

QUALITATIVE ASSESSMENT OF VIDEO STABILIZATION AND MOSAICKING SYSTEMS

Chao Zhang, Prakash Chockalingam, Ankit Kumar,
Peter Burt, and Arvind Lakshmikumar

Sarnoff Corporation
201 Washington Road, Princeton, NJ 08540, USA

Abstract — Image stabilization is a key preprocessing step in dynamic image analysis, which deals with the removal of unwanted motion in a video sequence. It is principally understood as the warping of video sequences resulting in a total or partial removal of image motion. Stabilization is invaluable for motion analysis, structure from motion, independent motion detection, geo-registration and mosaicking, autonomous vehicle navigation, model-based compression, and many others. Given the usefulness of image stabilization for many applications, a variety of algorithms have been proposed to perform this task, and many real-time systems have been built to stabilize the real-time videos and provide motion data for tracking and geo-registrations. However, even though there are on-line libraries that provide test videos, there has been no established methods or industrial standards based on which the performance of a stabilization algorithm or system can be measured. This paper aims to address this gap and suggests an evaluation methodology which would provide us the ability to qualitatively measure the performance of a given stabilized system. We propose a performance measurement system and define the performance metrics in this paper. We then apply the assessment to two typical stabilization systems. The discussed methods can be used to benchmark video stabilization systems.

Index Terms — Evaluation, image registration, image restoration, motion analysis

1. INTRODUCTION

Image registration is one of the fundamental preprocessing steps in image processing [1]. It is used to match two or more images of the same scene taken at different times, from different viewpoints, and/or from different sensors. Image stabilization is commonly regarded as an image registration of the video sequence from the same sensor over time, which results in the total or partial removal of the image motion due to unwanted camera motion. Image mosaicking is a direct and cascaded result of stabilization. It reconstructs a larger image from the alignment between pairs of frames. The quality of the alignment directly affects the performance of the mosaicking result.

In the last few years, electronic stabilization systems have emerged for aerial video surveillance, indoor/outdoor object tracking and border monitoring. Their real-time performances have been benefited from the advanced development of image registration algorithms, and continuing efforts of embedding software into hardware. Sarnoff's Acadia PCI vision processing board [2] is the first commercial product that has offered image stabilization, mosaicking and fusion for NTSC and PAL systems. Stable Eyes [3], the 2004 winner of the IFSEC Security Industry Award, is a video stabilization product

offering from UK based Ovation Inc. The stabilization algorithms in their system purport to compensate for demanding image stabilization situation such as noisy video sources, low-light cameras, movement of large objects within the scene and operations with PTZ cameras. DynaPel Systems Inc [4] released the SteadyEye digital real-time video stabilizer in 2005. Their system can be mounted on ships and vehicles, tall buildings, airplanes, highway posts to smooth out shaky videos. In the same year, another company called Barco launched their real-time EO sensor video processing FlexiVision III [5]. The FlexiVision III family provided image stabilization, and image fusion.

With lot of commercial systems in the market, there is an immediate need for the evaluation of the stabilization systems. In the past, a few papers have looked at evaluating the performance of video stabilization systems. Balakirsky et al. [6] compared the performance of different stabilization algorithms based on the accuracy of a real-time object tracker. Morimoto et al. [7] considered the maximum displacement velocity in pixels per second, computed as the product of frame-rate and maximum image displacement between frames. They also evaluated the fidelity of image stabilization techniques by using the peak SNR between stabilized frames [8]. The problem with using these methods to evaluate alignment of successive frames is that, even if perfectly aligned, successive frames will not match perfectly due to frame-to-frame changes in global illumination, camera response, and aliasing. Skerl et al. [9] has discussed a similarity measure in medical imaging, which can be used for image-to-image registration between different modalities, or same modality but different lighting conditions. West et al. [10] also discussed the evaluation of image-to-image registration of different modalities between CT, MR, and PET. The validity of their performance depends on the *golden standard registrations* that were obtained from simulations. Zitova et al. [1] in their survey paper, discussed localization error, matching error, and alignment error. They preferred *golden standard method* as a comparative method and probably the best method to be used in the application area. The method plays a role similar to the ground truth. However, unlike typical medical image registrations that align non-rigid multi-modal images, a video stabilization system works on temporal image

sequences in a simpler motion model. The system focuses on removing camera jitters in a temporally coherent way. The system may not necessarily align the consecutive images with respect to the same reference image; it may be designed to reduce high frequency jitter motion, while keeping the low frequency motions. We felt that there has been no established method based on which the performance of stabilization can be measured. The key contribution of this paper is to define the metrics that could rationalize the performance of stabilization algorithms and systems and provide the methodology for customer assessment and industrial standard.

