Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 93 (2016) 359 – 366

Procedia Computer Science

6th International Conference On Advances In Computing & Communications, ICACC 2016, 6-8 September 2016, Cochin, India

A Hybrid filtering approach of Digital Video Stabilization for UAV using Kalman and Low Pass filter

Lakshya Kejriwal[a], Indu Singh[b] *

[a]Computer Science and Engineering, Delhi Technological University, Delhi 110042, India
[b]Department of Computer Science and Engineering, Delhi Technological University, Delhi 110042, India

Abstract

In this paper a new video stabilization algorithm for unmanned aerial vehicles (UAV) has been presented which is used to stabilize the video being transmitted from UAV to the ground station. First, the corner points are extracted using Good Features to Track corner detection algorithm and the extracted points are used to compute the optical flow between two consecutive frames. Next, the points detected from optical flow are used to estimate the motion parameters using an affine transform model. Subsequently, a hybrid filter consisting of Kalman and low pass filter is used to smooth the estimated motion parameters and the frames are warped using the smoothed parameters to obtain a stabilized video sequence. The experimental results show that the algorithm can remove the unwanted vibration more effectively than the one that only uses either a Kalman Filter or a low pass filter.

Peer-review under responsibility of the Organizing Committee of ICACC 2016

Keywords: digital video stabilization; kalman filter; low pass filter; optical flow estimation; image compensation

1. Introduction

In recent years, unmanned aerial vehicles have found extensive use in commercial applications such as urban surveillance, disaster management, aerial photography, delivery services, inspection of pipelines etc. The UAV's today are equipped with a variety of payloads for autonomous navigation, data acquisition, real time surveillance for which the UAV has a streaming camera on board.

* Corresponding author. Tel.: +0-000-000-0000 ; fax: +0-000-000-0000 .
E-mail address: lakshya5340@gmail.com, indu.singh.dtu14@gmail.com

Peer-review under responsibility of the Organizing Committee of ICACC 2016
doi:10.1016/j.procs.2016.07.221

In order to execute the mission the operator needs to have a clear video displayed on the ground station. However due to the mechanical vibrations caused by the engine and the effect of turbulent wind, the video is distorted and unstable. Therefore the video frames have to be stabilized for immediate object detection and real time surveillance. Video stabilization is a process to reduce the unwanted motion of the camera so as to remove the jitter and instability in the video frame images[1].

The stabilization can be implemented either on the structure holding the camera inside the UAV with the help of Mechanical stabilization or on the ground station terminal using Digital stabilization technique[2]. Mechanical stabilization employs mechanical devices like gimbal to reduce the vibration of the structure holding the camera and to adjust for the movement of the UAV. However this method is not able to reduce all the vibrations, involves complex hardware, is costly to implement and can be time consuming. The Digital Stabilization technique, however employs image processing based methods and can be used to remove the unwanted vibration that cannot be removed by Mechanical stabilization method. It costs less than the Mechanical technique as it does not involve complex hardware and requires only software to stabilize the video. Therefore digital stabilization is a better method of stabilization since it's inexpensive and easy to implement.

The remaining structure of the paper is organized as follows: Section 2 cites some related work on video stabilization. Section 3 describes in detail our proposed methodology. Experimental result are presented in Section 4. Finally, the conclusion is given in Section 5.

2. Related Work

The overview of some of the approaches to deal with the advent growth of digital video stabilization methods are given in this section.

Shen et al[3] illustrated a digital video stabilization technique that applied PCA-SIFT (Principal Component Analysis-Scale Invariant Feature Transform) method to extract features and an adaptive particle filter was used to refine the results of global motion estimation. A cost function called SIFT-BMSE (SIFT Block mean square error) was proposed in the particle filter framework to filter out the foreground object pixels.

Y.Wang et al[4] proposed the cubic spline smoothing technique to separate the unwanted vibrations from intentional one's and eventually produce a stabilized video sequence. But when cubic spline interpolation method is used, knots selection plays a significant role on the result of video stabilization. If the knots are located on 'noisy' points, it will deteriorate the stabilization effect.

A methodology for moving camera stabilization was presented in[5] which adopted Kalman filter and least squares fitting algorithm to remove the motion of unwanted jitter and retain the motion of moving camera in order to produce the stabilized sequence. This method used least square fitting which is very sensitive to outliers and thus can deteriorate the stabilization effect.

