Adaptive Motion Smoothening for Video Stabilization

Paresh Rawat [1]
[1] Department of Electronics & Comm. Engineering,
[1]Truba College of Science &Tech, India

Jyoti Singhai [2]
[2] Department of Electronics Engineering
[2]Maulana Azad National Institute of Technology,, Bhopal, India

ABSTRACT
Persons using low coast hand held cameras are untrained, thus captured videos suffers from severe handshakes and jitters. Hence digital video stabilization algorithm is required to acquire video sequences by removing these unwanted handshakes and jitters. This paper proposed an adaptive motion smoothening method that removes high frequency jitters, by filtering the accumulated global motion vectors. A modified method for motion vector validation implemented on global motion vectors using adaptive threshold. Proposed method not only removes the jitters but also preserves the scene information. The proposed method reduces the missing image areas significantly. This allows using a simplified edge completion method to generate the full frame stabilized video, which performs better in the presence of large moving object in the scene.

Keywords:
Video stabilization, Motion Smoothening, Motion Estimation, Motion vector validation, Edge completion

1. INTRODUCTION
Exponentially increasing demand of the digital camcorders and mobile phones with video capture capability, demands the video stabilization systems, which can improve visual quality of videos. In video stabilization algorithms, first motion vectors for each pair of video frames are predicted using motion estimation. Then motion smoothening is used to filter the estimated noisy (unwanted) motion vectors to generate stabilized video sequence. For successful video stabilization, the accuracy of the motion estimation must be as high as possible. The motion estimation methods can broadly classified as feature-based approaches [2, 3, 4, and 5] or direct pixel based approaches [1, 7, 9, and 15]. Most of the previous methods of video stabilization uses feature based methods with block motion estimation, as they are faster. However, these are specific to the local effects and their efficiency varied with methods of feature point selection. The direct pixel based methods are efficient in case of slow and smooth motion as in the case of hand held mobile videos. These methods consider entire frame to calculate the motion vector hence are computationally complex, thus it is required to reduce the computation cost of the algorithms.

The use of motion filtering removes the jitters and undesired motions present in the captured video. These undesired motions usually considered as high frequency motion such as vibration, and can effectively remove by low pass filters. Thus, motion filtering also called as motion smoothing. In recent past researchers proposed various low pass filters such as IIR, FIR, Kalman [4], [8], [11], and particle filter [9]. Motion vector integration (MVI) by the first order IIR filters [4]. Gaussian kernel smoothing used to filter the intentional camera motions, and defined as weighted average of past k and future k motions [1], [6].

In this paper a direct pixel based hierarchical differential motion estimation [1, 6 and 7], used to calculate motion vectors. Then it proposed to use adaptive IIR filter to smooth the unwanted motions and jitters [4, 10]. The proposed method adopts the smoothness parameter by validating the motion vectors to identify the frames with smooth and rough regions. Thus will reduce the missing frame areas significantly, thus reduce the computation coast and also simplify the edge completion requirement The rest of this paper is organized as follows;

After introduction, a brief literature review of existing methods with their limitations described in section 2. The proposed video stabilization algorithm discussed briefly in section 3. Proposed adaptive motion smoothening method is described in section 4. Edge completion method discussed in Section 5. Results are presented in Section 6 and the paper is concluded in section 7.

2. LITERATURE REVIEW
In the last two decades, the researchers have developed many video stabilization algorithms. For any video stabilization technique the inter frame motion may be modeled using 2D or 3D motion model. The 2D motion models [1, 2, 3 and 7], are computationally less complex and efficient enough to estimate inter frame motion. Due to its higher efficiency and low computation, cost the 6-parameter 2D affine motion model commonly used to estimate motion vectors. \

Most of the previous methods of video stabilization [3, 8, and 4], have used feature based methods with Blok motion estimation because it is faster. However, they are less accurate. Hany Farid [7] used the direct pixel based method for estimating the motion using Taylor series expansion to minimize the quadratic error function, and 1D separable kernel filters used to find the temporal derivatives.