In Section 2, we will setup a performance measurement system. The performance metrics, engineering notions and mathematical formulas of the measurement metrics will be discussed in Section 3. In Section 4, based on the proposed metrics, we evaluate the performance of the intensity based algorithm on an Acadia I PCI Vision Accelerator board and a feature based method which uses SIFT features and RANSAC to eliminate outliers and fit the homography. The last section will conclude our discussion of the evaluation methodology and suggest further study based on this approach.

Since we focus on the evaluation methodology, this paper does not compare the performance of the Acadia board with other stabilization systems. Rather, we pick two typical stabilization systems that respectively fall into two categories: the intensity-based method, and the feature-based method, and hope our discussion may cover most of the systems available on the market.

2. PERFORMANCE EVALUATION SYSTEM

An electronic stabilization system should have one primary input that takes the jitter video, and one primary output that removes the fluctuation of the camera motion. Most commonly, stabilization is achieved by warping each frame of the source video to align to a prior frame. Alternatively all images can be aligned to a separately provided reference image. To generate a mosaic, all images need to align to a common anchor reference. Since there may be no overlap between the anchor reference and many of the images in the video sequence, the electronic stabilization systems compute the alignment from the current image to a neighbor reference image, and cascade the motion when reference is updated due to the need to maintain enough overlap in alignment.

The stabilization system has three basic processing components: an offset estimator, an intended motion estimator and a warper as shown in Fig. 1. The offset estimator estimates the offset between a current video frame and a prior frame, a neighbor reference or an external reference. This *observed* offset is stated as a vector displacement within an assumed image motion model (e.g., translation, rotation, zoom, affine, etc.). A separate offset is

generated for each image frame. A confidence measure may be generated with each such estimate indicating the likely error bounds for the measure. The intended motion estimator estimates that portion of an observed offset that is likely due to intended camera motion and that portion likely to be due to unintended camera vibration or instability. Again, a separate input may indicate intended motion or the module can estimate this from the sequence of observed offsets. The module generates a *correction* offset for each frame that represents the difference between the observed offset and the estimated intended offset. The warper warps each input image frame from its observed position to its (estimated) intended position. This becomes a frame in the output video. Warping entails resampling the source image to the desired output sample grid. A good quality interpolation function is required to maintain image sharpness and, if the source is interlaced, to avoid aliasing.

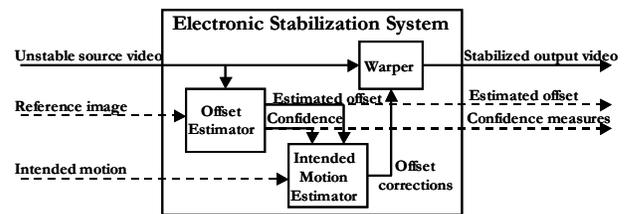


Fig. 1. An electronic stabilization system has three basic processing components: an offset estimator, an intended motion estimator and a warper. The unstable source video is aligned to the reference image by the offset estimator, and warped to the output by the warper. The intended motion estimator is used to correct the offset based on the intended motion.

Our assessment of how well a given stabilization algorithm performs entails determining how accurately and reliably it estimates image offsets, how well it discriminates intended motion from unwanted components of motion and how well it maintains image quality during warps. A basic method for measuring the performance of a stabilization system is suggested in Fig. 2. The measurement procedure begins with the construction of test sequences for which unwanted motion components are known precisely. This is done by introducing defined instabilities into stable reference sequences. The stable reference is provided to the measurement system as a sequence of frames. The instabilities are provided as a corresponding sequence of offsets, one from each frame. The unstable test sequence is then obtained within the system by warping each reference frame in accordance with the defined offset. Gaussian noises or Salt and Pepper noises can be added to the test sequence before it is provided to the stabilization system under study. This system returns a stabilized output sequence. It may also provide its estimates of the unwanted offsets as an output.

