are used for the global ME and MC compensation in next steps.

After global ME, the next essential task of a DVS system is MC in which the unwanted camera jitters are separated and removed from the intentional camera movement. Among the various MC algorithms proposed in the literature, smoothing of the GMV by low-pass filtering is the most popular. For instance, anMV integration method is used in [9,31] which utilizes a first-order infinite impulse response (IIR) low-pass filter to integrate differential motion and to smoothen the global movement trajectory. A frame position smoothing (FPS) algorithm, based on smoothing absolute frame positions that achieve successful stabilization performance with retained smooth camera movements, is utilized for MC in [17,32-39]. Off-line discrete Fourier transform (DFT)

domain filtering is proposed for FPS-based stabilization in [32]. Kalman filter and fuzzy systems have widely been used in DVS applications [33-39]. Real-time FPS-based stabilizer using Kalman filtering of absolute frame positions has been proposed in [17,33]. It is shown that the stabilization performance can be improved by a fuzzy adaptive Kalman filter; introducing a stabilization system that is adjusted according to the camera motion characteristics in [34]. Fuzzy stabilization systems improve the stabilization performance when their membership functions (MFs) are optimized to motion dynamics [35]. A membership selective fuzzy stabilization, in which the stabilization system selects between a pre-determined set of MFs according to instantaneous motion characteristics is proposed in [36]. A MF adaptive fuzzy filter for video stabilization is presented in [37] and a fuzzy Kalman system consists of a fuzzy system with a Kalman filter is presented in [38].

Regarding to the MC task of DVS system, almost all published algorithms try to smoothen the global movement trajectory by a kind of low-pass filtering. An important drawback of the low-pass filtering is that smoothened movement trajectory is delayed with respect to the desired camera displacements. A stricter filtering provides more stabilization at the expense of more trajectory delay and vice versa. More trajectory delay means losing more image content after stabilization.

A good MC unit should remove the unwanted camera motion while tracks the intentional motion without any delay. For this purpose, it should discriminate the unwanted and intentional camera motions while adjust the smoothing filter adaptively according to the amount of unwanted and intentional camera motions. The studied published MC algorithms lack some of these features. For example, algorithms presented in [27,37] suffer from the lack of discrimination of unwanted and intentional camera motions. Moreover, the proposed adaptive algorithm in [27] suffers from a continuous and well adaptation. They use an adaptive filter with a smoothing factor that is switched between only two values and therefore it leads to undesirable jumps in frame position. The proposed algorithm in [38] shows a high performance but still suffers from well adaptation.

In this article, we propose a DVS algorithm with new features in ME and MC units. The ME unit estimates a GMV based on the BMVs which are estimated by the video encoder. Therefore, accurate motion information is used without extra computation cost. Moreover, an adaptive thresholding algorithm is used to remove the noisy invalid LMVs. The MC unit of the proposed DVS system applied a fuzzy adaptive IIR filter to smooth the camera movement trajectory adaptively to the characteristics of unwanted and intentional camera motions. The fuzzy system adjusts the IIR filter by using two novel

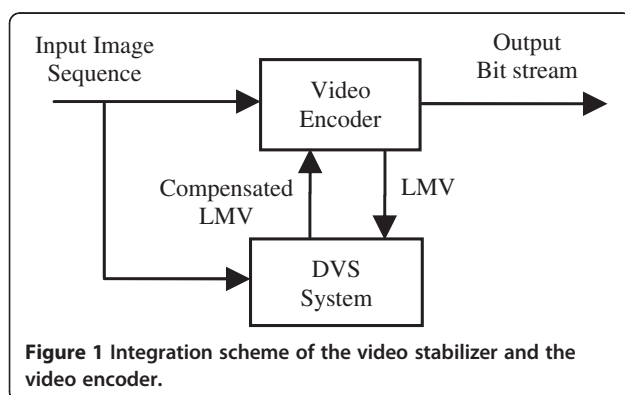


Figure 1 Integration scheme of the video stabilizer and the video encoder.

inputs which are quantitative representations of the unwanted and the intentional camera motions. Experimental results show a good performance for the proposed DVS algorithm.

The remainder of this article is organized as follows. The details of the proposed video stabilization algorithm are described in Section 2. Some experimental results are presented in Section 3, and the article is concluded in Section 4.

2. The proposed method

A flowchart of the proposed DVS system is depicted in Figure 2. The details of the proposed system are described in the sequel.

2.1. Block-based ME

The block-based ME is used to generate the LMVs. Since the ME is done by the video encoder, the computational complexity of the DVS is very low. In this article, to test the proposed DVS system independent of the encoder, a full search ME algorithm with full-pixel resolution is taken for 8×8 blocks over a search range of 33×33 pixel to achieve the BMVs.

The ME algorithm works as follows. First, the current frame is divided into a number of $N \times N$ blocks and an MV for each block is computed. The resulting MV points to the most correlated reference block in the previous frame within the search area. To measure the goodness of each candidate MV (x, y) , the mean absolute difference (MAD) measure is used as

$$MAD(i, j) = \frac{1}{N^2} \sum_{k=0}^{N-1} \sum_{L=0}^{N-1} |C(x+k, y+L) - R(x+i+k, y+j+L)| \quad (1)$$

where $C(x+k, y+L)$ and $R(x+i+k, y+j+L)$ denote the block pixels in the target frame and the displaced block pixels in the reference frame, respectively. The candidate MV (i, j) with the smallest MAD is chosen as the MV of the current block according to

$$V = \arg \min MAD(i, j), \quad -p \leq i, j \leq p-1 \quad (2)$$

where p defines the motion search range.

2.2. LMV validation

The ME unit plays an important role in DVS system and its estimation accuracy is a decisive factor for the overall performance of stabilization system. Block ME process typically computes some wrong MVs which are not in coincidence to the real motion direction of the blocks. Although, such MVs can be useful for the motion compensation in encoder, they include noise and should not