After estimation to smooth the undesired camera motion in the global transformation chain, various approaches have been proposed [2, 8, 9, 10, 11, and 12]. Buehler *et al.* [10] in 2001 proposed Image based rendering algorithm to stabilize video sequence. The camera motion estimated by non-metric algorithm, and then image-based rendering technique applied to smooth the camera motion. This method produced very accurate results in most of the cases, but it required tuning of motion model parameters to match with the type of camera motion in the video. Matsushita *et al.* [1] developed an improved method called Motion in painting for reconstructing undefined regions and Gaussian kernel filtering used to smooth the camera motion. Sunglok *et.al* [11] in 2009, proposed Adaptive RANSAC method to solve this problem, based on Maximum Likelihood Sample Consensus (MLESAC). For motion, smoothening Kalman filter used to relieves high frequency motion. However, the Kalman filter required to self adjusts for advanced adaptation.

Fatma *et.al* [12] in 2004 proposed an image sequence stabilization system with absolute frame displacement filtering using adaptive polynomial filtering with LMS and RLS algorithms. Polynomial filter based on RLS algorithm perform better but they has a larger computational complexity

Yang *et.al* [13], 2006 has proposed an adaptive image stabilization technique. Method uses sub-sampled multi-resolution block motion estimation to calculate the global motion vectors. An adaptive IIR filter proposed to smooth the unwanted motions. Method also incorporates a two-dimensional compensation method that eliminates the motion fluctuation perpendicular to the panning direction. Tanakian *et.al* [14] in 2011 have proposed a digital video stabilization system by adaptive motion vector validation and filtering. Method uses block motion estimation and adaptive IIR filter filters the unwanted motions. Cai *et.al* [4] in 2009, proposed a robust video stabilization algorithm using feature point selection and delta optical flow. After estimating the motion vectors, a first order IIR filter used for motion smoothening in real time. For post processing applications, they suggested to use non-causal filters. Tang *et.al* [3] in 2009, have proposed a fast video stabilization algorithm based on block matching and edge completion. After estimating the global motion vectors a statistical methods is used to get the global vibrant motion vector

In this paper an adaptive filter is proposed for global motion estimation. Adaptive IIR filters have yet not implemented with direct pixel based global motion estimation. These methods use entire frame at a time hence it is difficult to identify the smooth and rough frames.

3. PROPOSED METHOD

The block diagram of proposed digital video stabilization (DVS) system is depicted in Fig.1. For predicting the motion vectors the direct pixel based hierarchical differential global motion estimation is used as explained in section 3.1. Proposed adaptive motion smoothening method is explained briefly in section 3.2 the proposed smoothening method minimizes the missing frame areas significantly by removing the unwanted motions.

3.1 Hierarchical Differential Global Motion Estimation Motion

In this paper the hierarchical motion estimation framework is suggested to describe the changes (motion) between consecutive frames of the video sequence. 2D affine motion model [7] assumed between video frames

An L=3 level Gaussian pyramid is built for each frame, $f(x, y, t)$ and $f(x, y, t-1)$ as in Fig.2. Global motion vectors can be estimated by differential global motion estimation, using Taylor series expansion as in [6, and 7]. The motion estimated at pyramid level L is used to warp the frame at the next higher level L − 1, until the finest level (the full resolution frame).of the pyramid is reached. For warping the frames bi-cubic interpolation is used.