To avoid black borders when warping from a reference sequence based on the defined offset, we generate the unstable test sequence from a high-resolution image. A16-

mega-pixel image of Bridgewater, NJ is used to generate a test sequence. 720x480 (NTSC) or 800x800 are the two common sizes we have used for the performance test.

The generation of the instability offset sequence should be designed to compensate for motion at several levels of complexity. The simplest case is that of a camera mounted on a pole. The instability can often be modeled as simple translations. If the camera is able to pan and tilt, and zoom, then such motion can be modeled as similarity. For aerial video surveillance, we consider higher order motion types, the affine motion or the projective motion. Some systems may smooth out the high frequency jitter while keep the low frequency motion. Some of them compensate for pan, rotation, zoom and shear as well as jitter, then periodically reset with the accumulated displacements being compensated become too large.

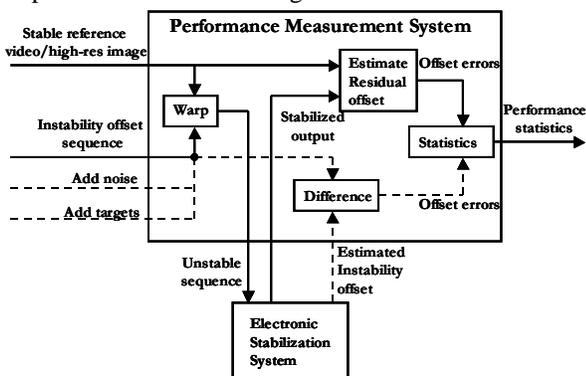


Fig. 2. A performance evaluation system for the electronic stabilization. The performance is estimated either from the residual offset between the reference image and the stabilized output, or between the true instability offset and the estimated offset from the stabilization system.

The tolerances of noises and foreground targets are other important indices to the image stabilization systems. Based on different camera modalities, Gaussian noises, or Salt-and-Pepper noises can be used to study the robustness of the stabilization system. Different level of noises adds to the designed offset sequence. The performance of the system becomes a function of the noise level. The performance is associated with the ratio of the foreground energy over the background energy as well.

For the most part we will be concerned with stabilization systems designed for surveillance applications. These typically compensate up to affine motions or just for translational jitter; they also need to stabilize a scene when there are noises from the camera or there are small foreground moving objects. Hence, we focus on these cases with our baseline measurement systems.

3. PERFORMANCE METRICS

To measure the performance of a stabilization system, we propose the following metrics:

- **Accuracy** – This metric represents the accuracy of the unwanted components of motion estimated and removed. The designed sequence should test the mean error of frame-to-frame motion in a sequence, and should also test the drift error from the cascaded motions when reference images are periodically updated due to large displacements (Sec 4.2).
- **Capture range** - It represents the maximum limit of the frame to reference motion. For feature-based methods, the feature points matching can be done in a small overlap region. This implies a large capture range for the feature-based methods. For global methods, larger overlaps are required (Sec 4.1). For an unknown stabilization system, the capture range defines a threshold beyond which the confidence measure of the alignment is low (Sec 4.7).
- **Robustness** - This metric reflects the reliability of the system over different imaging conditions. The designed sequence should test the accuracy of alignment when image feature varies, or degree of motion changes, or noise level increases (Sec 4.5). It may also test the accuracy when the complicated motion model such as affine parameters is defined as random numbers (Sec 4.3).
- **Tolerance for foreground moving objects** - This metric captures the object size and speed limit till which the objects still maintain a lock on the background. The foreground object is designed to have increased features in the input sequence. The stabilization test would show a threshold below which the alignment is performing well (Sec 4.6).
- **Ability to separate wanted from unwanted components of motion** - The metric reflects the ability of the system to remove jitters while preserving low frequency motions from pans and rotations (Sec 4.4).

