Stabilization of jittery videos using matching techniques

S.Rajeswararao, (M.Tech),
Department of Electronics and Communication Engineering.,
Sri Sivani College Of Engineering,
Srikakulam, A.P, India

P. Mohan Krishna, M.Tech
Assistant professor,
Department of Electronics and Communication Engineering.,
Sri Sivani College Of Engineering,
Srikakulam, A.P, India

M. Ravi Kishore, M.Tech
Assistant professor,
Department of Electronics and Communication Engineering.,
Sri Sivani College Of Engineering,
Srikakulam, A.P, India

Abstract- This study proposed an algorithm to stabilize jittery videos directly without the need to estimate camera motion. The poor image quality of many videos of surveillance cameras will lead to unanticipated effects. This project introduces a new approach to stabilize image sequence, initially two frames are considered and salient points are identified and then optimized which provide stabilization to create a new video sequences where the motion between the frames has effectively been removed. The newly developed algorithm presented in this work provides a fast and robust stabilization method.

Keywords—*Video stabilization, image processing, jittery videos, corner detection, quality measurement.*

1. INTRODUCTION

Now a days due to the rapid growth of the usage of the handheld cameras, telescopes, lens in many applications. The captured video of the non professional users doesn't get a quality image. So to overcome this problem many researchers have developed many techniques. There are hardware equipment attached directly to the camera and also software like pre-processing and post-processing techniques, but still there are draw backs to overcome when multiple objects are moving in the background.

There are mainly three methods of image stabilization namely digital image stabilization, optical image stabilization, and mechanical image stabilization [1]. In digital image stabilization there are two methods one is zooming and the other is maintaining the lens of the camera larger than the image to compensate the image motion when camera moves in one direction the image is moved in the opposite direction. In optical image stabilization based on the deflection of the rays the CCD is set to move in the image direction. In mechanical image stabilization the mechanical components like gyroscope are used to stabilize the image.

The stabilization process goes through three phases namely motion estimation, motion smoothening and image composition [2].

In motion estimation the motion between the motion between the frames is calculated, and the high frequency distortions of output of first phase are removed and in the last phase wrapping is done by image composition. These are the basic steps of any video stabilization algorithms [3].

2. METHOD AND OVERVIEW

This flowchart presents an overview of the proposed algorithm step by step implementation in the below provided figure1

ISSN: 2278 – 909X

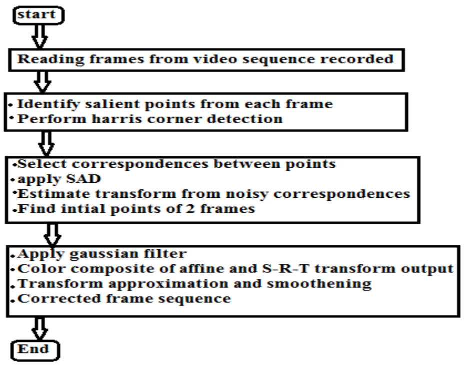

Figure 1: Overview of the proposed method of video stabilization.

A. Identifying feature points

Before performing the transformation the input for the transformation must be two frames with corresponding points between them.

First of all two successive frames must be considered and one can image relating two images by matching only locations in the image that are some way interesting[4]. Such points are called interest points and are located using an interest point detection algorithm.

B. Corner detection by using Harris corner detection

There are number of ways detection methods. Some detectors find points of high local symmetry, others find areas of highly varying texture, some corners. Corner points are interesting as they are formed from two or more edges there are number of corner detector methods like Moravec, Harris/Plessey, Trajkovic and Hedley [5]. In these the Harris corner detection method is fast and robust algorithm in detecting corner points.

B. Affine Transformation

In many imaging systems, detected images are subject to geometric distortions introduced by perspective irregularities where in the position of the camera with respect to the scene alters the apparent dimensions of the scene geometry [9].

Applying an affine transformation to a uniformly distorted image can correct for a range of perspective distortions by transforming the measurements from the ideal co-ordinates to those actually used. An affine transformation is an geometric transformation which maps variables (x_1,y_1) to (x_2,y_2) by applying a linear combination of translation, rotation, scaling and shearing.