.The estimated global motion vectors is x and y direction are given as

$$GMV(t) = [GMV_X(t), GMV_Y(t)] \quad (1)$$

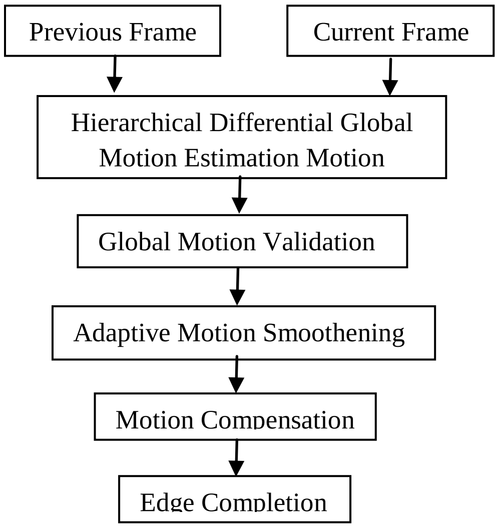


Fig. 1 Block diagram of proposed method

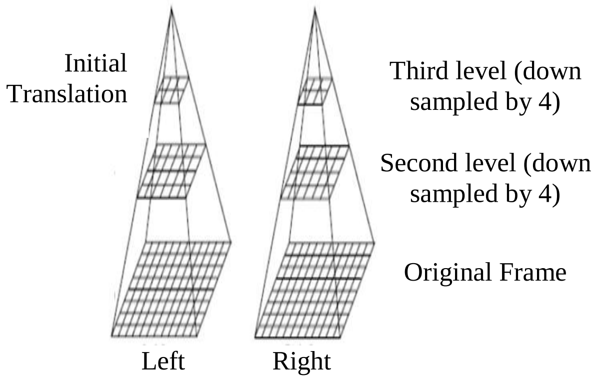


Fig.2 Pyramid Construction

3.1.1 Global Motion Vector Validation

Smooth and rough frames can identify by using adaptive thresholds. The noisy motion vectors obtained for two types of frames as; smooth frames with small object motions with less features (edges), and rough frames with large object motion with more features. As the motion vectors are calculated using direct pixel based method, which consider the entire frame at a time, so in this paper the higher value of the smoothening parameter is chosen to maintain the visual quality.

3.1.1.1 Smoother Frames

To identify the noisy motion vectors of relatively smoother frames, select the threshold of global motion vectors $GMV(t)$ adaptively as;

$$Mean\left(GMV^x(t)\right) < th_1 \quad (2)$$

3*.1.1.2 Rough Frames*

To identify the noisy motion vectors of relatively rough frames, adaptively set the threshold of global motion vectors $GMV(t)$ as;

$$\left(GMV^x(t)\right) > th_2 \qquad (3)$$

Where the threshold th_1 and th_2 are defined as;

$$th_1 = Min\left(GMV^x(t)\right) + 0.45 * Mean\left(GMV^x(t)\right) \qquad (4)$$

$$th_2 = 0.45 * Max\left(GMV^x(t)\right) \qquad (5)$$

Thus, the average $Mean\,(GMV^x(t)$ and minimum $Min\,(GMV^x(t)$ values of the estimated global motion vectors is used to calculate the thresholds. An example of the smoother and rough frame is shown in Fig. 3 for highway video 6.

a) Smoother Frame b) Rough Frame

Fig.3 Smoother and Rough frames for Highway Video 6

3.2 Motion Smoothing

To smooth the undesired camera motion in the global transformation chain, various approaches have proposed in last two decades. Most of the methods were either computationally complex or performs well only with simple and slow camera motion videos. But there efficiency degrades in the presence of large objects and camera motions, as the case in hand held mobile videos. Thus it is required to effectively smoothing the jitters and estimation error from estimated motion vectors $GMV(t)$.

An estimated global motion vectors $GMV(t)$ consist of an intentional motion component viz. camera rotation and panning, and unintentional motion component due to handshake. A good motion correction algorithm should only remove the unwanted motions while maintain the intentional camera motions. The unwanted motions are corresponding to the high frequency components.

In the proposed method a first order adaptive IIR filter [5, and 6], is used for smoothing the accumulation error. The IIR filters are used for motion smoothening because they require less memory space hence can be easily implemented in the real time systems at lower computation coast. These filters produce the videos which are perceptually pleasant to human visual system.