The performance statistics would include a single frame-to-frame measure, average frame-to-frame measure, the confidence measure, and the overall performance measure. Since we are using affine motion as the representative motion for stabilization systems, we define the mean-square error (*MSE*) between the estimated motion and the ground truth offsets:

$$MSE(t) = (d_{13}(t) - \hat{d}_{13}(t))^2 + (d_{23}(t) - \hat{d}_{23}(t))^2 + (d_{11}(t) - \hat{d}_{11}(t))^2 \cdot x_c^2 + (d_{12}(t) - \hat{d}_{12}(t))^2 \cdot y_c^2 + (d_{21}(t) - \hat{d}_{21}(t))^2 \cdot x_c^2 + (d_{22}(t) - \hat{d}_{22}(t))^2 \cdot y_c^2, \quad (1)$$

where t is the t^{th} frame of the sequence, d_{ij} is the component of the estimated motion, and \hat{d}_{ij} is the ground truth motion. x_c and y_c are the image center. As a generic affine motion, d_{ij} and \hat{d}_{ij} satisfies the following relation:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{pmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \end{pmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (2)$$

The *MSE* becomes the single frame-to-frame alignment measure. The *RMS* error is the average of the square root of *MSE* over the whole sequence. If multiple sequences (say r

sequences) are used, then the average is over all the sequences as well:

$$RMS = \frac{1}{N_r N_t} \sum_{r,t} \sqrt{MSE(r,t)} \quad (3)$$

where N_r represents the total number of sequences and N_t represents the total number of frames in a sequence.

4. EXPERIMENTS

We use the intensity-based stabilization program on the Acadia PCI board [11,12] and a feature based method as an example to analyze the system performance. The tests cover the performance metrics discussed in the last section. Similar to the golden standard methods, we generate various video sequences with ground truth. Section 4.1, 4.2 and 4.3 test the alignment accuracy under different motion model. Section 4.2 also tests the drift error of the system if cascading motion is frequently requested for mosaicking in aerial surveillance. Section 4.4 tests the capability of the system to smooth out high frequency jitter while keeping the low frequency motion. Section 4.5 and 4.6 are tolerance tests when noise and foreground objects are present. Section 4.7 discusses the confidence measure of the system and relates it to the performance.

4.1. Frame-to-frame alignment accuracy and capture range

The basic performance test of a stabilization system or algorithm is to find out the magnitude of error in the pure frame-to-frame motion. It is also interesting to know the range in which the system is able to handle the offsets. We have used the test image sequence generated from the Bridgewater image. The reference is the first frame of the sequence. The alignment will break when there is less and less overlap between the current image and the reference image. This metric is to measure the performance of the system or algorithm that handles frame-to-frame motions. The motion is a 2-dim translation:

$$d_x(a,t) = at, d_y(b,t) = bt \quad (4)$$

where $d_x(a,t)$ and $d_y(b,t)$ are translational offsets at time t with respect to the reference image, and a, b are integers.

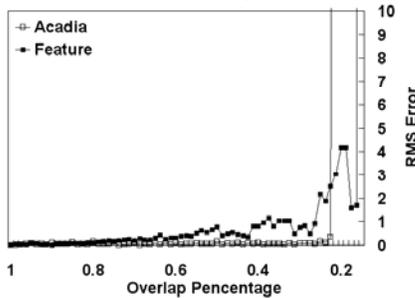


Fig. 3. The relation of RMS error and the overlap percentage in the Acadia PCI board and feature based method. ($a = b = 6$ from Eq. 4)

The minimum overlap percentage in Acadia PCI board is around 23%. For a system implementing the direct method,

this result is surprising. The reason is that Acadia program has taken into account that the next initial guess of the alignment uses the result from the current alignment. For a system implemented with feature-based method, overlap percentage is around 16% which is better than that of an Acadia board. But the average RMS error is around 1/2 of a pixel which is higher compared to the performance of an Acadia board.

4.2. Drift measure

The stabilization system needs to handle very large motion between the current frame and an anchor frame. To accomplish this, intermediate reference frames are used and motions are cascaded. The reference image sequence is designed that it uses the images from the input sequence that updates every 20 pixels. This measurement is to test the drift of cascaded motions and find out average frame-to-frame motion in the system. We have used the Bridgewater image sequence generated from Fig. 3. The instability offset is designed as a function of a, f , and t :

$$d_x(a, f, t) = a(t) \sin(2\pi ft) \quad (5)$$

where f is the frequency, a is the amplitude, and t is the frame number. $a(t)$ is defined as a linear relation to t , as illustrated in Fig. 4(a).