C. Select correspondences between points

In this step we pick up correspondences between the points derived in the before step. For each point take 9 x 9 blocks. The matching cost we use between the points is SSD between their respective image regions. Points in frame A and B are matched. SSD is used to compare frame A and B to find the shortest distance between them in pixels [6].

3. RESULTS AND DISCUSSIONS

In this section the results are discussed based on the proposed methodology

A. Read frames from a video file

The data related intensity of images has been separated from colour. The intensity data is fed as input to stabilization algorithm. It is also observed that speed of processes has been increased drastically because of the usage of gray scale images. Two consecutive frames of images after processing are shown in the below figure 2 real cyan colour composite is produced to illustrate the pixel wise difference between them. It can be seen from the figure that there is a large vertical and horizontal offset between the two frames.

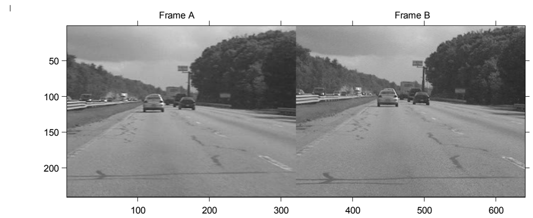

Figure 2: two successive frames taken frame A and frame B

D. Collect salient points from each frame

Our goal is to determine a transformation that will correct for the distortion between the two frames. We can use the geometric transform for this, which will return an affine transform. As input we must provide this function with a set of point correspondences between the two frames [8]. To generate these correspondences, we must first collect points of interest from both frames, which are salient points in the above selected frames as shown in the figure 3 below in two frames then select likely correspondences between them. As shown in the below figure 3

ISSN: 2278 – 909X

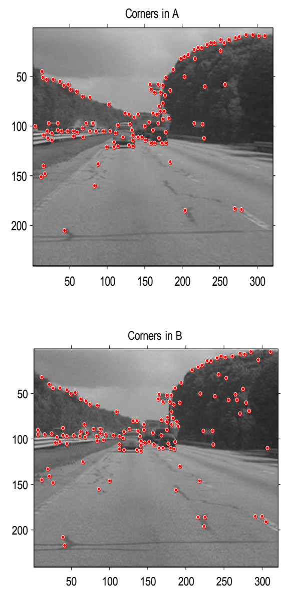

Figure 3: Salient points in frame A and frame B

B. Select correspondences between points

In this step we calculate the correspondences between points which are identified in the above section. For each point take 9 x 9 blocks. The matching cost we use between the points is SSD sum of squared differences between their respective image regions [7]. Points in frame A and frame B are matched. The below shown figure 4 depicts a colour composite, but added are the points from frame A in red, and the points from frame B in green, yellow lines are drawn between points to show the correspondences selected.

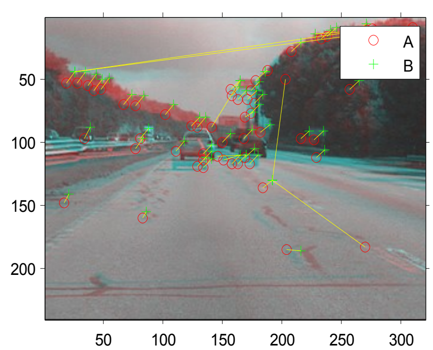

Figure 4: Correspondence points in both the frames represented by yellow lines

D .Estimating transform from noisy correspondences

Many of the points correspondences obtained are corrected but still we can derive a robust estimate of the geometric transform between the two images using Random sample consequence (RANSAC) algorithm.

This object when given a set of point's correspondences will search for the valid inliers correspondences. From these it will then derive the affine transform that makes the inliers from the first set of points match most closely with the inliers of the second set so we can calculate the following parameters like scaling, rotation, shearing, translation. This transform can be used to wrap the images such that their corresponding feature will be moved to the same image location [10]. The figure 6 depicts the output result of the following section as shown below.