To smooth motion vectors, the accumulated global motion vectors $AGMV(t)$ were calculated as;

$$AGMV(t) = \sum_{i=2}^{t} GMV(t-1) + GMV(t) \qquad (6)$$

Where $GMV(t)$, are global motions vectors at t^{th} frame as calculated by motion estimation using Taylor series expansion and $AGMV(t)$ is defined by vector,

$$AGMV(t) = \{AGMV^x(t)\,, A\,GMV^y(t)\} \qquad (7)$$

Apply first order IIR filter on eq. (25) to get the smoothed accumulated motion vectors $ASMV(t)$ defined as;

$$\mathrm{ASM(t)} = \tau * \mathrm{ASMV(t-1)} + (1-\tau) * \mathrm{AGMV(t-1)} \qquad (8)$$

Where t is representing the current frame. Where τ is defined as the smoothing parameter having the rage as $0 \leq \tau \leq 1$. In this paper τ is set to;

$$\tau = \begin{cases} 0.96 & for\ smoother\ frames \\ 0.9 & for\ rough\ frames \end{cases} \qquad (9)$$

Where the accumulated smoothened motion vectors $ASMV(t)$ are defined a s;

$$ASMV(t) = \{ASMV^x(t)\,, ASMV^y(t)\} \qquad (10)$$

The smoother motion vectors are defined as

$$\mathrm{SVX(t)} = ASMV^x(t) - ASMV^x(\mathrm{t}-1) \qquad (11)$$

$$\mathrm{SVY(t)} = ASMV^y(t) - ASMV^y(\mathrm{t}-1) \qquad (12)$$

The proposed smoothening system tunes the smoothing factor τ of the IIR filter adaptively according to the mother and rough video frame.

The adaptive FIR filters [16] are also widely used to smooth estimated motion vectors. These filters smooth a data sequence using a digital filter which works for both real and complex inputs. These filters are a direct form II transposed implementation of the standard difference equation [16]. In this paper proposed method performance is compared with an adaptive FIR filter with 31 order filtered numerator coefficient vector b and first order denominator coefficient vector a. It is found that for video stabilization the proposed method with motion compensation perform better than the FIR filter.

4. MOTION COMPENSATION

In this paper for motion compensation Gaussian kernel filtering [2] is used.. Use of Gaussian kernel filtering minimizes the missing frame areas. The motion compensated transform from frame i to j can calculated as,

$$MC_t = \sum_{i \in N_t} T_j^i * G(k) \qquad (13)$$

Where, $G(k)$ is the Gaussian kernel filter as defined by [2], and N_t is the neighborhood given as;

$$N_t = m : k - n \leq m \leq k + n \qquad (14)$$

The motion compensated frames f_{ct} can warped from the original frame f_t by

$$f_{ct}(x,y,t) = MC_t * f_t(x,y,t) \qquad (15)$$

5. EDGE COMPLETION

After compensation, still some missing areas exist in the video frames. Some of the existing methods have used trimming to obtain the common frame areas in all video frames. But this method looses information on the corner frame areas. Block based methods suffer from large mussing frame areas. But proposed method with Gaussian frame compensation minimizes the missing areas significantly. Thus in this paper the seamless stitching of missing frame areas is achieved by filling up the missing frame areas using simplified edge completion method. The method is slightly similar to mosaicing method. It utilized the neighborhood of present frames to fill the missing frame areas.

$$f_k(i,j) = median\,[f_N(i,j)] \qquad (16)$$

Where $N \in [(k-n, \ldots, k+n]$ (17)

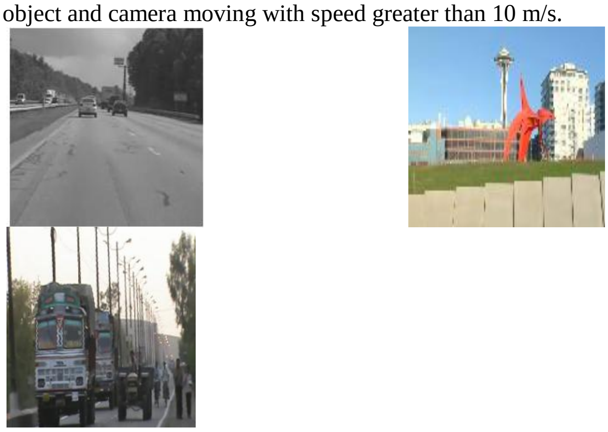
Where $f_k(i,j)$ is defined as the (i,j) pixel of the missing frame and k represents the current frame. In the previous method of edge completion [3] have used n=10, But in the current paper n is selected within 3-5, because missing frame areas are less.