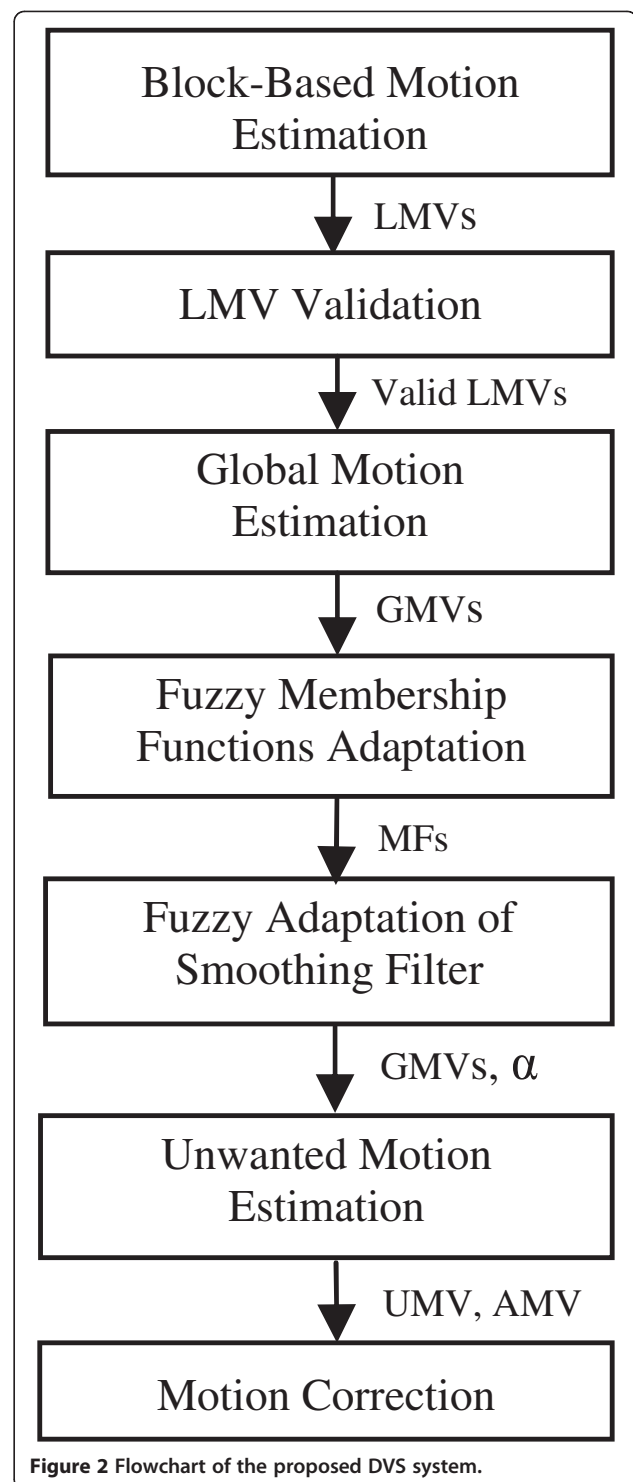


Figure 2 Flowchart of the proposed DVS system.

be used for the global motion compensation and video stabilization operations. The noisy MVs are mostly obtained from two types of regions including: very smooth regions with lack of features and very complex uneven regions [27-30]. Inspiring from the algorithm presented in [27], two qualifying tests, namely "Smoothness Test" and

“Complexity Test”, are used to detect and remove the noisy MVs by an adaptive thresholding method as follows.

2.2.1. Smoothness test

The noisy MVs corresponding to the smooth regions such as sky image are detected by thresholding of the average of MAD as

$$MAD_{avg}^n < th_1, \quad (3)$$

where MAD_{avg}^n denotes the average of calculated MADs within the search area, during ME of n th block. th_1 is also defined as

$$th_1 = MAD_{min}^n + T_1 \times \text{Mean}(MAD_{avg}^n), \quad (4)$$

where MAD_{min}^n and MAD_{Avg}^n denote the minimum and the average values of computed MADs, respectively, during ME of n th block within the search area. T_1 is an experimentally defined constant coefficient about 0.45 and $\text{Mean}(MAD_{avg}^n)$ denotes the average of MAD_{avg}^n over all blocks of the frame. In fact the threshold th_1 includes a global average value over the frame plus a margin.

2.2.2. Complexity test

The noisy MVs corresponding to the complex texture regions are identified by another thresholding as

$$MAD_{min}^n > th_2, \quad (5)$$

where threshold th_2 is defined adaptively as:

$$th_2 = T_2 \times \text{Max}(MAD_{min}^n), \quad (6)$$

where T_2 is an experimentally defined constant coefficient about 0.45, and $\text{Max}(MAD_{min}^n)$ denotes the maximum value of MAD_{min}^n , over all blocks of the frame. According to the equations above, the MAD_{min}^n is compared against a portion of its global maximum over a frame.

It is notable that MAD is computed during ME by encoder. Therefore, the smoothness test and complexity test have no additional computational complexity cost for the proposed DVS system.

A similar thresholding approach is presented in [27-30], in which fixed values for thresholds th_1 and th_2 are used. Our simulation results on different video contents show that using fixed thresholds for different video contents may cause a remarkable amount of invalid noisy LMVs remain or a notable amount of valid LMVs be removed. To solve this problem, the values of thresholds th_1 and th_2 are adjusted adaptively based on the video content for each frame. Note, if ME is executed by a fast search algorithm rather than full-search algorithm at the

encoder, the MADs calculated during ME are used for adaptation of thresholds th_1 and th_2 .

Original LMVs and validated LMVs for a sample frame are presented in Figure 3. This figure shows that many noisy LMVs have been removed by the LMV validation process.

2.3. Global ME

The global ME unit produces a unique GMV for each video frame, which represents the camera movement during the time interval of two frames. Since the LMVs obtained from the image background tend to be very similar in both magnitude and direction, we used a clustering process to classify the motion field into clusters corresponding to the background and foreground objects. The global motion induced by camera movement is determined by a clustering process that consists of the following steps.

- Step 1. Construct the histogram H of the valid LMVs. The value of $H(x, y)$ is incremented by one each time the $LMV(x, y)$ is encountered.
- Step 2. As long as the scene is not dominated by moving objects, the cluster corresponding to background blocks has the maximum votes in the clustering process. The position (x, y) of the largest cluster or histogram bin is considered as the GMV.

As an example, Figure 4 shows the largest histogram bin at coordinates (5, 12), yields the GMV.

2.4. Unwanted ME and correction

An estimated GMV may consist of two major components: an intentional motion component (e.g., corresponding to camera panning) and unintentional motion component (e.g., corresponding to handshake). A good MC algorithm should only remove the unwanted motion while maintain and track the intentional motion. Assuming that the unwanted motion is corresponding to the high-frequency components, the proposed algorithm uses a low-pass filter to remove the unwanted motion component. AnSMV is resulted by a low-pass filtering on the GMVs that resembles the intentional camera movement. An adaptive first-order IIR filter is applied as

$$SMV(n) = \alpha(n)SMV(n-1) + [1 - \alpha(n)]GMV(n). \quad (7)$$

where the index n indicates the frame number. The parameter α , ($0 \leq \alpha \leq 1$), can be regarded as the smoothing factor of the filter which is adjusted by the fuzzy system for each frame. A larger smoothing factor leads to a



(a)



(b)

Figure 3 Example of noisy LMV removal on a frame of the avenue sequence. (a) Original MVs. (b) Valid MVs.

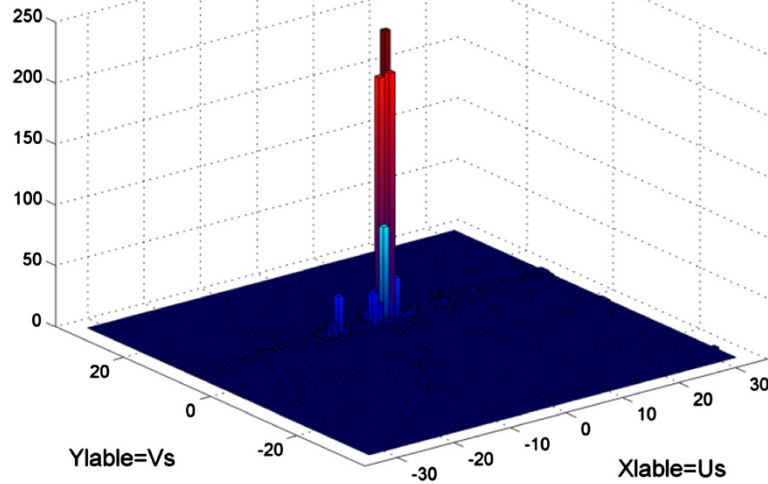


Figure 4 Clusters of motion field.

smoother, but a larger lag during intentional camera motion that makes artificially stabilized, image sequence. Therefore, a fixed value of α hardly leads to good stabilized image sequences. To avoid the lag of intentional movement and to smooth the unwanted camera motion efficiently the following fuzzy adaptation mechanism of α is proposed.