The offset starts from 0, and ends at 0. The drift error = $d_x(a, f, T-1) - d_x(a, f, 0)$, which shall reflect the drift between the first frame and the last frame. The average drift error in this measure is averaged by the whole number of frames in the sequence. This drift error differs from the frame-to-frame alignment RMS error in a sense that frame-to-frame error can be either way, while the drift error is one-directional.

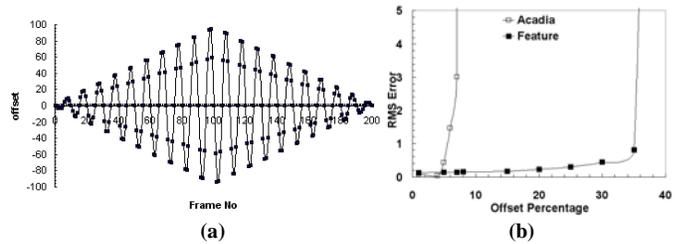


Fig. 4. (a) The instability offset is a linear relation coupled by a sine wave. $a(t) = t$, and $f = 0.1$. The maximum offset between two consecutive frames in this sequence is 8% of the image. (b) The relation of RMS error vs the offset percentage for complex motion tests in the Acadia PCI board and feature based algorithm.

For Acadia stabilization system, the drift error is 1 pixel after 100 frames, or 0.01 pixel per frame at $f = 0.1$. The average RMS error is around 1/7 of a pixel when initial motion provides no help or even does the opposite. Actually the frame-to-frame performance is better in the case when initial motion is set to identity. This suggests the system may have different settings when work for aerial surveillance or border securities. The RMS error is much larger than the drift error per frame. However, the frame-to-

frame alignment error may cancel out along the time, but the drift error accumulates. The drift error can become an important index to the system performance. A suggestion to reduce the drift error is to increase the threshold that updates the reference image so that cascading frequency reduces.

For feature-based method, the drift is about 2 pixels after 100 frames, most of which come from the accumulated errors by motion cascading. Since feature-based method is a global method, the update rate can be decreased to reduce the error. Other drifts may be caused from the fact that feature matching is on subsampled images. The average RMS error is the same as that has been discussed in Sec 4.1.

4.3. Measure of the complexity motion

The instability sequence generated by the affine model has been used to test the ability of the stabilization system to handle the complex motions. All 6 parameters defined in Eq. 2 are randomly offset by some percentage from its identical motion. The input sequence starts with an identity motion from the first frame which becomes the reference image. Then each new frame is warped from the random affine based on the Bridgewater image.

At an offset percentage of 7, the two scaling terms fall into the range of [0.93,1.07], and two shear terms in [-0.07,0.07]. The translation in x and y direction is in the range of [-35, 35]. The tests show that the average RMS error in good alignments is around 1/15 of a pixel for Acadia and 1/7 of a pixel for feature-based method. The average RMS error increases when the percentage of the offset from 6 parameters in the affine motion becomes larger. The tests are performed when previous motion in Acadia is used as the initial estimate for the current alignment, which degrades the performance. Fig. 4(b) also clearly indicates that Acadia stabilization system can handle offsets only up to 7% from the identical motion while the feature-based method can handle up to 35 % in our test settings.

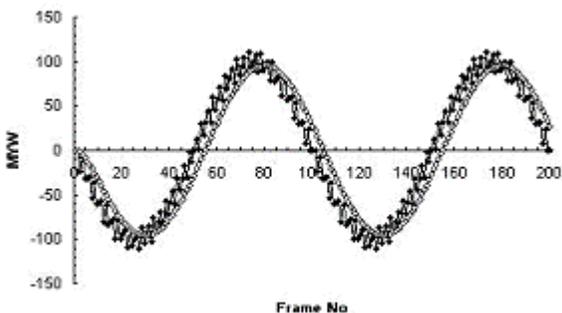


Fig. 5. Smoothed motion from Acadia Stabilization system. High frequency jitter is added to the input sequence as illustrated by the solid dots. The hollow dots represent the smoothed output motions that remove the jitters.