C.Wang et al[6] demonstrated a robust digital image stabilization method which was based on feature point tracking using Kanade-Lucas-Tomasi tracker and subsequently used the feature points to estimate the trajectory between two subsequent image frames. Unintended motion compensation was accomplished by an adaptive Kalman filter in order to produce the stabilized video sequence. Also a new method was employed to automatically change the reference frame by detecting a scene change.

A. Litvin et al[7] introduced a probabilistic estimation framework for separating the unwanted motion from desirable one and p-norm-based multi resolution approach to estimate the motion parameters which were described using a six parameter affine model .The undesirable motion was separated from intentional one by using recursive Kalman Filter. It also used mosaicking to construct undefined areas in each frame after motion compensation.

Matsushita et al[8] presented full frame video stabilization method which employed deblurring algorithms that transfers and interpolates sharper pixels of neighbouring frames instead of estimating point spread functions to produce good quality full frame stabilized videos. The image part which were missing after motion compensation were filled by aligning image data from previous and subsequent frames. However, this method relies on global motion estimation which is sensitive when an object covers large amounts of image area.

Hu et al[9] proposed a video stabilization method which used SIFT to extract feature points from input frames and a Gaussian smoothing method extended by parabolic fitting for motion smoothing. Also, the undefined areas in the output frames resulting from image compensation were reconstructed using mosaicking with dynamic programming. Since this method used SIFT for feature extraction and Gaussian smoother it is not suited for high resolution video.

3. Proposed Approach

The flowchart of the stabilization algorithm is illustrated in Fig. 1.The proposed algorithm consists of four steps. First, the corner points are extracted using the algorithm given by J. Shi et al[10] known as Good Features to Track and the extracted points are used to compute the optical flow between two consecutive frames. Next, the points detected from optical flow are used to estimate the motion parameters using an affine transform model. Subsequently, a hybrid filter consisting of Kalman and low pass filter is used to smooth the estimated motion parameters and the frames are warped using the smoothed parameters to obtain a stabilized video sequence.

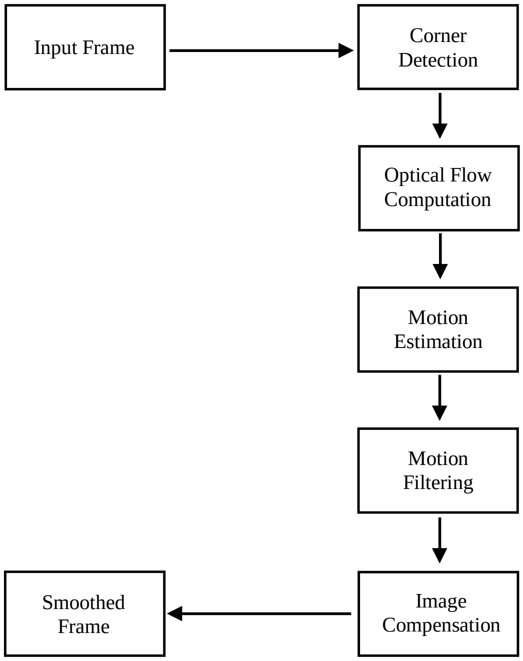

Fig. 1. Flowchart of video stabilization algorithm

3.1 Corner Detection and Optical flow Computation

The corner points are detected in input image frame using Good Features To Track. These corner points are used to compute the apparent motion, optical flow, of the points between two consecutive frames. The motion of the pixel points between two consecutive frames is computed by minimizing the function defined as follows:

$$\varepsilon(d) = \varepsilon(d_x, d_y) = \sum_{x=u_x-w_x}^{u_x+w_x} \sum_{y=u_y-w_y}^{u_y+w_y} (I(x,y) - J(x+d_x, y+d_y))^2 \quad (1)$$

where $\mathbf{u} = [u_x \; u_y]^T$ is an image point on the first image I, $\mathbf{d} = [d_x \; d_y]^T$ is the image velocity at $\mathbf{x}$, where $\mathbf{x}=[x \; y]^T$ and x and y are the pixel coordinates at a generic point $\mathbf{x}$ and J is the second image. The function is minimized using Pyramid Lucas Kanade Tracker[11] algorithm whose goal is to find the location $\mathbf{v} = \mathbf{u}+\mathbf{d} = [u_x + d_x \; u_y + d_y]^T$ on the second image J such that I($\mathbf{u}$) and J($\mathbf{v}$) are similar.