2.4.1. Fuzzy adaptation of smoothing filter

The smooth filtering is implemented on the vertical and horizontal components of the GMVs separately. The smoothing factor of filter, i.e., $\alpha(n)$ is adjusted by a fuzzy system continuously for MC of each frame. In facts, two fuzzy systems with a similar structure are used corresponding to the vertical and horizontal motion components. The fuzzy system has two inputs (Input1, Input2) and one output. The fuzzy inputs are defined as

$$x_1 = \frac{1}{M} \sum_{i=n-M+1}^n |GMV_x(i) - GMV_x(i-1)|, \quad (8)$$

$$x_2 = |GMV_x(n) - GMV_x(n-M)|. \quad (9)$$

$$y_1 = \frac{1}{M} \sum_{i=n-M+1}^n |GMV_y(i) - GMV_y(i-1)|, \quad (10)$$

$$y_2 = |GMV_y(n) - GMV_y(n-M)|. \quad (11)$$

where x_1 and x_2 denote the inputs of fuzzy system used for the adaptive filtering of the horizontal motion component and also y_1 and y_2 are the inputs of fuzzy system used for the adaptive filtering of the vertical motion component. $GMV_x(n)$ and $GMV_y(n)$ indicate the horizontal and vertical components of the GMV of last frame

and $M + 1$ is the number of last GMVs used for decision. The fuzzy system inputs, Input1 (x_1, y_1) and Input2 (x_2, y_2), are used as quantitative representations of unwanted and intentional camera movements, respectively. The value of Input1 is proportional to the noise amplitude and the value of Input2 is proportional to the intentional camera motion when it has an accelerating movement.

Defining suitable inputs for an adaptive DVS system has a great impact on the performance of system. Only relevant inputs can provide precise discriminating between unwanted and intentional camera motions to be used for the adaptation of smoothing filter. Different scenarios for the combination of unwanted and intentional camera motion can be considered. As examples, some scenarios are presented graphically in Figure 5. In graphs (a) and (b), camera has an intentional accelerating movement plus noise or unwanted motion. The noise amplitude is high in (a) while it can be ignored in (b). Graph (e) is corresponding to a camera movement path while panning in which the camera is moving with a constant velocity without any acceleration and noise. The explanations of all graphs are summarized in Table 1.

From the adaptive filtering point of view it is important to measure the amount of noise and the intentional camera movement velocity and acceleration. A stricter smoothing filter is needed when the noise amplitude is high to remove the noise. On the other hand, the strict smoothing filter prevents following of camera path when it has an intentional high acceleration. Therefore, the smoothing factor of filter should be tuned carefully proportional to the amount of noise and camera movement acceleration. According to this, we defined the fuzzy

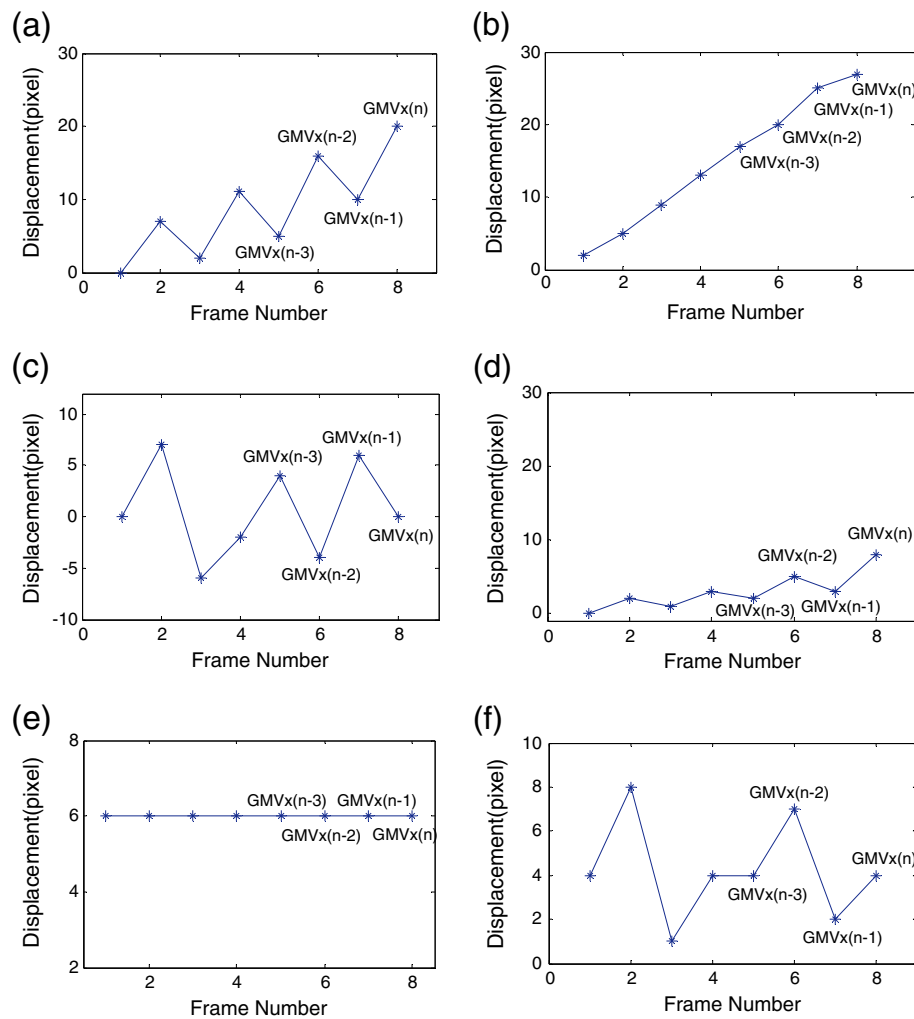


Figure 5 Sample scenarios for combination of unwanted and intentional camera motions: (a) high acceleration with high noise, (b) high acceleration with no noise, (c) high noise with no acceleration, (d) low acceleration with low noise, (e) constant velocity without noise and acceleration, (f) constant velocity with noise.

inputs so that Input1 gives information about the amount of noise and Input2 gives information about the amount of camera movement acceleration. It is notable that amount of camera movement velocity itself does not have any constrain on the filtering so it is not measured and used here.