6. RESULT AND DISCUSSION

The Core 2 Duo 2.4GHz processor and MATLAB software is used to develop the proposed video stabilization method. An adaptive motion smoothening method is proposed which removes the high frequency jitters, by filtering the accumulated global motion vectors.
The performance comparison of the proposed method is given in section 6.2. The edge completion is used to generate the full frame stabilized video sequences in section 6.3.

6.1 Input Data

Real time videos are generated using Nokia (6303) mobile phones and Sony digital camcorder and resized to resolution of 176 x 144. The performance of proposed video stabilization algorithm is tested open 32 real time videos. Three videos used for objective evaluation are shown in Fig. 4. First, two are the commonly used videos and third video demonstrating worst case of blur effects with fast and largest motion of object and camera moving with speed greater than 10 m/s.



a) b) c)

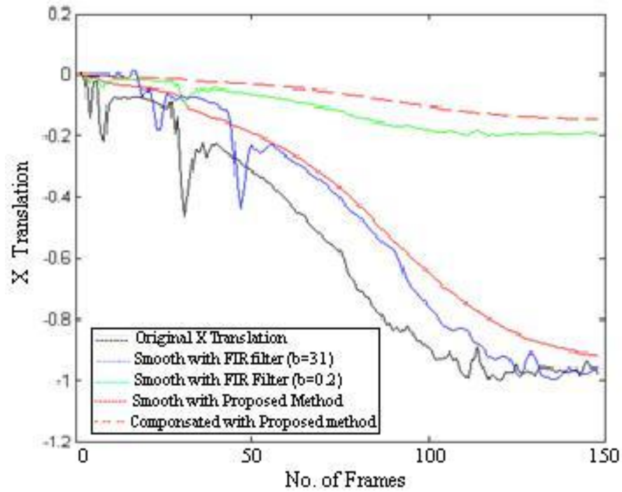
Fig.4 Input videos a) shaky car video, b) Content preserving video [17], c) Highway video 6.

6.2 Motion Smoothening

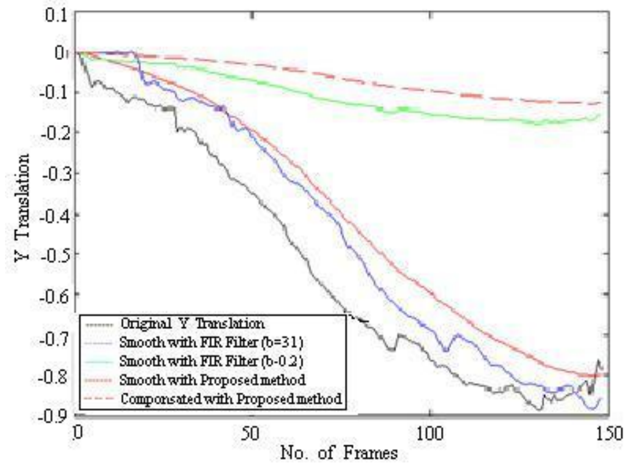
The proposed method first calculates the accumulated global motion vectors in X and Y directions, using the differential motion estimation. To smooth the estimated X and Y translations an adaptive IIR filter is used. Which smooth the undesired motions efficiently and thus stabilizes the video. The performance of proposed method is compared with an adaptive FIR filter [8], with 31 order filtered numerator coefficient vector b with denominator coefficient vector a=1 as in Fig.5.

