(IJAER) 2015, Vol. No. 10, Issue No. I, July e-ISSN: 2231-5152/ p-ISSN: 2454-1796

A NOVEL VIDEO STABILIZATION QUALITY ASSESSMENT ALGORITHM

H.NAGA VENI, M.V.H.BHASKARA MURTHY

Department of Electronics And Communication Engineering.

Aditya Institute of Technology And Management.

Tekkali, Andhra pradesh, 532201, INDIA.

ABSTRACT

The main objective of this paper is to introduce a new approach to stabilize image sequence. A logical and computationally efficient technique for video stabilization and enhancement has been presented. The recent significant advance in video stabilization is to create a new video sequence where the motion between frames has effectively been removed. Motion in video images is caused either by the object motion or by the camera movement. The newly developed algorithm presented in this work provides a fast and robust stabilization system, and improves real-time performance.

key words: — image processing, video stabilization image quality measurement

I .INTRODUCTION

The usage of handheld smart cameras has increased unprecedetly in the recent past in the field of video signal acquiry and capture. The role of cameras in the process of video capturing has been reported by [1]. It has been observed that usage of video cameras will lead to unanticipated effects such as image distortion, image blurring etc, if operated by nonprofessionals. The quality of captured video is found to be a function of the speed of the vehicle. While the motion of the vehicle is smooth, the acquired image has been reflecting the actual images with acceptable accuracy. The higher degree of jitter and vibrations are common while acquiring the images using small unnamed aerial imaging system and off road navigation of ground vehicles equipped with onboard cameras. Consequently, the video images acquired from these platforms have to be preprocessed to eliminate the jitter variations before human analysis. The stabilization algorithm involves two steps. The first step comprises of determining the affine image transformations between all neighboring frames of a video sequence using a Random Sampling and Consensus (RANSAC) procedure applied to point correspondences between two images. In the second step, wrapping of video frames will be done to achieve a stabilized video.

II. PREVIOUS WORK

Labeeb Mohsin Abdullah, Nooritawati Md Tahir and Mustaffa Samad [1]had proposed an algorithm to stabilize the image sequence by using Harris corner detection technique. Corner Detector System Object is used to find corner values using Harris Corner Detection which is one of the fastest algorithms to find corner values.

Aleksandr Shnayderman, Alexander Gusev, and Ahmet M. Eskicioglu had proposed to measuring SVD based gray scale image quality[2]. This is evaluated based on SVD based grayscale Image value and graphical measurement

Edward Rosten, Reid Porter and Tom Drummond, proposed "FASTER and better A machine learning approach to corner detection ".After the salient points from each frame are obtained the Correspondence between the points that are identified previously need to be picked [3]. For each point, the matching of lowest cost between the points that existed in frame A and B are also needed to be found for all points. Hence, it is necessary to divide the sequence of frames image into 9×9 block. The matching cost means the distance between frame A and B measured in pixel. To find this cost, the technique of Sum of Squared Differences (SSD) can be used between the consecutive frame images

FAST is an algorithm proposed originally by Rosten and Drummond [4] for identifying Interest points in an image.

Yue Wang, ZuJun Hou, Karianto Leman and Richard Chang had proposed "Real-Time Video Stabilization for Unmanned Aerial Vehicles[5]. the keypoints are located based on FAST corner detection and preliminarily matched. Secondly, the matched keypoints are then involved for estimation of affine transform to reduce false matching keypoints.

Elmar Mair Gregory D. Hager, Darius Burschka, Michael Suppa, and Gerhard Hirzinger" has proposed technique Adaptive and Generic Corner Detection Based on the Accelerated Segment Test"[6]

C. Harris and M.J. Stephens, "A combined corner and edge

The output video quality is also measured based on the proposed methods[7]. To cater for image regions containing texture and isolated features, a combined corner and edge detector based on the local auto-correlation function is utilised

Mohammed A. Alharbi, Captain, RSAF had proposed A thesis on Fast Video Stabilization Algorithms[8] An affine motion model is utilized to determine the parameters of translation and rotation between images. The determined affine transformation is then exploited to compensate for the abrupt temporal discontinuities of input image sequences

Harris corner detection has implemented for identifying salient points from each frame and achieved the measured Quality value is 40.50% and computational time is upto 9.28s only.