Table 1 Sample scenarios for combination of unwanted and intentional camera motion

| Graph | Noise | Velocity | Acceleration |
|-------|-------|----------|--------------|
| a | High | High | High |
| b | Low | High | High |
| c | High | Zero | Zero |
| d | Low | Low | Low |
| e | Zero | High | Zero |
| F | High | High | Zero |

The proposed fuzzy system tunes the smoothing factor of the IIR filter adaptively according to the amount of noise and the camera intentional accelerating movement. In the proposed fuzzy system, trapezoidal and triangular MFs are used for the inputs and the outputs, respectively. The number of MFs has been selected so as to obtain decent performance with as few MFs as possible to maintain low system complexity. The experimentally designed input and output MFs and also the surface of desired outputs are shown in Figure 6. According to experimental results, the performance of used IIR filter is more sensitive to α 's changes where α has a large value. Therefore, more MFs of the fuzzy output are concentrated in this operating area. The constructed rule base is containing 30 rules as presented in Table 2. The proposed fuzzy system was implemented while the *min* function was used for the *fuzzy implication* and the *max*

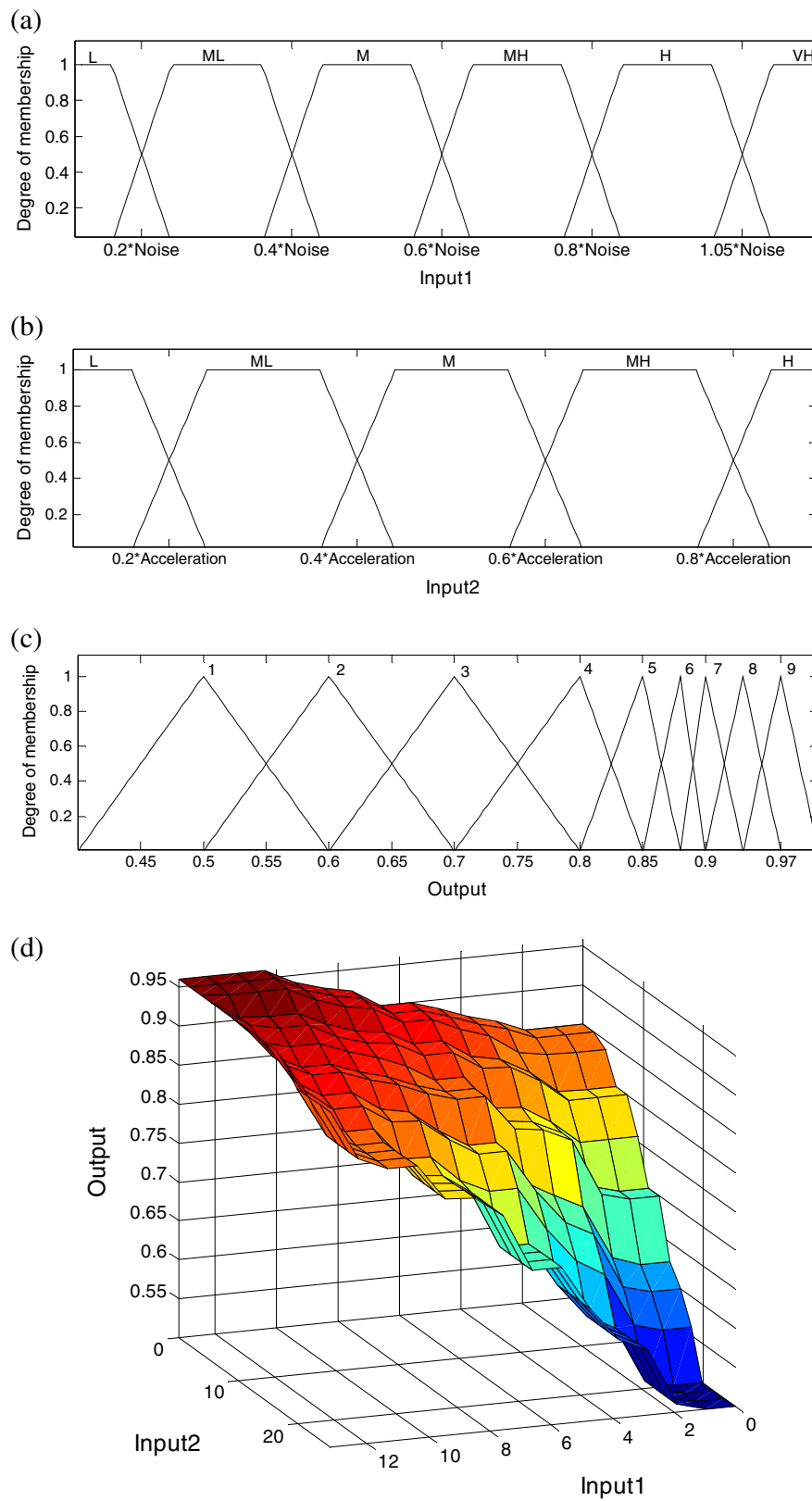


Figure 6 The experimentally designed inputs and output MFs and also surface of desired input and output of the proposed fuzzy system. (a) MFs of fuzzy Input1 x_1 & y_1 , (b) MFs of fuzzy Input2 x_2 & y_2 , (c) MFs of fuzzy output, (d) surface of desired outputs.

Table 2 Central values of fuzzy system output

| Input2 | Input1 | | | | | |
|--------|--------|------|------|------|------|------|
| | L | ML | M | MH | H | VH |
| L | 0.85 | 0.87 | 0.9 | 0.94 | 0.97 | 0.97 |
| ML | 0.8 | 0.85 | 0.87 | 0.9 | 0.94 | 0.97 |
| M | 0.7 | 0.8 | 0.85 | 0.87 | 0.9 | 0.97 |
| MH | 0.6 | 0.7 | 0.8 | 0.85 | 0.87 | 0.97 |
| H | 0.5 | 0.6 | 0.7 | 0.8 | 0.85 | 0.94 |

L, low; ML, medium low; M, medium; MH, medium high; H, high; VH, very high.

function used for the *fuzzy aggregation*. Furthermore, the *centroid defuzzification* method was applied. The output of fuzzy system defines the smoothing factor of the IIR filter, i.e., $\alpha(n)$, for MC of each video frame.

2.4.2. Adaptive fuzzy MFs

Study on a number of video sequences has shown that the range of fuzzy inputs (Input1, Input2) is very variable on different video contents. Therefore, fixed MFs for the inputs of fuzzy system cannot provide a good stabilization performance over all video contents. In order to have a good performance for the proposed DVS system over different video contents, it is proposed to adjust the MFs of fuzzy inputs adaptively to recently

received video frames. The range of MFs for the fuzzy inputs, i.e., (0, Input1(max)) and (0, Input2(max)) are modified adaptively as

$$\text{Input1(max)} = \text{Max of input 1 over } K \text{ recent frames,} \quad (12)$$

$$\text{Input2(max)} = \text{Max of input 2 over } K \text{ recent frames,} \quad (13)$$

where Input1 and Input2 are clipped to a range from 1 to 10% of video frame height in term of pixel, and the K corresponds to the number of frames received in last few seconds, e.g., 2 s. This means that the system is adapted to the time-varying noise conditions while the frame size and frame rate are considered.

2.4.3. MC

After computing the smoothing factor $\alpha(n)$ by the fuzzy system, SMV is calculated by Equation (7). For the first three frames, a fixed large value for $\alpha(n)$ is used. After computing SMV, the unwanted motion vector (UMV) is obtained by