3.2 Motion Estimation

The pixel points detected by computing the apparent motion between two consecutive frames is subsequently used to estimate the motion parameters between the two frames. For simplicity a 2D affine transform model with four unknown parameters: scale, rotation and translation in x and y axis is used. Let I(x,y) and J(x',y') be the pixel location of the point given by optical flow in the two consecutive frames of the video sequence, I and J. The transform matrix between the two points can be given by:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} S.\cos\theta & -S.\sin\theta \\ S.\sin\theta & S.\cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} T_x \\ T_y \end{bmatrix} \tag{2}$$

where S is scale, θ is rotation, T_x and T_y is the translation in x and y axis respectively.

3.3 Image Compensation

In order to obtain a stabilized video, it is assumed that the first frame is stable which is used as a reference frame to stabilize the next frame. The second frame is subsequently used as a reference frame to stabilize the next frame and the process continues until the last frame. The parameters estimated from equation (2) contain both desired and undesired motion. The desired motion which is the intended movement of the UAV is slow and has low frequency. The undesired motion of the camera sensor is fast, has high frequency and is unpredictable. In order to compensate the current frame the desired motion and the undesired motion has to be separated.
Let us consider a set of frames A_i, for i=1,2,3,……n, the above two steps are used to estimate the motion parameters between two consecutive frames A_i and A_{i+1} and represent them as affine transform M_i as follows:

$$M = \begin{bmatrix} S.\cos\theta & -S.\sin\theta & T_x \\ S.\sin\theta & S.\cos\theta & T_y \\ 0 & 0 & 1 \end{bmatrix} \tag{3}$$

The cumulative transform $M_{cummulative}$, which is the product of all the affine transforms until the current frame and represents all the camera motion after the first frame, is estimated for each frame. To stabilize the video sequence, a hybrid filter is applied to the motion parameters obtained from cumulative transform and constructed into a full transform, $M_{smoothed}$. In order to obtain a smoothed frame, the current frame is then warped using equation (4)[12].

$$A_{smoothed} = M_{smoothed} .(M_{cummulative})^{-1}.A_{current} \tag{4}$$

3.4 Motion Filtering

The motion filtering technique using Kalman Filter is based on the assumption that there is only white Gaussian noise in the unstable video. However, this may not be the actual situation and the video may still be unstable after being processed by Kalman Filter. In this paper a new motion filtering algorithm has been presented. Our method is that the motion parameters estimated by Kalman Filter are passed into a low pass filter to remove the additional noise in the jitter video. The parameters obtained from low pass filter are again estimated by Kalman Filter to remove any undesired motion still present after low pass filtering to obtain the smoothed motion parameters. Here the filter is called a Hybrid Filter. The architecture of the hybrid filter is shown in Fig. 2.

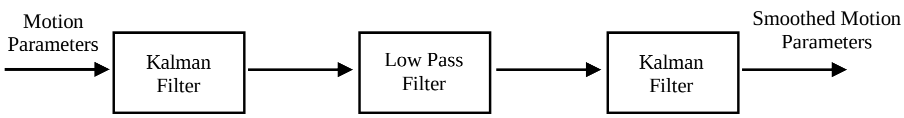

Fig. 2. Hybrid Filter

Kalman Filter is a linear recurrence filter that predicts the next state of the system based only on the previous state of the system, but not dependent on sequence of past states, so it can be used for real time processing. X_k is the state of system at time k, A is the state transfer matrix which describes the transition relationship of the system between time, k and k+1, W_k is the system noise.

$$X_{k+1} = AX_k + W_k \tag{5}$$

Z_k is the observation state of system state X_k at time k, H is the observation matrix, V_k is the observation noise at time k. The final observation state Z_k is described as below:

$$Z_k = HX_k + V_k \tag{6}$$

The specific procedure of Kalman Filter consists of two steps: prediction and correction state and the overall aim is to regain the value of X_k. Here Q is the covariance matrix of process noise, R is the covariance matrix of observation noise whose value is set to 10^{-5}, x_k is the estimated value of X_k and P_k is the estimated value of covariance of X_k.

(i) The aim of prediction state is to predict the values of X_k and P_k.

$$\begin{aligned} x'_k &= Ax_{k-1} \\ P'_k &= AP_{k-1}A' + Q \end{aligned} \tag{7}$$

Where x'_k and P'_k are the estimated values of X_k and P_k respectively.

