University of Sindh Journal of Information and Communication Technology
(USJICT)
Volume 2, Issue 3, July 2018

ISSN-E: 2523-1235, ISSN-P: 2521-5582
Website: http://sujo.usindh.edu.pk/index.php/USJICT/

Robust Technique for Object Tracking by interference of Global Motion Estimation and Kalman Filter

Javaria Maqsood[1], Asma Katiar[2], Lubna Ali[3]

[1]Department of Information and Technology, [2] Department of Mechatronics Engineering,
Mehran University of Engineering and Technology, Jamshoro, Pakistan
[3]Department of Computing, Indus University Gulshan Campus, Karachi Pakistan
javariamaqsood143@gmail.com, asma.katiar@indus.edu.pk, lubna.ali@indus.edu.pk

Abstract: In today's modern world of computer vision there are many techniques for object tracking. But still there is great capacity available for further research. A robust technique for object tracking is proposed in this paper. In this work a fusion of global motion estimation and Kalman filter-based tracking algorithm is implemented which detects and tracks all the moving objects in the video. The algorithm detects corners in a frame and tracks the moving ones in the subsequent frames of the input video. The movement of a moving object is traced by persisting the motion trajectory of corner points on that object. Video stabilization is also implemented so that the moving or shaky video can be processed and amended so that Kalman filter can be implemented. The proposed methodology achieved a precision of 94.73 percent which is quite good in comparison to other published techniques.

Keywords: Tracking, Stabilization, Global Motion Estimation, Kalman Filter;

I. Introduction

In this modern era of life computer vision has controlled day to day life hacks. One of those is object tracking which is one of the most essential need of computer vision world. Object tracking proves to be great help in many of the ways. Now a days it is being utilized for Security Management, Traffic Management, Speed Monitoring, Traffic Violation Monitoring, Road blockage identification and many such others.

The Jittery Videos and Good estimation are the two biggest challenges in this research. So even today after many years of research, researchers are working on improvement in the algorithms being proposed. Researchers have widely presented multiple algorithms with eye catching results. But still there is great gap available for the research and introduction to new algorithms in this field. Authors have proposed a new robust algorithm for this problem in this paper.

The unwanted movements, vibrations and noise caused by unstable camera support can degrade the quality of the video, need to be solved for the quality enhancement of the video. This process of video stabilization is done as post operational task of the video processing in some situation and could be an essential part of video recording and processing as concurrent operation in some other scenarios where live operation are being performed such as object tracking and recognition. It is well known that, problem identified is problem solved! So is the case here, as engineering intelligentsia has started to contemplate over the idea of stabilizing video using analytical as well as statistical algorithms. Various video stabilization techniques used for the purpose include. One of the significant techniques for video stabilizations is based on Global Motion estimation (GME).

The motion estimation algorithms may further be divided into three broad streams, Block matching technique, Transform domain methods, and Feature matching algorithms [1-2].In general various other methods include the filtering methods such as Kalman filter and low pass filter, FIR (finite impulse response) filter method, local subspace method, content preserving warping (CPW) method[3-5].Software of video stabilization usually deploy such algorithms in post processing stages, which is done on already recorded videos.

In this proposed methodology a fusion method for increase the accuracy of video stabilization to get the better results is being proposed. Here Video stabilization will be done by integration of two algorithms, these are: global motion estimation and Kalman filter method.

Ying, y., et al, conceived the algorithm to detach the undesirable camera movements and termed it as "Global Motion Estimation". The basic premise of which is to cleave the global motion estimation into Motion approximation and Motion correction [6].

Juanjuan Zhu, et al, proposed the notion of "video stabilization" primarily constituting 3 steps, Motion estimation, Motion filter and Motion compensation. To serve this purpose he employed Kalman filters along with low-pass filters. Reason behind was to gauge the quality of image stabilization and to establish a hypothesis under which noise in motion could be termed as white noise [7].

S. Erturk, et al, proposed a filtering method used low-pass filter, the camera motion path in Fourier domain. The solution of this method gives a smooth stabilized motion and can be used for off-line stabilization. But unfortunately for a real time implementation on a device the solution is unacceptable due to requirements of the large memory to reserve the several frames of the video and to minimize the memory (causal low-pass filter) is preferred [8].

In [9] author has proposed a methodology for object tracking which is based on Position Vectors and Pattern Matching however the next iteration is very much depended on previous results so it's really important to have good results in start.

Neuro-Fuzzy Based Approach is also being used for this problem [10]. In [11] author has proposed a system for real time object tracking. That system works on multi-inertial sensing data and the system is placed on drone.

In papers [12-17] authors are focusing to improve the robustness of the object tracking technique as it is very helpful when it is needed in vehicles.