The compensated translation after Gaussian Kernel filtering for proposed method represented by (- - -) lines as in Fig.5.It is clear that proposed method gives better-stabilized results. A step by step result of proposed video stabilization method is shown for each 25th frame of the simulated distinct Highway video 6 in Fig.6. The large accumulation error is present after the estimation as in Fig 6(b), which filtered effectively by proposed method of motion smoothing which reduce the missing frame areas shown in Fig.6 (c). the result of the edge completion method is shown in Fig. 7 for Highway video 6.

The Inter-frame Transformation Fidelity (ITF) calculated for stabilized videos, which is a measure of the PSNR between successive frames.



a) Comparison X Translation



b) Comparison of Y Translation

Fig 5 Performance comparison of the smoothen Translation for Highway Video 6

$$ITF = \frac{1}{N_{frame} - 1} \sum_{j=1}^{N_{frame} - 1} PSNRR(f_j, f_{j+1}) \quad (18)$$

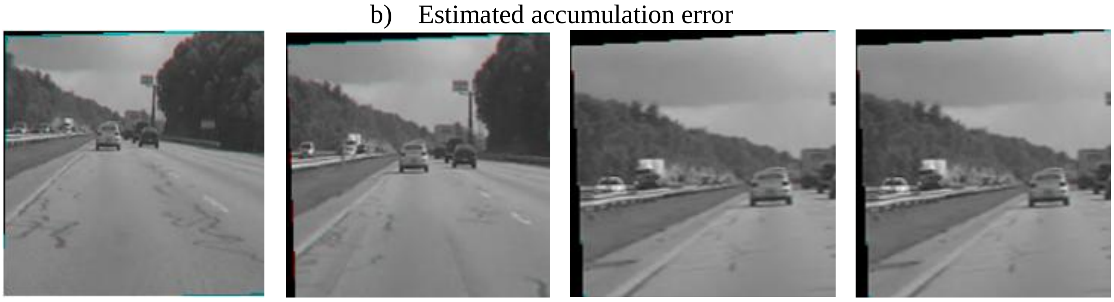
The comparison of the ITF is shown in Table 1. It is clear that the proposed method performs better than the previous adaptive FIR filters at lower computation coast.

Table 1 Comparison of ITF

Video Name	Inter-frame Transformation Fidelity (ITF)		
	With FIR filter [8] (b=31)	With FIR filter (b=0.2)	With Proposed Method
Shaky Car	18.1417	18.319	18.37
Content 3D video 2	18.0392	19.0174	20.4594
Highway video 6	19.815	20.6289	21.268

a) Input frames (25, 50, 75, and 100)