(ii) The aim of correction state is to use the value of current measurement, Z_k to update the value of x'_k

$$\begin{aligned} K_g &= P_{k-1}H'(HP_{k-1}H' + R)^{-1} \\ x_k &= x'_k + K_g(Z_k - Hx'_k) \\ P_k &= (I - K_g H)P'_k \end{aligned} \tag{8}$$

Where K_g is the Kalman Gain and I is the identity matrix[13].

The initial value of x_0 and P_0 is selected at random because we can't get the initial information of the video sequence before we process it. The values are adjusted in the first few frames after which the Kalman Filter can operate properly. For a low pass filter a discrete Gaussian filter is selected which is an approximation by sampling the continuous Gaussian of equation (9).

$$h(t) = \frac{1}{\sqrt{2\pi}.\sigma}.e^{-\frac{t^2}{2\sigma^2}} \tag{9}$$

The discrete Gaussian impulse response is sampled in a finite number of frames from –r to r using equation (10).

$$g[k] = \frac{1}{G}e^{-(\frac{k}{\sigma})^2} \tag{10}$$

Where $G = \sum_{i=-r}^{r} e^{-(\frac{i}{\sigma})^2}$ and k = [-r , r]. The latency and computational effort of the low pass filter depends on the value chosen for r. In our case we take $\sigma = 4$ and r=10 to minimize latency. The motion parameters are finally discrete convolution operated with $g[k]$ to get new motion parameters which are again passed to Kalman Filter to further remove the undesired motion and subsequently obtain the smoothed parameters.

4. Experimental Results

The algorithm was implemented in C++ using OpenCV modules. The experimental video on which the proposed algorithm is tested is captured from RC UAV that transmits a 1.3 GHz wireless video from a board camera which has a resolution of 640x480 pixels. For Good Features To Track the maximum number of corners is set to 200 and the quality level to 0.01.The corner detection results are shown in Fig. 4. (a), the two images are the consecutive frames of a video sequence as shown in Fig. 3. The optical flow was computed with a window size of 21 and the results are shown in second row of Fig. 4. (b).

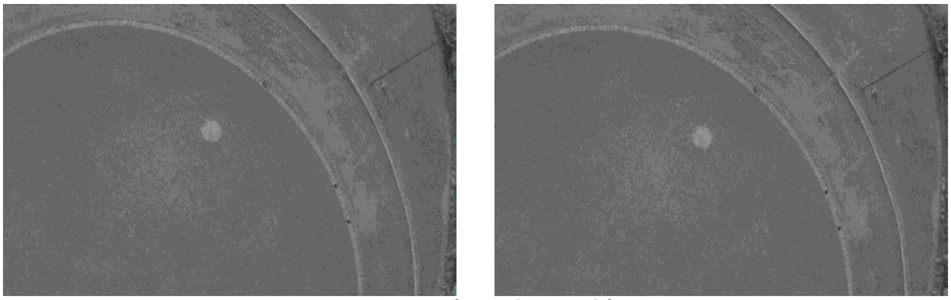

Fig. 3. (a) First frame; (b) Second frame

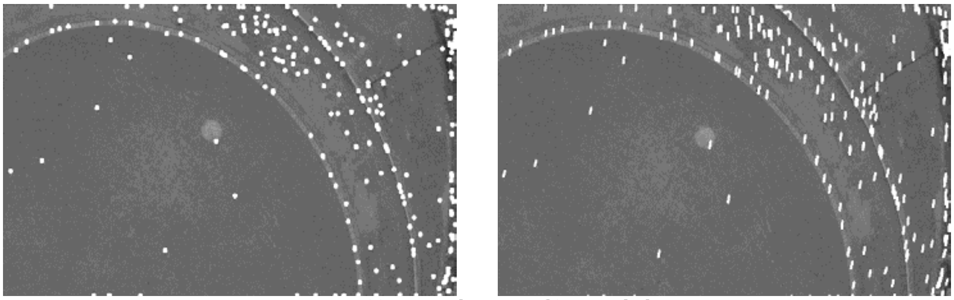

Fig. 4. (a) Corner detection; (b) Optical Flow

Tests are run on a number of video sequences each close to 3 minutes (5400 frames). The smoothed parameters after motion filtering by our algorithm are shown in Fig. 5., which are compared with the smoothed parameters obtained using other filters. From the figure it can be seen that the curve for hybrid filter is more smooth as compared to curve for other filters.