(IJAER) 2015, Vol. No. 10, Issue No. I, July e-ISSN: 2231-5152/ p-ISSN: 2454-1796

III. PROPOSED METHOD.

In this section, the overview of the proposed method of video stabilization and the aspects of implementation of the algorithm has been described and is shown in figure1.

A. Read Frames From A Video File:

The data related intensity of images has been separated from color. The intensity data is fed as input to the 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 using above criteria are shown in figure 2. A real cyan color composite is produced to illustrate the pixel wise difference between them. It can be read from the frames represented in figure 2 that there is obviously a large vertical and horizontal offset between the two frames

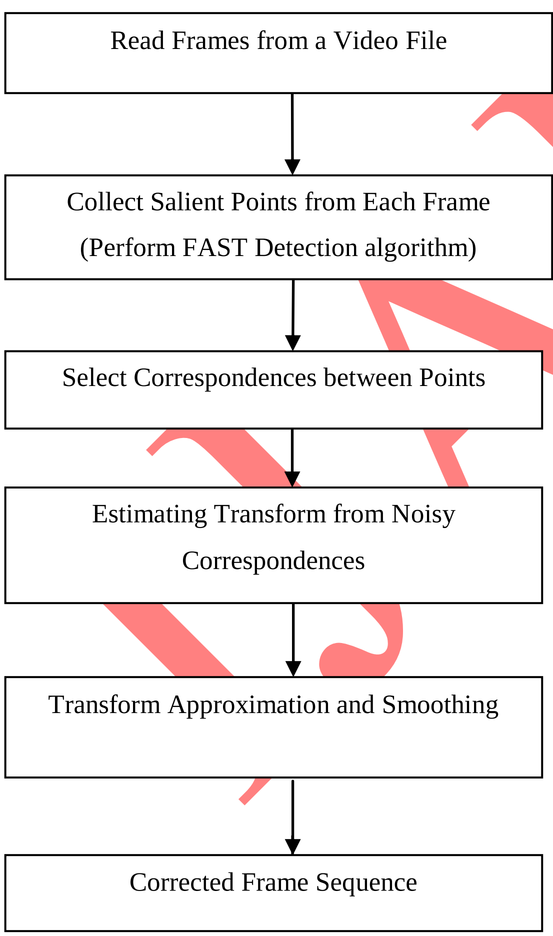


Figure 1: Flow chart of the proposed method of video stabilization

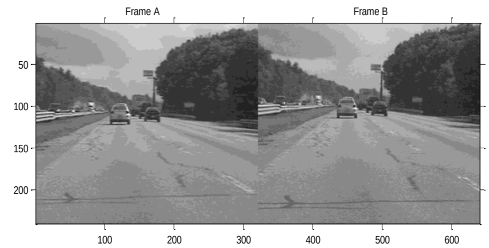


Fig 2: Reading the two frames from a video sequence

B .Collection of Salient Points From Each Frame:

In the process of determination of a transformation that will correct for the distortion between the two frames, the Geometric Transform can be used, which return an affine transform. The input for this function shall comprise a set of point correspondences between the two frames. To generate these correspondences, collection of points of interest from both frames is performed, and then selection likely correspondences between them are done.

In this step we produce these candidate points for each frame. To have the best chance that these points will have corresponding points in the other frame, points around salient image features such as corners are to be identified. For this purpose, there is several corner detection algorithms are established like: Moravec corner detection algorithm, Harris & Stephens corner detection algorithm, SUSAN corner detector.

The reason behind the work of the FAST algorithm was to develop an interest point detector for use in real time frame rate applications like SLAM on a mobile robot, which have limited computational resource.

High measured quality video and less computational time than base paper achieved by proposing an another matching technique named as FAST (Features from Accelerated Segment Test) detection

FAST Detection Algorithm

FAST is an algorithm proposed originally by Rosten and Drummond [4] for identifying Interest points in an image and is shown in figure3.. An interest point in an image is a pixel which has a well-defined position and can be robustly detected. Interest points have high local information content and they should be ideally repeatable between different images [10]. Interest point detection has applications in image matching, object recognition, tracking etc.