b) Estimated accumulation error



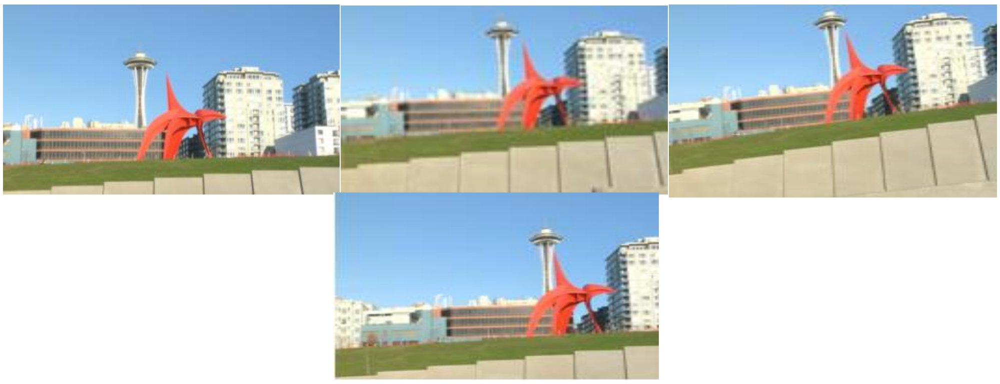
c) Smoothened and compensated frames

Fig. 6 Step wise result of proposed smoothening algorithm for shaky car video

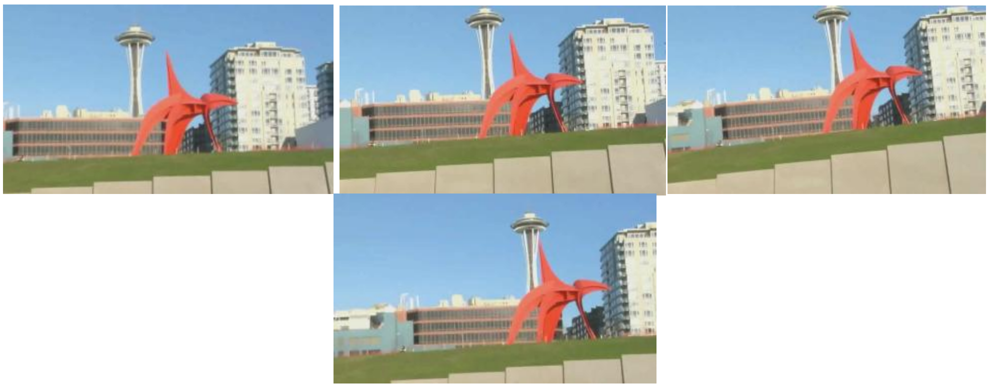
a) Input video frames (25^{th}, 50^{th}, 75^{th} and 100^{th})

b) Stabilized result with Proposed method

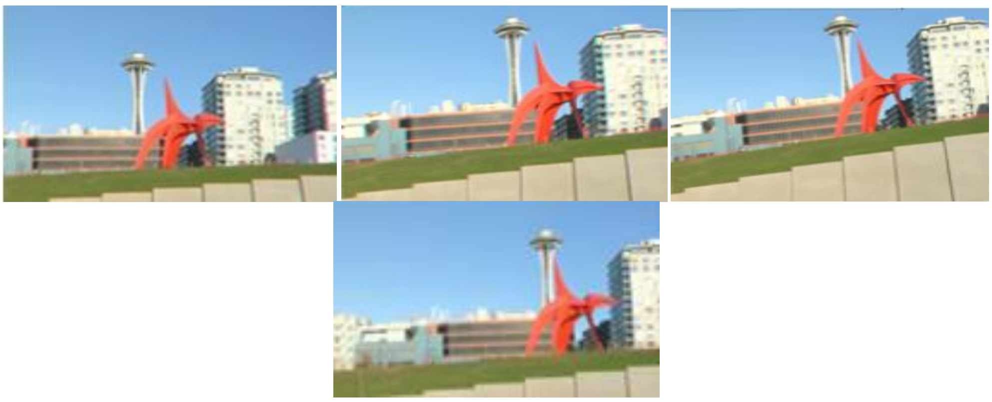
Fig7 Final Stabilized and Edge completed Results for Highway video 6



c) Input video frames (25^{th} , 50^{th} , 75^{th} and 100^{th}) by [5]



d) Stabilized with Deshaker [16] method



e) Stabilized result with proposed Smoothening method

Fig.8 Comparison of the stabilized results with standard Deshaker method (for Content 3D video 2, [5])

Finally, results of the proposed method compared with the standard Deshaker method [16] as in Fig.8. Although Deshaker method is parity consistent but it gives the slightly trimmed video frames. The proposed method stabilizes the video but also preserve the content much better.

7. CONCLUSION AND FUTURE WORK

In this paper, an adaptive motion smoothening method has proposed which efficiently removes the high frequency jitters, by filtering the accumulated global motion vectors at lower computation coast. Smooth and rough frames has identified by using adaptive thresholds to validate the motion vectors. The validation method has proposed for direct global method by considering the entire frame at a time. To accomplish the better visual quality for direct method the smoothening parameter τ is set to higher values. Extensive experiments have conducted on a wide variety of video sequences. The proposed method of motion smoothening performs better for various types of motion and jitters. The proposed method is capable of stabilizing the worst and large motion videos efficiently, where multiple moving objects are in the scene or moving object covers large frame areas. Method can also implement easily in real time.