In [18] author have discussed multiple method for object tracking in real time but despite of all these methods still a great gap is available for further research. As object tracking in real time need to be really robust but all the proposed and published work still can't tackle the real time constraints.

The main objective of this paper is to achieve following objectives in this research domain.

- Need to increase the accuracy of the video stabilization.
- Fusion of both algorithms.
- To produce a better performance of our video stabilization algorithm by applying global motion estimation (GME) method and Kalman filter method.

II. Type Style and Fonts

In the proposed methodology there are two main implementations. The first one is video stabilization process and then comes the implementation of Kalman filter for estimation purpose. The two of the following are discussed in following headings.

A. Vedio Stabilization

The first and foremost requirement in this project was to stabilize the video to run the Kalman algorithm. This was necessitated to get rid of the optical flow on stationery points caused by the camera motion. The Kalman tracker produced far better results when it was applied on the stabilized video and unwanted optical flow was eliminated.

The theme of video stabilization hinges on the idea of searching for the "background plane" in a video arrangement [9] and to use the observed distortion to correct for camera motion. This stabilization algorithm involves two steps. First, the affine image transformations between all neighboring frames of a video sequence is determined using the 'Global Motion Estimation' function applied to point correspondences between two images. Second, the video frames are warped to achieve a stabilized video. Stepwise details of the algorithm are given below:

i. Declaim in the leading two frames of a video arrangement as an image intensity as there would be no impact of color intensity over stabilization. Furthermore if the gray scale image would be used it would give a higher speed.

ii. Finding area of concern (which would be mostly corners) from leading two frames, and then select expected similarities amongst both them. For this purpose Fast features would be extracted which is one of the fastest corner or edge detector algorithm.

iii. Selection of similarities amongst the corner or edges points found in the previous step. For each point, extraction of a Fast Retina Key point (FREAK) descriptor would be done and it would be positioned around it. The matching cost being utilized amongst the points will be the Hamming distance.

iv. A robust approximation of the geometric transformation amongst the twofold frames is set up by means of the M-estimator sample Consensus (MSAC) algorithm, which is a modification of the RANSAC algorithm and is applied in the 'Global Motion Estimation' function. With the help of these it computes the homography amongst two frames that marks the inliers from the first set of points match most closely with the inliers from the second set. This affine transformation will be made by a 3-by-3 matrix of the form:

$$\begin{matrix} a1 & a3 & 0 \\ a2 & a4 & 0 \\ tx & ty & 1 \end{matrix}$$

The parameter 'a' describes scale, rotation, and shearing properties of the conversion, while the constraints 't' are conversion constraints. This transformation is used to change the images such that their conforming features will be stimulated to the similar image position.

v. Assumption is made here that the background plane will be moved or altered consequently amongst the frames, then at that point this transformation will be capturing the camera motion. Therefore correction for this will stabilize the video. This scenario will uphold as long as the camera motion amongst corresponding frames is small enough, or, the data rate of video is high which would make it difficult to process the algorithm.

vi. For all the frames of the video T_i = 0,1,2......., utilization of the overhead technique for approximation of the falsification amongst all frames T_i and T_{i+1} as affine transformation, *Hi* will be made. Therefore the accumulative falsification of *i*th frame comparative to the previous frame will be the product of all the previous inter-frame transformation. Which can be presented by,

$$Hcumulataive, i = \prod_{j=0}^{i-1} Hi$$

vii. To achieve arithmetical straightforwardness and steadiness, the matrix can be converted as a simpler scale, rotation, translation transformation. This would have individual four free constraints paralleled to the complete affine transform's six: one scale factor, one angle, and two translations. It has the form:

$$\begin{matrix} s.cos(\phi) & s.-sin(\phi) & 0 \\ s.sin(\phi) & s.cos(\phi) & 0 \\ tx & ty & 1 \end{matrix}$$

viii. The homography and the subsequent re-alignment of two video frames introduce black borders in the video frames which account for camera motion. These borders are eliminated by cropping all the video frames to achieve a smooth video.

B. Abbreviations and Acronyms

Kalman filter is used to estimate the object movement. The Kalman filter works under the phenomena of prediction and estimation. As the below figure suggests the working of Kalman filter.