$$\text{UMV}(n) = \text{GMV}(n) - \text{SMV}(n). \quad (14)$$

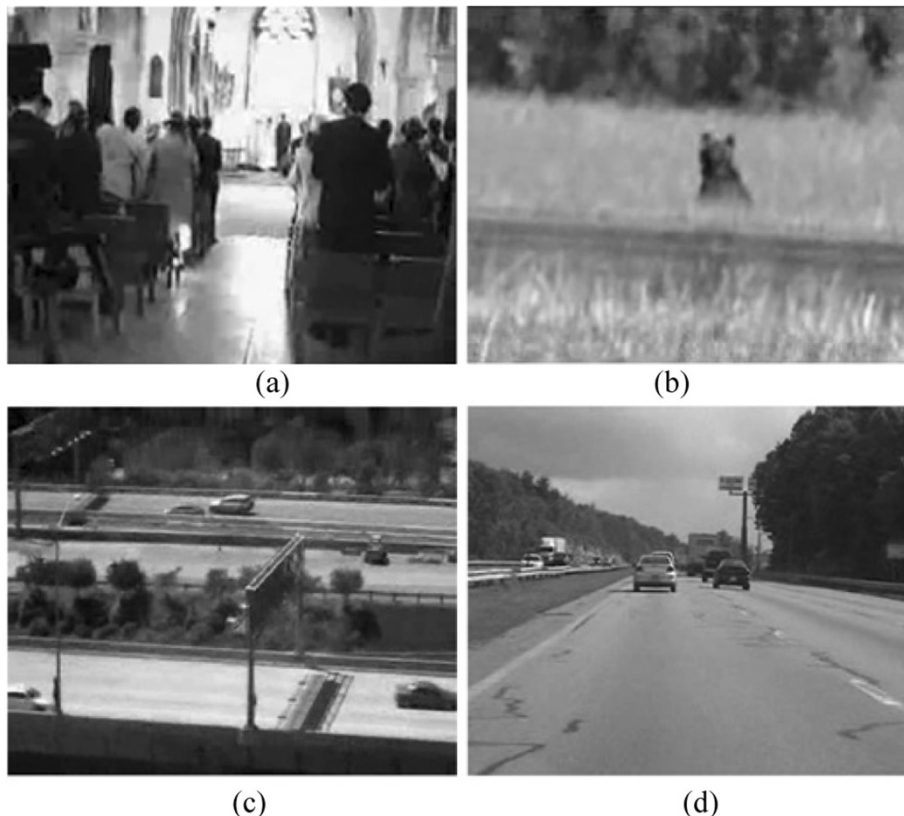


Figure 7 Sample frames of used video sequences. Images taken by a camera (a–c) held by a hand; (d) in a moving vehicle.

To restore the current frame to its stabilized position, we offset the current frame by the accumulated UMV, AMV, defined by

$$AMV(n) = \sum_{i=m}^n UMV(i) \quad (15)$$

where m is the frame number of the last scene cut frame.

3. Experimental results

The performance of the proposed DVS method is evaluated against 15 video sequences covering different types of scenes.

Since there is no well-known video sequence in this research field, the algorithm is tested on a number of sequences which are easily available. For example, some used video sequences are available at [40,41]. These sequences have a frame rate of 25 fps and a picture size

of 352×288 pixels. Sample frames of used video sequences are shown in Figure 7. Moreover, the performance of proposed DVS system has been evaluated with several datasets extracted from movement curves published in the literature, not from ME on video sequences. Sample simulation results are presented in Figures 8 and 9. The results presented in Figure 8 correspond to a case in which the camera has no intentional movement, whereas the results presented in Figure 9 correspond to a case in which the camera has intentional camera movements. We worked with both gray-scale and color test sequences where in both cases ME is implemented on the luminance component. Good experimental results are obtained with $M = 3$. However, a larger M provides more smoothness at the expense of more tracking delay and vice versa. The stabilizer performance is assessed according to the smoothness of the resultant global motion compared to the original sequence and the gross movement preservation capability. The results of the

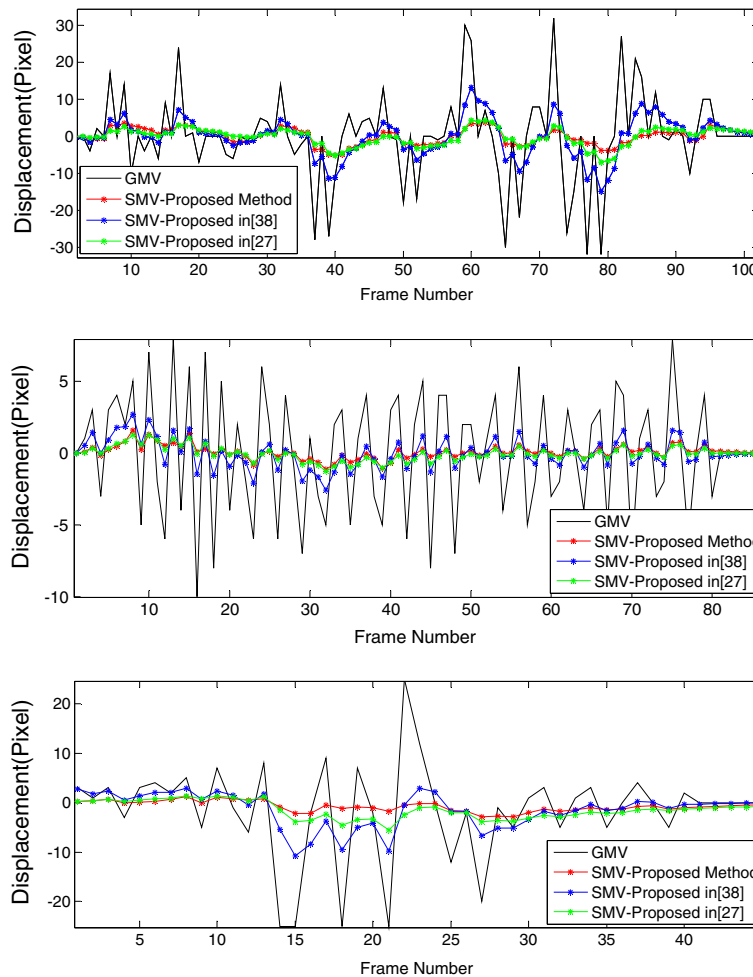


Figure 8 Comparison results of DVS algorithms: the absolute frame positions before and after stabilization correspond to a case in which the camera has no intentional movement.