Features detection using FAST:

1. Select a pixel "p" in the image. Assume the intensity of this pixel to be I_P. This is the pixel which is to be identified as an interest point or not.
2. Set a threshold intensity value T, (say 20% of the pixel under test).
3. Consider a circle of 16 pixels surrounding the pixel p.

4. "N" contiguous pixels out of the 16 need to be either above or below I_P by the value T, if the pixel needs to be detected as an interest point.
5. To make the algorithm fast, first compare the intensity of pixels 1, 5, 9 and 13 of the circle with I_P. As evident from the figure above, at least three of these four pixels should satisfy the Threshold criterion so that the interest point will exist.
6. If at least three of the four pixel values - I_1, I_5, I_9, I_{13} are not above or below I_P + T, then P is not an interest point (corner). In this case reject the pixel p as a possible interest point. Else if at least three of the pixels are above or below I_P + T, then check for all 16 pixels and check if 12 contiguous pixels fall in the criterion.
7. Repeat the procedure for all the pixels in the image

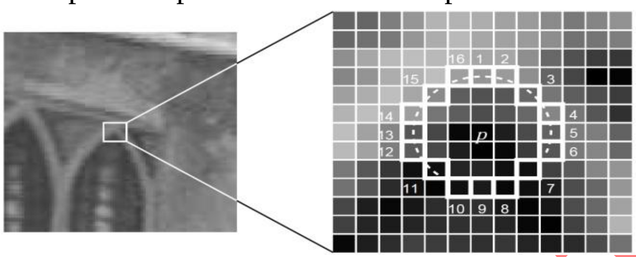


Figure 3: Image showing the interest point under test and the 16 pixels on the circle

The detected points from both frames are shown in the figure4. Observe how many of them cover the same image features, such as points along the tree line, the corners of the large road sign, and the corners of the cars.

Collection of salient from frame is shown Figure 4.

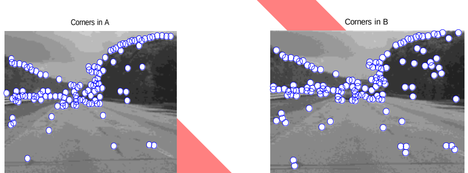


Figure 4: Collection of salient points from both frames

In the process of determination of a transformation that will correct for the distortion between the two frames, the Geometric Transform can be used, which return an affine transform. The input for this function shall comprise a set of point correspondences between the two frames. Togenerate

C. Select Correspondences between Points:

An interest point in an image is a pixel which has a well-defined position and can be robustly detected. Interest points have high local information content and they should be ideally repeatable between different images Interest point detection has applications in image matching, object recognition, tracking etc.

In this step, correspondences between the points derived in the step B have to be picking up and consolidation shall be done. For each point, we extract a 9-by-9 block centered on it. The matching cost that is used between these points is the sum of squared differences (SSD) between their respective image regions.

The figure5 shows the same color composite of figure 6, 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 by the above procedure. Many of these correspondences are correct, but there are also a significant number of outliers.

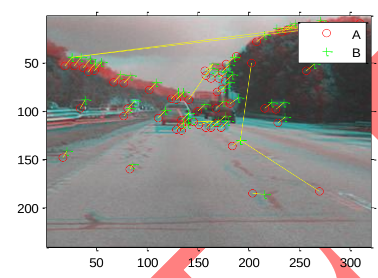


Figure 5: Select the Correspondences Points from Both Frames

D. Estimating Transform from Noisy Correspondences

Many of the point correspondences obtained in the previous step are identified with limited accuracy. To rectify this problem, a new robust estimate of the geometric transform between the two images using the Random Sample Consensus (RANSAC) algorithm, which is implemented in the Geometric Transform Estimator System object is developed and presented.

Affine Transform: It is any transformation that can be expressed in the form of a matrix multiplication (linear transformation) followed by a vector addition (translation).

We can use an Affine Transformation to express:

a)Rotations (linear transformation)

b)Translations (vector addition)

c)Scale operations (linear transformation)

In essence, an Affine Transformation represents a relation between two images.