4.4. Smoothing measure in the presence of jitter

Many stabilization systems have the ability of smoothing the unwanted motion. We generate a test sequence that adds

the high frequency jitter on top of the low frequency motion. The designed offset is a function of a, b, c, f , and t :

$$d_x(t) = ct, \quad (6)$$

$$d_y(a, b, f_L, f_H, t) = a \sin(2\pi f_L t) + b \sin(2\pi f_H t),$$

where c is a constant, $a \gg b, f_H \gg f_L$.

We set the parameters as follows: $c=1, a=100, b=12, f_H=0.6$, and $f_L=0.01$. We also set the smoothing factor to be 0.9. (If set to 1.0, no smoothing is applied, images are completely aligned to the reference image.) Fig. 5 plots the input sequence in solid dots, and the smoothed output with respect to the reference image in hollow dots. There is a slight phase difference between the input and output.

4.5. Noise tolerance

Noise is an important factor in degradation of image alignment. This test uses Gaussian noise or speckle noise (salt and pepper noise) to measure the performance degradation. We have used the example from Section 4.1 as the input sequence and added Gaussian noise as a function of variance. The plot of RMS error vs the S/N ratio of the Acadia stabilization system and the feature-based method is shown in Fig. 6(a). In feature-based method, the error shoots up at SNR of around 4 while the Acadia stabilization system has excellent performance which can handle sequences with SNR upto 2.5. Another interesting observation in Acadia system is that a small amount of noise can reduce the RMS error. This also happens to the foreground-background performance test. The reason could be that the noise reduces the drift error in the system. This may not be the case for other stabilization systems.

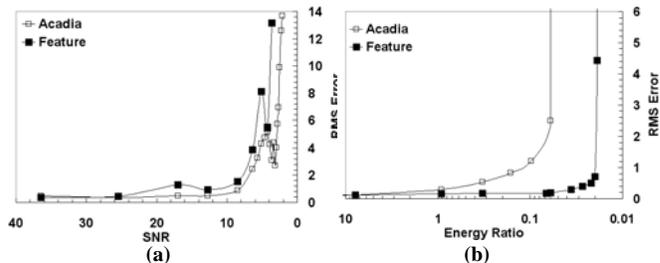


Fig. 6. (a) Different Gaussian noises are added to the input sequence from Section 4.1 and the average RMS error at different noise level is plotted. (b) The plot shows the average RMS error at different energy levels of the target.

4.6. Stabilization measure in the presence of moving foreground objects

In this experiment, we evaluate the performance of the system in presence of moving foreground objects. We generate a test sequence in which a square textured object moves in the foreground and the background has an independent jitter motion. It is observed that as the ratio of the texture content in the background to the texture content in the texture object decreases, the mean square error starts to increase. The energy of the texture is calculated using a co-occurrence matrix of the gray scale values as given in

[13]. Fig. 6(b) shows a plot between the RMS error and the energy ratio. In Acadia program, the error shoots up at an energy ratio of 0.067 whereas the feature-based method has a superior performance that can handle till an energy ratio of 0.019.

4.7. Frame-to-frame alignment confidence measure

The above measures assume a motion model and have used video sequences that bear the ground truth. In this section, we are adapting a metric that can be used for arbitrary motion, and is insensitive to geometric qualities and yet tolerant of photometric differences, such as differences in brightness and contrast due to an adjustment of camera gain between images. This was first proposed by Pope et al. [14], which aimed at evaluating and validating estimates of image alignments for real-time videos. The idea is to find the matches from orientation, not intensities. The equation can be written as:

$$CM = \frac{1}{N_V} \sum_{(x,y) \in V} |\theta_I(x,y) - \theta_R(x,y)| \quad (7)$$

$$= \frac{1}{N_V} \sum_{(x,y) \in V} \left| \tan^{-1} \left(\frac{\nabla_y I_I(x,y)}{\nabla_x I_I(x,y)} \right) - \tan^{-1} \left(\frac{\nabla_y I_R(x,y)}{\nabla_x I_R(x,y)} \right) \right|$$

where V is the set of valid pixels from both images. The valid pixels includes those pixels in the overlapped region of both images, and

$$|\nabla I_I(x,y) + \nabla I_I(x,y) + \nabla I_R(x,y) + \nabla I_R(x,y)| > T_1,$$

where T_1 is a threshold which determines the minimum gradient at which gradient direction can be reliably measured and compared.