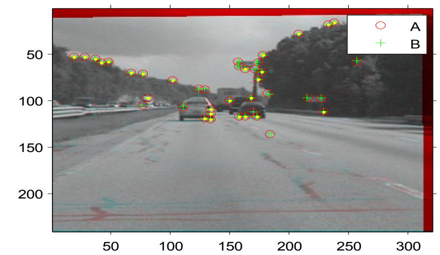

Figure 6: After the transformation is applied the correspondence points move to same position

E. Transform Approximation and Smoothening

In this section we construct a new S-R-T transforms as given in the steps below:

- We extract scale and rotations part of sub-matrix from affine transformation of matrix 3 x 3
- Compute the data from mean of two possible arctangents.
- Translation is remain same as original S-R-T transform
- Finally reconstruct new S-R-T transform

The output of the above following steps is shown in the below figure 7

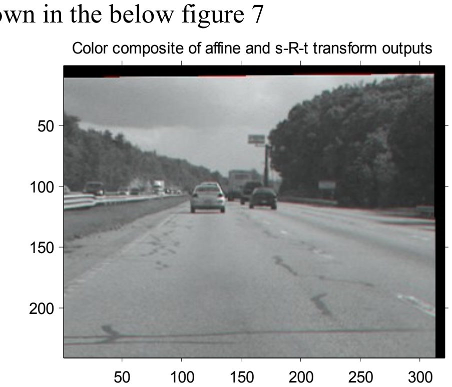

Figure 7: Colour Composite of affine and S-R-T transform outputs

F. Corrected frame sequence

ISSN: 2278 – 909X

During the computation, we computed the mean of raw video frames and of the corrected frames. These mean values are shown below in the figure 8. The left image shows the mean of the raw input frames, proving that there was a great deal of distortion in the original video [12]. The mean of the corrected frames on the right which shows with almost no distortions.

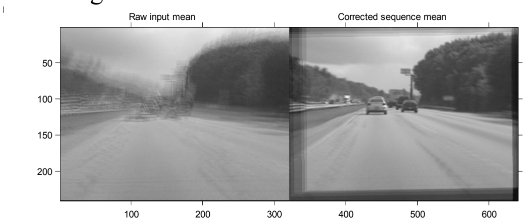

Figure 8: correct frame sequence according to RANSAC

G. Measurements

The output video quality is also measured based on the proposed methods. This is evaluated based on SVD base gray scale image value and graphical measurement [11].

i) *SVD based gray scale image quality* for accurate measurement and measuring the value in 2D measurement a new measurement method is used that is Singular value decomposition(SVD) which expresses the quality of the distorted images graphically or numerically as a scalar measurement, both near and above the visual threshold[13]. Equation 1 represents the computed value:

$$M - SVD = \frac{\sum_{i=1}^{(\frac{k}{n})x(\frac{k}{n})} |D_i - D_{mid}|}{(\frac{k}{n})x(\frac{k}{n})}$$

Where:

D_{mid} represents the midpoint of the sorted D_i

K is the image size

N is the block size

M-SVD is the measurement of singular value decomposition

The below figure 9 shows the quality video of the calculated based on the above value

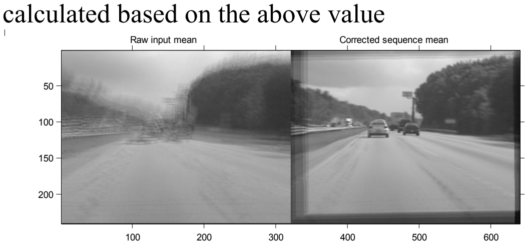

Figure 9: The quality video of the raw input

ii) *Graphical* Measurement

Graphical measurement will indicate the condition of video due to distortion the graphical measurement of our video is shown in the below figure 10:

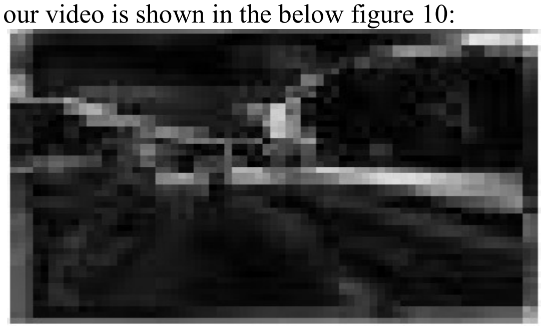

Figure 10: Graphical measurement of the video

For example if we consider our video and calculate the numerical value and quality value for a sample picture as shown in the figure 11 then the values would be like this:

Figure 11: the video to which the numerical value and quality value are measured

Computational speed: 8.078

Numerical measure: 30.4%

Quality value: 69.6%

4 .Conclusion and Future scope

The proposed technique is logically and computationally efficient approach in terms of stabilizing high jittery videos suffered from distortion. This technique should prove useful in enhancing the quality of low grade video surveillance cameras. This technique is particularly helpful in identifying people, license plate etc. From low quality video surveillance camera.

In future, we can find better feature detector and overcome the consequences of extreme shaking of handheld camera in feasible real time implementation for video stabilization.

Comparison between the existing technique and proposed technique:

Parameter value	Existing technique	Proposed technique
Computational value(sec)	9.28	8.078
Quality value	40.50%	69.6%

ISSN: 2278 – 909X

REFERENCES

[1]Video stabilization and enhancement by Hany Farid and Jeffrey B. Woodward, Department of Computer Science, Darthmouth College, Hanover NH 03755.

[2]A thesis on onboard video stabilization for unmanned air vehicles by Nicholas S. Cross.

[3]M. Gleicher and F. Liu.,"Re-cinematography: improving the camerawork of casual video,"ACM Transactions on multimedia computing, Communications, and Applications, 2008.

[4]Aleksander Shnayderman, Alexander Gusev, and Ahmet M. Eskicioglu "An SVD-based gray scale Image Quality Measure for Local and Global Assesment", IEEE 2006.

[5]A thesis on Fast Video Stabilization Algorithm by Mohammed A. Alharbi, Captain, RSAF AFIT/GCS/ENG/06-02.

[6]C. Morimato and R. Chelleppa, "Evaluation of Image Stabilization algorithms", Proc. of IEEE int. Conf. on Acoustics, speech and signal processing, Vol-5,pages 2789-2792, 1995.

[7]C. Morimoto, D. De Menthon, L. Davis, R. Nelson "Detection of independently moving objects in passive video", In Proc. of Intelligent Vehicles workshop, pages 270-275, Detroit, MI, September 1995.

[8]Z. Duric and A. Rosenfield "stabilization of images sequences" Technical report CAR-TR-778, Center for Automation Research University of Maryland, College, Park, 1995.

[9]T. Vieville, E. Clergue, and P. Facao., "Computation of ego-motion and structure from visual and internal
Sensors using the vertical Cue", In. Proc. International Conference on Computer Vision, Pages 591-598, Berlin, Germany, 1993.

[10]M. Irane, B. Rousso and S. peleg., "Recovery of ego-motion using image stabilization., In .Proc of IEEE Conference on Computer Vision and Pattern Recognition, pages 454-460, Seattle, WA, June 1994.

[11]O. Kwon, R. Chelleppa, and C. Morimoto.,
'Motion Compensated Sub band Coding of Video acquired from a moving platform" In. Proc. of IEEE International conference on Acoustics, Speech, and Signal Processing, pages 2185-2188, Detroit, MI, January 1995.

[12]P. Brut and P. Anandan., "Image Stabilization by registration to a reference mosaic", In Proc. DARPA Image Understanding Workshop, pages 425-434, Monterey, CA, November 1994.

[13]M. Hansen, P. Anandan, K. Dana, G. Vanderwal, and P. J. Brut., "real-time Scene stabilization and Mosaic Construction". In. Proc. DAPRA Image Understanding Workshop pages 457-465, Moneterey, CA, November 1994.

[14]http://www.mathworks.com.

ISSN: 2278 – 909X