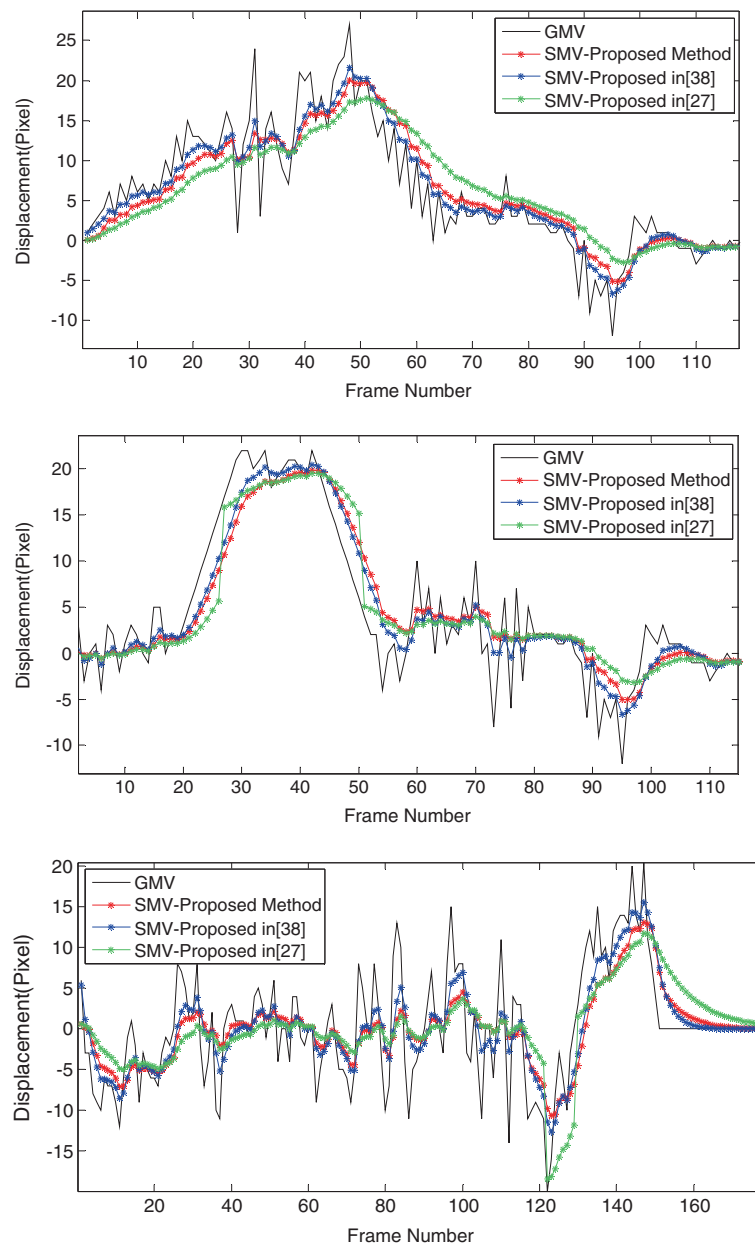


Figure 9 Comparison results of DVS algorithms: the absolute frame positions before and after stabilization correspond to a case in which the camera has intentional movement.

proposed DVS algorithm are compared with results provided by the presented algorithms in [27,38], as the most relevant anchor algorithms. An adaptive IIR filtering technique and a fuzzy Kalman system are proposed for MC in [27,38], respectively. Some graphical comparison results are presented Figures 8 and 9. The results provided by the presented algorithm in [27] show that it leads to smooth camera movement trajectory but at expense of a relative large tracking delay when the camera has intentional accelerating movement. The results provided by the presented algorithm in [38] show that it

closely tracks the intentional camera movements but at the cost of slightly reduced stabilization capabilities. Whereas results demonstrate that our proposed DVS system provides expanded stabilization, while enables the close tracking of the intentional camera movements. Small-scale subjective quality test also demonstrated that human eyes have better visual perception to the stabilized videos by the proposed DVS system than the original videos in all cases.

One way of visualizing the effectiveness of image stabilization is to subtract consecutive frames of the

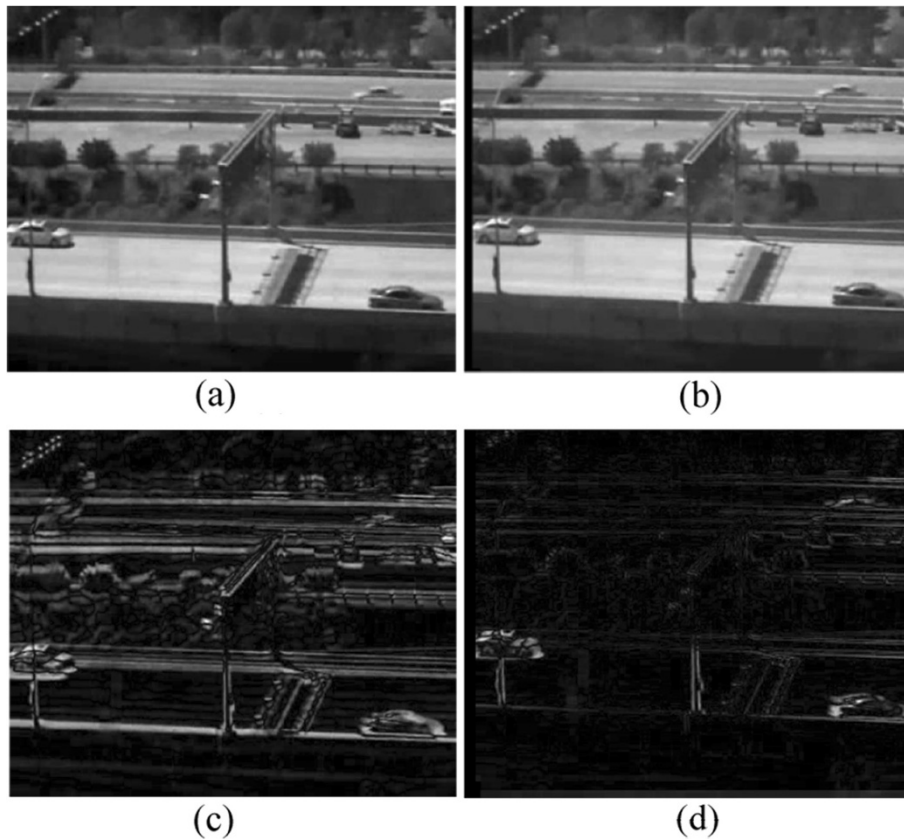


Figure 10 Visualizing the effectiveness of image stabilization. (a) Original frame, (b) stabilized frame, (c, d) difference between two frames before and after MC, respectively.

original and stabilized image sequences. Figure 10 shows an example of this technique. The upper left image in (a) shows a frame of the original sequence while the upper right image in (b) is the stabilized frame at the same time instance. The bottom left image in (c) shows the difference between the frame in (a) and its previous frame from original sequence. Similarly, Figure 10 shows the difference between the stabilized frame shown in

(b) and its previous frame in the stabilized sequence. The decrease in luminance of the difference signal means removing unwanted camera motion. Expressing the difference of consecutive frames in the original and the stabilized image sequences in terms of squared error, for a video sequence, namely road [40], is plotted in Figure 11. As expected, the frame difference measure has been reduced considerably in the stabilized image sequence.

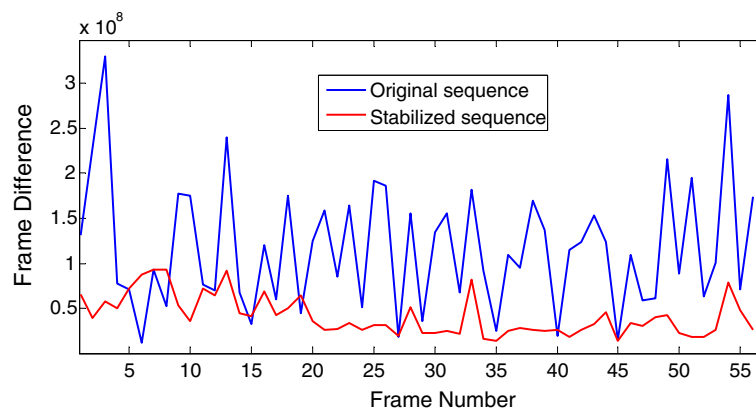


Figure 11 Frame difference for the road video sequence before and after stabilization.

Table 3 MSE resulted by MC units of different DVSS

| Sequences | Adaptive fuzzy filter (proposed method) | Kalman filter [38] | IIR-filter [27] |
|-----------|---|--------------------|-----------------|
| 1 | 4.6887 | 27.5002 | 5.4913 |
| 2 | 0.1804 | 1.1427 | 0.2219 |
| 3 | 1.901 | 16.2712 | 5.6552 |

Numerical performance assessment of a DVS system is a difficult task since the ground truth of unwanted motions as reference is not available. To solve the lack of reference problem, we choose some video sequences in which the camera has been fixed without any intentional movement. Therefore, the reference GMVs of these sequences are zero. Samples of such video sequences are shown in Figure 8. Finding a reference path for intentional camera motion, we computed the meansquare error (MSE) between the smoothed GMVs and the GMVs of reference as a numerical measure to compare the MC performance of proposed adaptive fuzzy DVS algorithms with those of presented algorithms in [27,38]. The MSE measure is defined as

$$MSE_{MV} = \frac{1}{N} \sum_{n=0}^N [GMV_R(n) - SMV(n)]^2, \quad (16)$$

where $GMV_R(n)$ and $SMV(n)$ denote the n th GMV of the reference and the smoothed video sequence, respectively. Numerical performance assessment results for the three video sequences, as shown in Figure 8, are presented in Table 3. The numerical results show lower MSE values for our proposed method that means a higher performance in this case.