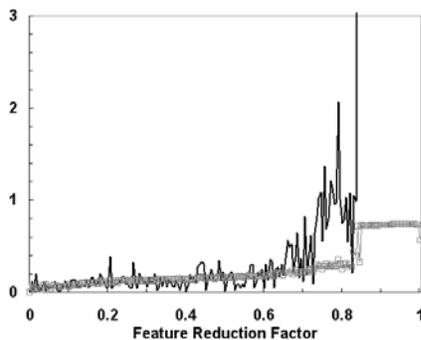


Fig. 7. Plot of Feature Reduction Factor vs Confidence measure and RMS error for Acadia stabilization system. The hollow dots represent the confidence measure of each alignment. The RMS errors are plotted as a solid line. A high value from the confidence measure indicates a bad alignment.

We have used the same input sequence as from Section 4.2. We then modify the sequence such that the amplitude with respect to the DC value of each image is scaled by $(1-t/T)$, where t is the frame number, and T is the total number of images in the sequence. Thus the features of the image at time t are reduced by a feature reduction factor of t/T . The reference image is the 1st frame of the sequence. From Fig. 7, it can be clearly observed that the confidence

measure strongly correlates to the RMS error. In Acadia PCI board, the error shoots up when the features are reduced by a factor of 0.98, whereas the feature based method has a low performance with an error shoot up at a feature reduction factor of 0.85 itself.

6. CONCLUSION

In this paper, we have laid down a procedure to evaluate the performance of video stabilization algorithms. The performance measure can also be used to improve the algorithm. For example, the confidence measure can be used in the Acadia stabilization to signal any false alignment. The motion can either be extrapolated from the previous motions, or use a different alignment method for a better motion.

The superiority of a system will change with different metrics. For example, in our experiments, Acadia system performs excellently in handling noises and vast photometric differences and the sub-pixel precision for all the experiments are very good. For feature-based method, though the sub-pixel precision is not as good as Acadia system, they perform very well in handling random affine motions and moving foreground targets. Hence, based on the type of application, the stabilization system can be evaluated for those specific metrics. We believe that these proposed methods can become a gold standard to benchmark video stabilization algorithms.

REFERENCES

- [1] Zitova, B. and Flusser J.: Image registration methods, a survey. *Image and Vision Computing*, **21**(11), 977-1000, (2003).
- [2] Acadia PCI vision acceleration board, <http://www.pyramidvision.com/products/acadia/index.asp>
- [3] Stable Eyes, <http://www.ovation.co.uk/Video-Stabilization.html>
- [4] SteadyEyes, <http://www.dynapel.com>
- [5] FlexiVision III, <http://www.barco.com/corporate/en/products/product.asp?genr=1550>
- [6] Balakirsky, S. and Chellappa, R.: Performance Characterization of Image Stabilization Algorithms. *University of Maryland Technical Report*, (1996) (CAR TR-822).
- [7] Morimoto, C. and Chellappa R.: Fast electronic digital image stabilization. *Int. Conf. Pattern Recognition*, (1996).
- [8] Morimoto, C. and Chellappa, R.: Evaluation of image stabilization algorithms. *IEEE ICASSP*, (1998)
- [9] Skerl, D., Likar, B., and Pernus F.: A Protocol for Evaluation of Similarity Measures for Rigid Registration. *IEEE Trans. Medical Imaging*, **25**(6), 779, (2006).
- [10] West, J., et al.: Comparison and evaluation of retrospective intermodality brain image registration techniques, *J. Computer Assisted Tomography*, **21**, 554-566, (1997).
- [11] van der Wal, G.S., Hansen, M.W., and Piacentino, M.R.: The Acadia Vision Processor. *Proc. IEEE Int. Workshop Computer Architectures for Machine Perception*, 31-40, (2000).
- [12] Bergen, J., P. Anandan, K. Hanna, and R. Hingorani: Hierarchical Model-Based Motion Estimation. *ECCV*, Santa Margarita Ligure, 237-252, (1992).
- [13] Sonka, M., Hlavac, V., and Boyle R.: *Image Processing, Analysis, and Machine Vision*, 2nd edn.
- [14] Pope, A., and Zhang, C., Invention disclosure SAR-15512.