In order to better evaluate the efficiency of our algorithm, the mean square error between two consecutive frame images is computed for which the mean squared error of every pixel value between two consecutive frames is calculated using the equation in (10).

$$\mathrm{MSE}(I,J)=\frac{1}{X*Y}\sum_{i=1}^{X}\sum_{j=1}^{Y}(I(i,j)-J(i,j))^2 \tag{11}$$

Where $I(i,j)$ and $J(i,j)$ represent pixel values of location (i,j) in frames I and J respectively.

Table 1. Error of two consecutive frames by different filtering algorithms

Test frame sequences	Original Video	Stabilized video by our algorithm	Stabilized video by Kalman Filter	Stabilized Video by Low Pass Filter and Kalman Filter
65th and 66th frame	812.849	28.919	227.122	125.625
Maximum	3021.274	106.542	846.271	466.381
Average	756.548	26.479	211.834	116.508

From Table 1 it can be seen that the proposed algorithm has reduced the error by 24.38% when compared to Kalman Filter and by 11.89% when compared to Kalman Filter and low pass filter.

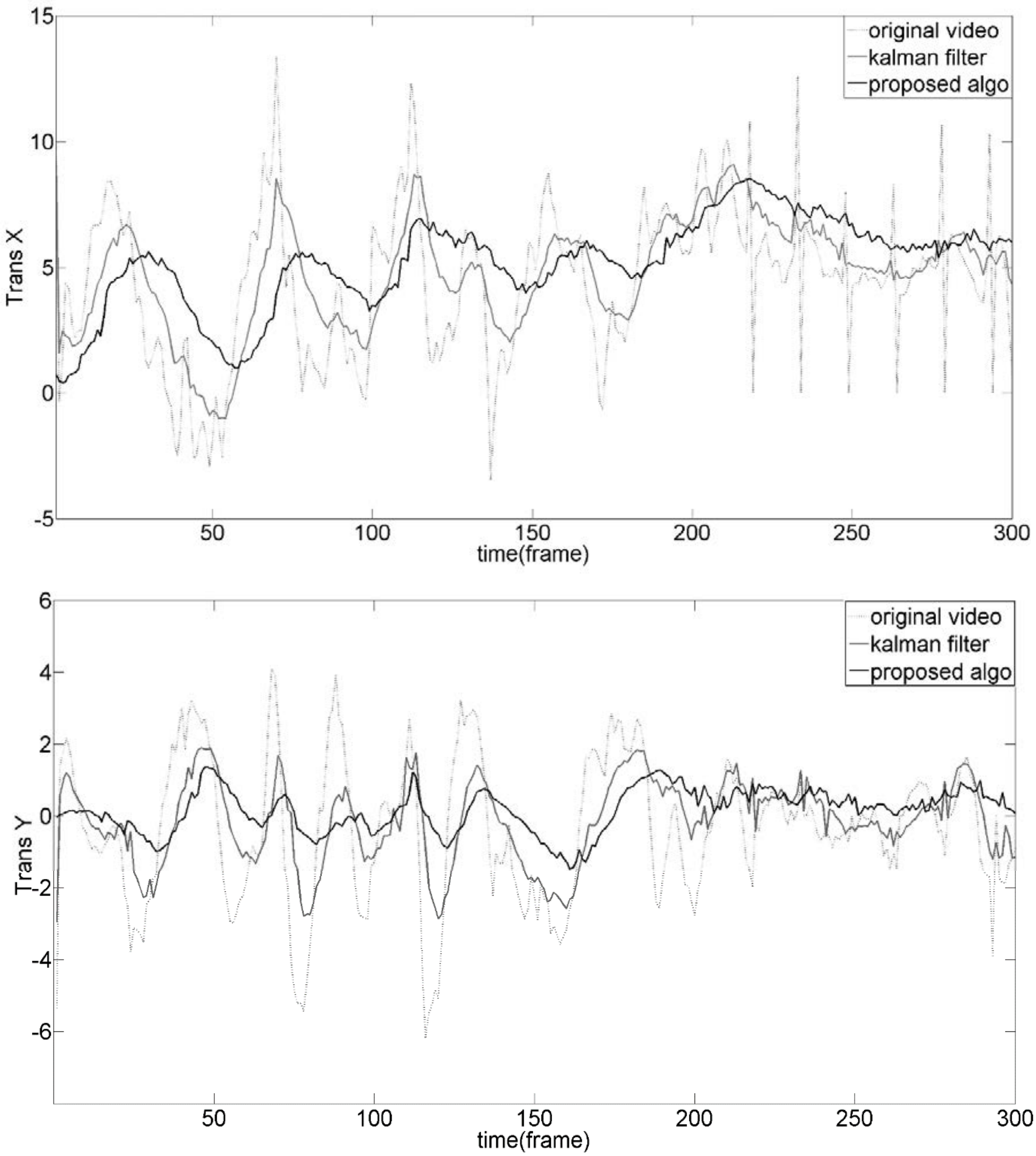

Fig. 5. Smoothed Parameters

5. Conclusion

A new improved video stabilization algorithm has been presented in this paper. The method demonstrated is that the motion parameters obtained by computing the optical flow between two subsequent frames are filtered by a hybrid filter consisting of Kalman Filter and a low pass filter. The proposed algorithm is tested on a video captured from RC

UAV that transmits a 1.3 GHz wireless video from a board camera to ground station. The experimental results show that the algorithm is more effective in removing the unwanted vibration as compared to the one using only a Kalman Filter or low pass filter. Our future work will focus on enhancing our model by using a real time technique of mosaicking to compensate for the uncovered area when the frame is warped to obtain a stable sequence.

References

1. Saranyu Navayot and Nuttaka Homsup, Real-Time Video Stabilization for Aerial Mobile Multimedia Communication, International Journal of Innovation, Management and Technology, Vol. 4, No. 1,pp 26-30, February 2013
2. H. Shen, Q. Pan, Y. Cheng and Y. Yu, Fast video stabilization algorithm for UAV**,** IEEE International Conference on Intelligent Computing and Intelligent Systems,ICIS 2009,pp 524-546,Shanghai,China,2009