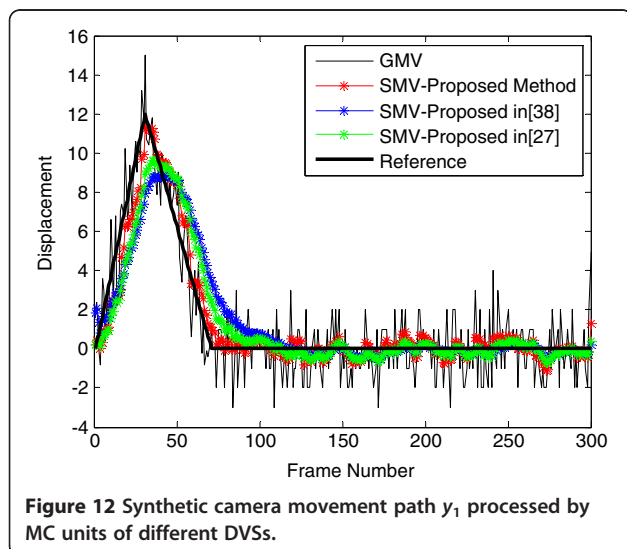


Figure 12 Synthetic camera movement path y_1 processed by MC units of different DVSSs.

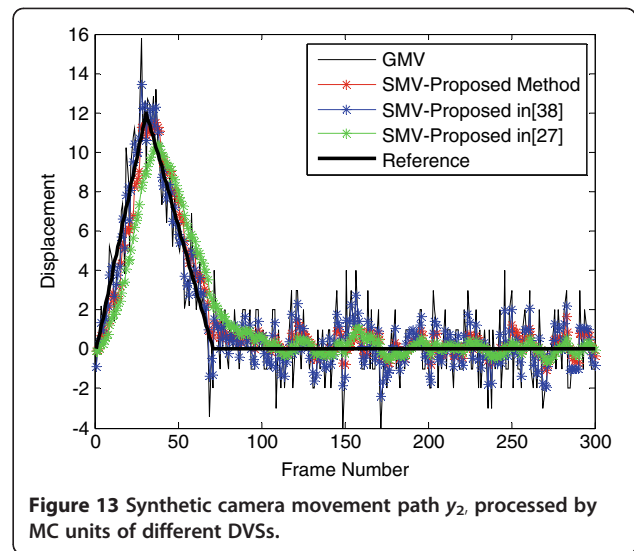


Figure 13 Synthetic camera movement path y_2 processed by MC units of different DVSSs.

To evaluate the performance of the proposed algorithm numerically in case of intentional camera movement, we produced a synthetic signal as reference. The reference signal includes three parts corresponding to intentional camera movements with positive, negative, and zero value accelerations. Moreover, a sequence of normal random numbers was generated and the random numbers were rounded to integer values to simulate the unwanted camera motion in terms of pixel displacement. Provided sequence was added to the reference signal to obtain a synthetic signal including both the intentional and the unwanted camera motions. Two copies of the synthetic signal (y_1, y_2) were generated to be processed by MC units of the proposed DVS algorithm and also by the anchor algorithms presented in [27,38]. While the reference signal is known that the performance of compared algorithms can be evaluated numerically by computing the MSE measure between the reference signal and the processed signals. It is noted that the performance of the proposed algorithm in [38] depends on the values of Q parameter (*process noise covariance*) of Kalman filter, so it cannot provide a well adaptation. Therefore, the two copies of synthetic signal were processed by two values of Q (0.005 and 0.6) and graphical simulation results are shown in Figures 12 and 13, respectively. Moreover, numerical comparison results in terms of MSE for the synthetic signals are presented in Table 4. According to the graphical results, the proposed algorithm in [27]

Table 4 Resulted by MC units of different DVSS

| Sequences | Adaptive fuzzy filter (proposed method) | Kalman filter [38] | IIR-filter [27] |
|-----------|---|--------------------|-----------------|
| y_1 | 0.5886 | 0.9369 | 1.3577 |
| y_2 | 0.6561 | 2.0282 | 1.4728 |

removes the unwanted camera motions in case of no intentional camera movement. It results a considerable tracking delay in case of intentional accelerating camera movement. The proposed algorithm in [38], depending on the Q value, performs only good filtering of the unwanted camera motion or only the tracking of intentional camera movement as shown in Figures 12 and 13, respectively. The proposed algorithm in this article shows a high performance in both the removing of unwanted motion and the tracking of intentional camera movement. Furthermore, the numerical results presented in Table 4 confirm the graphical results shown in Figures 12 and 13. The least MSE for our proposed algorithm means the best performance.

4. Conclusion

In this article, we proposed a computationally efficient DVS algorithm using motion information obtained from a hybrid block-based video encoder. Since some of the obtained MVs are not valid, an adaptive thresholding was developed to filter out valid MVs and to compute an accurate GMV for each frame. The GMVs are smoothened with an IIR low-pass filter that is tuned adaptively to unwanted and intentional camera movements. The filter is adjusted by a fuzzy system with two inputs which quantify the unwanted and intentional camera movements. The proposed method fulfills two apparently conflicting requirements: close follow-up of the intentional camera movement and removal of the unwanted camera motion. In order to improve the stabilization performance, inputs MFs of the fuzzy system are continuously adapted according to the motion properties of a number of recently received video frames. Simulation results show a high performance for the proposed algorithm. With a low degree of computational complexity, the proposed scheme can effectively be used for the mobile video communications as well as for the conventional video coding applications to improve the visual quality of digital video and to provide a higher compression performance.

Competing interests

The authors declare that they have no competing interests.