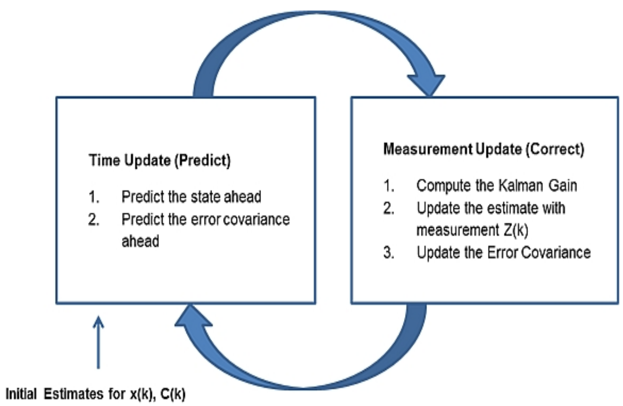

Figure 1: Kalman Filter Process

Figure 1 clearly depicts the working of Kalman filter. Kalman filter first predicts the state using initial estimates and further uses the previous estimates of the system. Below are the equations being used for the estimation of the object.

$$\left.\begin{aligned} \hat{x}_k^- &= A\hat{x}_{k-1} + Bu_{k-1} \\ P_k^- &= AP_{k-1}A^T + Q \end{aligned}\right\} \text{Prediction}$$

$$\left.\begin{aligned} K_k &= P_k^- H^T (HP_k^- H^T + R)^{-1} \\ \hat{x}_k &= \hat{x}_k^- + K_k(z_k - H\hat{x}_k^-) \\ P_k &= (I - K_k H)P_k^- \end{aligned}\right\} \text{Correction}$$

The first equation of prediction is known as Prediction of Estimate which is used to get state observation. Where A is the state transition matrix and B is the state noise matrix. 'u' Presents the zero-mean excitation or driving noise with covariance.

The second equation of prediction is calculating Minimum prediction mean squared error. Which is then used to calculate Kalman gain 'K' as shown in first equation of correction. This process than further calculates the next state. Second equation of correction shows how the correction of estimate is made and then new minimum mean square error is calculated using last equation of correction.

- This process is followed for using Kalman filter in the proposed system. Integration of Kalman filter in the overall system is shown in the figure 2 below.

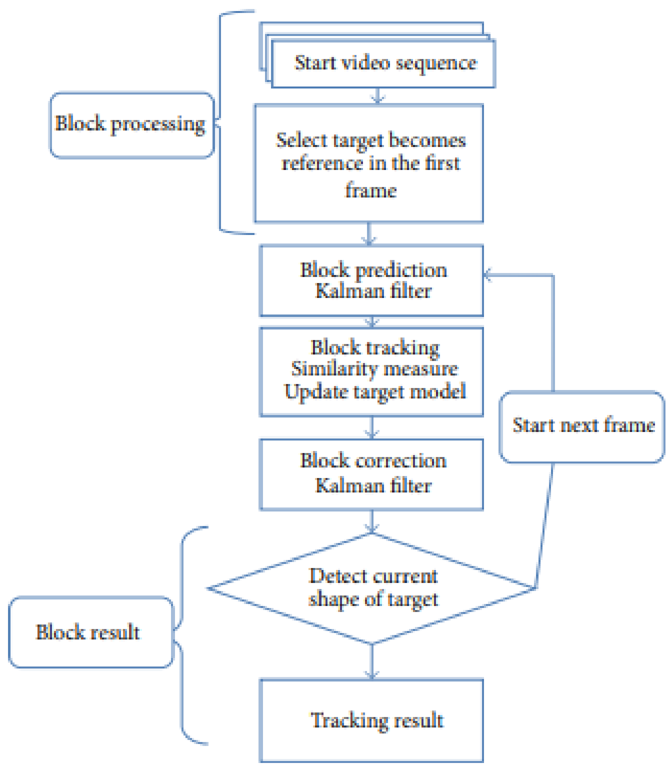

Figure 2: Overall Kalman Proposed system

Kalman filter is being processed on the upcoming frame of the video, And then The guessed estimate is being checked and if the set test is not passed the process of Run Moment Matching Optimization is processed to make the estimate valid and the estimation of object can be made with minimum mean squared error.

III. RESULTS

For comparing the proposed algorithm a testing approach by author of [9]. Same Video was taken and the 3245 frames sampled at 10 Hz were prepared. The author [9] took three quality parameters for checking authentication of the proposed algorithm. Those were,

Precision: Will give percentage that obstacle was identified correctly out of all positives and it is relevant.

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

Recall: Also known as the True Positive Rate, will give percentage that correctly identified object retrieved in a search

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negatives}$$

False Positive Rates: Gives us the percentage that the object wasn't there but the algorithm recognised it as an object.

$$False\ Positive\ Rate = \frac{False\ Positive}{False\ Positive + True\ Negatives}$$

Where,

True Positive represents that the object was there and it was tracked as an object

True Negative represents that condition doesn't hold while in fact it does.

False Positives represent that object wasn't there but it was detected as an object.

False negative presents that object was there but it was not detected.

Table 1: Result comparison

Technique	[19]	Proposed
Precision	92.39%	94.73%
Recall	86.82%	88.33%
False Positive Rate	7.61%	5.27%

Table 1 shows the comparison of proposed result with already published result. To achieve these results Kalman filter was processed till 7th iteration. It can clearly be seen that better results were achieved by using the proposed technique.