The methodology followed to perform the implementation of above technique is as follows. For added robustness, the Geometric Transform Estimator System object will be run multiple times to calculate a cost for each result. This cost is obtained by projecting frame B onto frame A according to the derived transform, and taking the sum of absolute difference (SAD) between the two images. From this, the best transform as the one that minimizes this cost is taken as outcome.

Figure 6 shows a color composite showing frame A overlaid with the re-projected frame B, along with the re-projected point correspondences. It is clear from this figure that the results are favorable, with the inliers correspondences nearly exactly coincident. The cores of the images are both well aligned, such that the red-cyan color composite becomes almost purely black-and-white in that region.

Further, it can also be observed from this figure that how the inliers correspondences are all in the background of the image, not in the foreground, which itself is not aligned. This is because the background features are distant enough that they behave as if they were on an infinitely distant plane. Thus, even though the affine transform is limited to altering only the imaging plane, in the proposed technique it is sufficient to align the background planes of both images. Furthermore, if it is assumed that the background plane has not moved or changed significantly between frames, then this transform is actually capturing the camera motion. Therefore correction for this will stabilize the video.

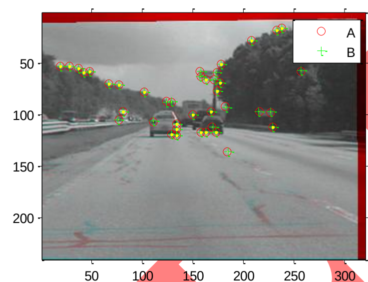


Figure 6: color composite showing frame A overlaid with the reprojected frame B

E. Transform Approximation and Smoothing

In this step, construction of a new S-R-T transforms is performed. For this

- Extraction of scale and rotation part of sub- matrix from affine transform of matrix 3 by 3 as shown in the above step.
- Compute theta from mean of two possible arctangents.
- Compute scale from mean of two stable mean calculations
- Translation will remain same as original S-R-T transform and
- Finally reconstruct new S-R-t transform and is shown in Figure 7.

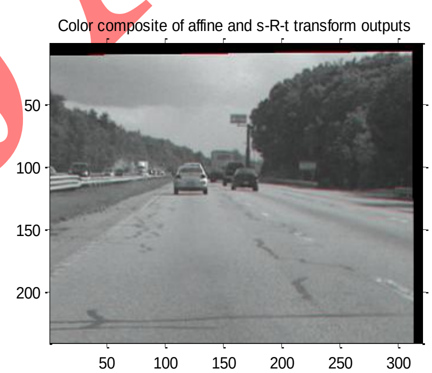


Figure 7: Color Composite of Affine and S-R-T Transform

F. Corrected Frame Sequence:

During computation, computation of the mean of the raw video frames and of the corrected frames was done. These mean values are shown side-by-side in Figure 8 for better comparison. The left image shows the mean of the raw input frames, proving that there was a great deal of distortion in the original video. The mean of the corrected frames on the right, however, shows the image core with almost no distortion. While foreground details have been blurred this shows the efficacy of the stabilization algorithm.

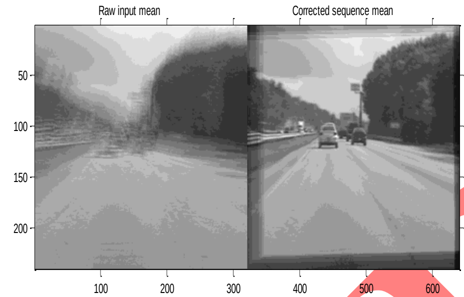


Figure 8: Corrected frame sequence of the jittery video

IV.RESULTS AND DISCUSSIONS

MEASUREMENTS

The output video quality is also measured based on the proposed methods. This is evaluated based on Singular Value Decomposition (SVD) based grayscale Image value and graphical measurement.