Received: 8 November 2011 Accepted: 6 November 2012

Published: 7 December 2012

References

1. A Engelsberg, G Schmidt, A comparative review of digital image stabilizing algorithms for mobile video communications. *IEEE Trans. Consum. Electron.* **45**(3), 592–597 (1999)
2. YC Peng, CK Liang, HA Chang, HH Chen, CJ Kao, Integration of image stabilizer with video codec for digital video cameras, in *Proceedings of the International Symposium on Circuits and Systems*, 5th edn. (Kobe, Japan, 2005), pp. 4781–4784
3. CK Liang, YC Peng, HA Chang, CC Su, H Chen, The effect of digital image stabilization on coding performance, in *IEEE Proceedings of the 2004 International Symposium on Intelligent Multimedia, Video and Speech Processing*, 2004, pp. 402–405. doi:10.1109/SIMP.2004.1434085
4. HH Chen, CK Liang, YC Peng, HA Chang, Integration of image stabilizer with video codec for digital video cameras. *IEEE Trans. Video Technol.* **17**, 801–813 (2007)
5. L Marcenaro, G Vernazza, CS Regazzoni, Image stabilization algorithms for video surveillance applications, in *Proceedings of the International Conference on Image Processing*, 1st edn. (Thessaloniki, 2001), pp. 349–352
6. J Zhou, H He, D Wan, Video stabilization and completion using two cameras. *IEEE Trans. Circuit Syst. Video Technol.* **99**, 1 (2011)
7. M Oshima, T Hayashi, S Fujioka, T Inaji, H Mitani, J Kajino, K Ikeda, K Komoda, VHS camcorder with electronic image stabilizer. *IEEE Trans. Consum. Electron.* **35**(4), 749–758 (1989)
8. K Sato, S Ishizuka, A Nikami, M Sato, Control techniques for optical image stabilizing system. *IEEE Trans. Consum. Electron.* **39**(3), 461–466 (1993)
9. SJ Ko, SH Lee, SW Jeon, ES Kang, Fast digital image stabilizer based on gray-coded bit-plane matching. *IEEE Trans. Consum. Electron.* **45**(3), 598–603 (1999)
10. A Ertürk, S Ertürk, Two-bit transform for binary block motion estimation. *IEEE Trans. Circuit Syst. Video Technol.* **15**(7), 938–946 (2005)
11. S Ertürk, Multiplication-free one-bit transform for low complexity block-based motion estimation. *IEEE Signal Process. Lett.* **14**(2), 109–112 (2007)
12. NJ Kim, HJ Lee, JB Lee, Probabilistic global motion estimation based on Laplacian two-bit plane matching for fast digital image stabilization. *EURASIP J. Adv. Signal Process.* **43** (2008). doi:10.1155/2008/180582
13. W Nan, H Xiaowei, W Gang, Y Zhonghu, An approach of electronic stabilization based on binary image matching of color weight, in *2nd International Asia Conference on Informatics in Control, Automation and Robotics (CAR)*, 2nd edn. (Wuhan, 2010), pp. 155–158
14. S Battiato, G Puglisi, Fast block based local motion estimation for video stabilization, in *IEEE Computer Society Conference on Pattern Recognition Workshops (CVPRW)* (Colorado Springs, CO, 2011), pp. 50–57
15. JK Paik, YC Park, DW Kim, An adaptive motion decision system for digital image stabilizer based on edge pattern matching. *IEEE Trans. Consum. Electron.* **38**(3), 607–616 (1992)
16. AA Yeni, S Ertürk, Fast digital image stabilization using one bit transform based sub-image motion estimation. *IEEE Trans. Consum. Electron.* **51**(3), 917–921 (2005)
17. S Ertürk, Digital image stabilization with sub-image phase correlation based global motion estimation. *IEEE Trans. Consum. Electron.* **49**(4), 1320–1325 (2003)
18. S Ertürk, Image sequence stabilization: motion vector integration (MVI) versus frame position smoothing (FPS), in *Proceedings of the 2nd IEEE R8-EURASIP Symposium on Image and Signal Processing and Analysis, ISPA'01* (Croatia, Pula, 2001), pp. 266–271
19. ZY Xiang, W Jian, ZW Gong, Z Quan, D Rui, An improved algorithm of electronic image stability based on block matching, in *IEEE 5th Conference on Industrial Electronics and Applications (ICIEA)* (Taichung, 2010), pp. 1924–1927
20. J Zhu, B Guo, Fast layered bit-plane matching for electronic video stabilization, in *International Conference on Multimedia and Signal Processing (CMSP)*, 1st edn. (Guilin, Guangxi, 2011), pp. 276–280
21. JY Chang, WF Hu, MH Cheng, BS Shang, Digital image translation and rotation motion stabilization using optical flow technique. *IEEE Trans. Consum. Electron.* **48**(1), 108–115 (2002)
22. S Erturk, Translation, rotation and scale stabilization of image sequences. *IEE Electron. Lett.* **39**(17), 1245–1246 (2003)
23. NA Tsoligkas, S Xalkiadis, X Donglai, I French, A guide to digital image stabilization procedure—an overview, in *18th International Conference on Systems Signals and Image Processing (IWSSIP)* (Sarajevo, 2011), pp. 1–4
24. JM Wang, HP Chou, SW Chen, CS Fuh, Video stabilization for a hand-held camera based on 3D Motion model, in *16th IEEE International Conference on Image Processing (ICIP)* (Cairo, 2009), pp. 3477–3480
25. O Nестares, Y Gat, H Haussecker, I Kozinsev, Video stabilization to a global 3D frame of reference byfusing orientation sensor and image alignment data, in *9th IEEE International symposium on Mixed and Augmented Reality (ISMAR)* (Seoul, 2010), pp. 257–258
26. M Mohammadi, M Fathi, M Soryani, A new decoder side video stabilization using particle filter, in *18th International Conference on Systems signals and Image Processing (IWSSIP)* (Sarajevo, 2011), pp. 1–4

27. SH Yang, FM Jheng, An adaptive image stabilization technique, in *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics(SMC2006)*, 3rd edn. (Taipei, 2006), pp. 1968–1973
28. F Vella, A Castorina, M Mancuso, G Messina, Digital image stabilization by adaptive block motion vectors filtering. *IEEE Trans. Consum. Electron* **48**, 796–801 (2002)
29. S Battiato, G Puglisi, AR Bruna, A robust video stabilization system by adaptive motion vectors filtering, in *IEEE International Conference on Multimedia and Expo* (Hannover, 2008), pp. 373–376
30. JP Hsiao, CC Hsu, TC Shih, PL Hsu, SS Yeh, BC Wang, The real-time video stabilization for the rescue robot. 4364–4369 (2009)
31. K Uomori, A Morimura, H Ishii, Electronic image stabilization system for video cameras and VCRs. *J. Soc. Motion Picture Television Eng.* **101**, 66–75 (1992)
32. S Ertürk, TJ Dennis, Image sequence stabilization based on DFT filtering. *IEE Proc. Image Vis. Signal Process* **127**, 95–102 (2000)
33. S Ertürk, Real-time digital image stabilization using Kalman filters. *Real-Time Imaging* **8**, 317–328 (2002)
34. MK Güllü, E Yaman, S Ertürk, Image sequence stabilization using fuzzy adaptive Kalman filtering. *Electron. Lett.* **39**(5), 429–431 (2003)
35. MK Güllü, S Ertürk, Fuzzy image sequence stabilization. *Electron. Lett.* **39**(16), 1170–1172 (2003)
36. MK Güllü, S Ertürk, Image sequence stabilization using membership selective fuzzy filtering. *Lect. Notes Comput. Sci. (LNCS)* **2869**, 497–504 (2003)
37. MK Güllü, S Ertürk, Membership function adaptive fuzzy filter for image sequence stabilization. *IEEE Trans. Consum. Electron.* **50**(1), 1–7 (2004)
38. N Kyriakoulis, A Gasteratos, *A Recursive Fuzzy System for Efficient Digital Image Stabilization* (Hindawi Advances in Fuzzy Systems). (2008). doi:10.1155/2008/920615
39. B Pinto, PR Anurenjan, Video stabilization using speeded up robust features, in *International Conference on Communications and Signal Processing (ICCSP)* (Kerala, 2011), pp. 527–531
40. Road. <http://www.jnack.com/adobe/photoshop/videostabilization/>
41. Shaky Car. `matlabroot\toolbox\vipblks\vipdemos\shaky_car.avi`

doi:10.1186/1687-5281-2012-21

Cite this article as: Tanakian et al.: Digital video stabilizer by adaptive fuzzy filtering. *EURASIP Journal on Image and Video Processing* 2012 2012:21.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com