8 ACKNOWLEDGEMENTS

Authors especially want to acknowledge Mr. Feng Liu [5], and expresses thanks to made input videos globally available. Authors are grateful to everyone who have supported for our research work.

9. REFERENCES

[1] Y. Matsushita, E. Ofek, Weina Ge, X. Tang, and H.Y. Shum, "Full frame video stabilization with motion inpainting", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 28, No.7 pp. 1163-1178, July 2006.

[2] Pyung Soo Kim, FIR filtering based image stabilization mechanism for mobile video appliances Proceeding CIS'04 Proceedings of the First international conference on Computational and Information Science Pages 1106-1113, 2004

[3] Chongwu Tang, Xiaokang Yang,Li Chen,"A fast video stabilization algorithm based on block matching and edge completion", IEEE 13th International workshop on Multimedia signal processing MMSP 2011.

[4] J. Cai, R. Walker, "Robust video stabilization algorithm using feature point selection and delta optical flow", IET Computer Vision, Vol. 3, No. 4, pp. 176–188, 2009

[5] Feng Liu, M. Gleicher, H. Jin, Aseem Agarwala, "Content Preserving Warps for 3D Video Stabilization", Int. Conf. Proc. ACM SIGGRAPH 2009 papers, New York, NY, USA: ACM, 2009, pp. 1-9.

[6] Paresh Rawat, Jyoti Singhai, "Review of motion estimation and video Stabilization technique For hand held mobile video" International Journal of Signal & Image Processing (SIPIJ), Vol.2, No.2, 2011 .

[7] Hany Farid and J. B. Woodward, "Video stabilization and Enhancement", TR 2007-605, Dartmouth College, Computer Science, 1997

[8] Pyung Soo Kim, "FIR filtering based image stabilization mechanism for mobile video appliances", Computational and Information Science Lecture Notes in Computer Science Volume 3314, pp 1106-1113, 2005.

[9] J. Yang, D. Schonfeld, and M. Mohamed, " Robust Video Stabilization based on particle filter racking of projected camera motion", IEEE Transaction on, Circuits and Systems for Video Technology, Vol. 19, No. 7, pp. 945-954, July 2009.

[10]C. Buehler, M. Bosse, and L. Mcmillian, "Non-metric image based rendering for video stabilization", Proc. Computer Vision and Pattern Recognition, Vol. 2, pp.609–614, 2001.

[11] Sunglok Choi, Taemin Kim, and Wonpil Yu, Robust Video Stabilization to Outlier Motion using Adaptive RANSAC", International Conference on Intelligent Robots and Systems October 11-15, 2009 .

[12] Fatma Özbek, Sarp Erturpk, "Image sequence stabilization based on adaptive Polynomial filtering using the LMS and

RLS algorithms", in Journal Of Electrical & Electronics Engineering", pp. 1233-1241, June. 2006.

[13] Shih-Hsuan Yang, *Member*, *IEEE* and Fu-Min Jheng, "An adaptive image stabilization technique", IEEE International Conference on Systems, Man and Cybernetics Vol. 3_ 2006

[14] M.J. Tanakian, M. Rezaei, and F. Mohanna, "Digital video stabilization system by adaptive fuzzy filtering", 19th European signal processing conference, pp. 318-322, Sept. 2011.

[15] M.J. Tanakian, M. Rezaei, and F. Mohanna, "Digital video stabilization system by adaptive Motion vector validation and filtering", Proceeding of 1st International conference on communication engineering pp. 165-168, Dec. 2010.. \\

10. Authors

1. Prof. Paresh Rawat
B.E. 2000, M Tech 2007, Ph. D* (Pursuing)
Designation: Assistant Professor, TCST Bhopal
Research areas: Image and Video Processing

2. Dr. Jyoti Singhai
Qualification: B.E., M. Tech., Ph. D
Designation: Associate Professor, at MANIT Bhopal India. The main Research areas are: Audio ,Video & Image Processing