- **SVD Based Grayscale Image Quality**

A. Graphical Measure

The proposed graphical measure is a bivariate measure that computes the distance between the singular values of the original image block and the singular values of the distorted image block.

$$D_i = Sqrt\left[\sum_{i=1}^{n}(s_i - \hat{s}_i)^2\right]$$

Where s_i is the singular value of the original is block and $\hat{s}_i$ is the singular values of the distorted block, and 'n' is the block size. If the image size is k, we have $\left(\frac{k}{n}\right)\times\left(\frac{k}{n}\right)$ blocks. The set of distances, when displayed in a graph, represents a "distortion map." The block size used in our experiments is 8×8 for two reasons: It is a common block size in JPEG compression and other image processing applications. The graphical measurement for a sample video frame is shown in figure 9.

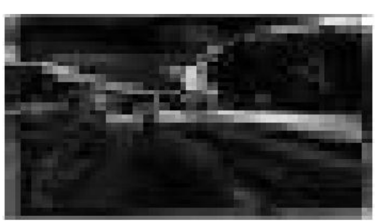


Figure 9: Graphical measurement for each sample video

B. Numerical Measure

Singular value decomposition (SVD) is developed as a new measurement that can express the quality of distorted images either graphically that is in 2D measurement or numerically as a scalar measurement, both near and above the visual threshold**.** The experiments here utilized SVD based measurement that outperformed the normally used PSNR [4] and corresponding results as shown in table 1.

The numerical measure is derived from the graphical measure. It computes the global error expressed as a single numerical value depending on the distortion type

$$M\_SVD = \frac{\sum_{i=1}^{\left(\frac{k}{n}\right)\times\left(\frac{k}{n}\right)} |D - D_{mid}|}{\left(\left(\frac{k}{n}\right)\times\left(\frac{k}{n}\right)\right)}$$

Where:

D_{mid} represents the midpoint of the sorted D’s

k is the image size

n is the block size

M_SVD is the measurement of Singular value decomposition.

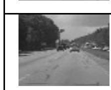
Sample frame	Feature detection	Computational time(s)	Error value	Quality value
	Harris corner	9.28	59.50 %	40.50 %
	Fast corner	8.63	31.16 %	69.84 %

Table I: Criteria of sample video & Results

V.CONCLUSION AND FUTURE SCOPE

The proposed technique is logical and computationally efficient approach in terms of stabilizing high jittery videos suffered from distortion compared to existing technique. The table 1 conclude that Fast corner detection technique for video stabilization gives better quality than Harris corner detection method. This technique should prove useful in enhancing the quality of low-grade video surveillance cameras.

In future, we can find better feature detector and to overcome the consequences of extreme shaking of handheld camera in real time implementation for video stabilization. This technique is particularly helpful in identifying people, license plates, etc. from low-quality video surveillance cameras.

REFERENCES

1. Labeeb Mohsin Abdullah, Nooritawati Md Tahir & Mustaffa Samad , "Video Stabilization based on Point Feature Matching Technique", 2012 IEEE Control and System Graduate Research Colloquium (ICSGRC 2012).
2. Aleksandra Shnayderman, Alexander Gusev, and Ahmet M. Eskicioglu"An SVD-Based Grayscale Image Quality Measure for Local and GlobalAssessment ",IEEE 15(2), 2006.
3. Edward Rosten, Reid Porter and Tom Drummond, "FASTER and better: A machine learning approach to corner detection" in *IEEE Trans. Pattern Analysis and Machine Intelligence, 2010, vol* 32, pp. 105-119
4. E. Rosten and T. Drummond, "Machine learning for high speed corner detection," in 9th Euproean Conference on Computer Vision, vol. 1, 2006, pp. 430–443
5. Yue Wang, ZuJun Hou, Karianto Leman and Richard Chang "Real-Time Video Stabilization for Unmanned Aerial Vehicles" inMVA2011 IAPR Conference on Machine Vision Applications, June 13-15, 2011, Nara, JAPAN
6. Elmar Mair Gregory D. Hager, Darius Burschka, Michael Suppa, and Gerhard Hirzinger" Adaptive and Generic Corner Detection Based on the Accelerated Segment Test"
7. C. Harris and M.J. Stephens, "A combined corner and edge detector",Proc of Alvey Vision Conference, pp 147–152, 1988.
8. A thesis on Fast Video Stabilization Algorithms by Mohammed A. Alharbi, Captain, RSAF AFIT/GCS/ENG/06-02.