3. Y. Shen, P. Guturu, T. Damarla, B. P. Buckles and K. R. Namuduri, Video Stabilization Using Principal Component Analysis and Scale Invariant Feature Transform in Particle Filter Framework, IEEE Transactions on Consumer Electronics, Vol. 55, No. 3, pp. 1714-1721, 2009
4. Y. Wang, Z. Hou, K. Leman and R. Chang, Real-Time Video Stabilization for Unmanned Aerial Vehicles, MVA2011 IAPR Conference on Machine Vision Applications, Nara, Japan, pp 336-339, June 13-15, 2011
5. H. Yu and W. Shang, Moving camera video stabilization based on Kalman filter and least squares fitting, Proceedings of 11th World Congress on Intelligent Control and Automation (WCICA), Shenyang, China, pp 5956-5961, June 29-July 4,2014
6. Chuntao Wang, Jin-Hyung Kim, Keun-Yung Byun, Jiangqun Ni, and Sung-Jea Ko, Robust Digital Image Stabilization Using the Kalman Filter, IEEE Transactions on Consumer Electronics, Vol. 55, No. 1, pp 6-14, February 2009
7. A. Litvin, J. Konrad, and W. Karl, "Probabilistic Video Stabilization Using Kalman Filtering and Mosaicking," Proc. IS&T/SPIE Symp. Electronic Imaging, Image, and Video Comm., pp. 663-674, 2003.
8. Yasuyuki Matsushita, Eyal Ofek, Weina Ge, Xiaoou Tang, Heung-Yeung Shumb, Full-Frame Video Stabilization with Motion Inpainting, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 28, No. 7, pp 1150-1163, July 2006
9. R. Hu, R. Shi, I. Shen, W. Chen,Video Stabilization Using Scale-Invariant Features, The 11th International Conference Information Visualization, pp. 871-877, 2007
10. J. Shi and C. Tomasi, Good Features to Track, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp 593-600, June 1994.
11. J.-Y. Bouguet, Pyramid implementation of the Lucas Kanade feature tracker: description of the algorithm, in Intel Corporation, Microsoft Research Labs, Technical Report, 2000.
12. Y. Wang, R. Chang, T. Chua, K. Leman, N. Pham, Video Stabilization Based on High Degree B-spline Smoothing, 21st International Conference on Pattern Recognition (ICPR), Tsukuba, Japan, pp 3152-3155, November 11-15, 2012
13. D. Simon, Optimal state estimation: Kalman, H infinity, and nonlinear approaches, John Wiley & Sons, pp. 125-129, 2006