IV. Conclusion

The proposed methodology is quite robust as only 7 iterations were made to achieve better results in comparison to previously used algorithms. A precision of 94.73% was achieved by the proposed methodology. The proposed methodology utilizes Global estimation and Kalman filter approach to correctly predict and track the object. Further before applying Kalman filter video was stabilized by an algorithm being discussed and proposed in the methodology.

V. References

[1] Ying, Y., Ji-Cheng, H.: Search-length based fast block-matching motion estimationalgorithm. Computer Engineering and Design 26(8), 2155–2157 (2005)

[2] Xu, L., Lin, X.: Digital image stabilization based on circular block matching. IEEETransactions on Consumer Electronics 52(2) (May 2007)

[3] Yang, A., Schonfeld, D., Mohamed, M.: Robust video stabilization based on particle filtering tracking of projected camera motion. IEEE Transactions on Circuits and Systems for Video Technology 19(7) (July 2009)

[4] Pinto, B., Anurenjan, P.R.: Video stabilization using speeded up robust features.In: International Conference on Communications and Signal Processing (ICCSP),pp. 527–531 (2011)A Robust Global Motion Estimation for Digital Video Stabilization 143

[5] Shen, Y., Parthasarathy, G., Thyagaraju, D., Bill, P.B., Kameswara, R.N.: Videostabilization using principal component analysis and scale invariant feature transformin particle filter framework. IEEE Trans. on Consumer Electronics 55(3)(August 2009).

[6] Ying, Y. et al. Search-length based fast block-matching motion estimation algorithm. Computer Engineering and Design 26(8), 2155–2157 (2005).

[7] Juanjuan Zhu,et al. Electronic image stabilization systembased on global feature tracking . Journal of Systems Engineering and Electronics, 2008, 19(2):228-233.

[8] S. Erturk, et al. "Image sequence stabilization based on DFT filtering,"IEE Proc. On Vision Image andSignal Processing, vol. 147, no. 2, pp. 95–102, (2000).

[9] Reddy, V. Purandhar, and A. Annis Fathima. "Object Tracking Based on Position Vectors and Pattern Matching." In Computational Signal Processing and Analysis, pp. 407-416. Springer, Singapore, 2018.

[10] Abrol, Marvi. "Neuro-Fuzzy Based Approach on Object Tracking." (2018).

[11] Chen, P., Dang, Y., Liang, R., Zhu, W. and He, X., 2018. Real-time object tracking on a drone with multi-inertial sensing data. IEEE Transactions on Intelligent Transportation Systems, 19(1), pp.131-139.

[12] Price, Eric, Guilherme Lawless, Roman Ludwig, Igor Martinovic, Heinrich H. Bülthoff, Michael Black, and Aamir Ahmad. "Deep neural network-based cooperative visual tracking through multiple micro aerial vehicles." IEEE Robotics and Automation Letters (2018).

[13] Du, Dawei, Yuankai Qi, Hongyang Yu, Yifan Yang, Kaiwen Duan, Guorong Li, Weigang Zhang, Qingming Huang, and Qi Tian. "The Unmanned Aerial Vehicle Benchmark: Object Detection and Tracking." arXiv preprint arXiv:1804.00518(2018).

[14] Koubâa, Anis, and B. Quershi. "Dronetrack: Cloud-based real-time object tracking using unmanned aerial vehicles." IEEE Access (2018).

[15] Dasgupta, S., Martirosyan, H., Koppula, H., Kendall, A., Stone, A., Donahoe, M., Bachrach, A. and Bry, A., Skydio, Inc., 2018. OBJECT TRACKING BY AN UNMANNED AERIAL VEHICLE USING VISUAL SENSORS. U.S. Patent Application 15/827,945.

[16] Nguyen, Bella, and Ioannis Brilakis. "Real-time validation of vision-based over-height vehicle detection system." Advanced Engineering Informatics 38 (2018): 67-80.

[17] Dasgupta, Saumitro, Hayk Martirosyan, Hema Koppula, Alex Kendall, Austin Stone, Matthew Donahoe, Abraham Bachrach, and Adam Bry. "OBJECT TRACKING BY AN UNMANNED AERIAL VEHICLE USING VISUAL SENSORS." U.S. Patent Application 15/827,945, filed June 7, 2018.

[18] DeAngelis, D.J., Evansen, E.G. and Reilly, G.M., ISOLYNX LLC,

[19] Dias, R., Cunha, B., Sousa, E., Azevedo, J. L., Silva, J., Amaral, F., & Lau, N. (2017, April). Real-time multi-object tracking on highly dynamic environments. In *Autonomous Robot Systems and Competitions (ICARSC), 2017 IEEE International Conference on* (pp. 178-183). IEEE.