실시간 영상 안정화를 위한 키프레임과 관심영역 선정

*배주한 *황영배 *최병호 **전재열
*전자부품연구원, ** 삼성테크윈
*jhbae@keti.re.kr

Adaptive Keyframe and ROI selection for Real-time Video Stabilization

*Bae, Ju-Han *Hwang, Young-Bae *Choi, Byung-Ho **Chon, Je-Youl
*Korea Electronics Technology Institute, **Samsung Techwin

Abstract

Video stabilization is an important image enhancement widely used in surveillance system in order to improve recognition performance. Most previous methods calculate inter-frame homography to estimate global motion. These methods are relatively slow and suffer from significant depth variations or multiple moving object. In this paper, we propose a fast and practical approach for video stabilization that selects the most reliable key frame as a reference frame to a current frame. We use optical flow to estimate global motion within an adaptively selected region of interest in static camera environment. Optimal global motion is found by probabilistic voting in the space of optical flow. Experiments show that our method can perform real-time video stabilization validated by stabilized images and remarkable reduction of mean color difference between stabilized frames.

1. Introduction

Video Enhancement has been widely used in importance with the increasing needs of digital media from various area. The primary purpose of video stabilization is on eliminating unwanted shaky or annoying motion from camera effectively. Even though most surveillance cameras are installed with hardware stabilizer, they are still suffering from unexpected weather condition (wind/rain). On condition above, video stabilization makes an important role in getting higher success rate of object detection and trace.

In general, video stabilization methods follow a three-step framework: motion estimation, motion compensation (called motion smoothing or motion filtering) and image composition[1]. Motion estimation calculates interframe motion between consecutive frames[2][5] which decrease high-frequency component and computes global transformation to generate a stabilized frame. Finally image composition warps the current frame according to the global transformation and generates the stabilized image sequence. Video stabilization involves motion compensation and it produces missing image area due to the unintended motion. Therefore, image composition can optionally have inpainting and deburring [3] to improve the quality of stabilized image sequence.

Most previous image stabilization methods made assumptions on motion model to estimate global motion. Planar homography model is popular due to its robustness compared to optical flow model but it still has restriction on the scene which contain multiple moving objects or great depth variation.

In this paper, we propose an efficient video stabilization method which aims to generating real-time stabilized video sequence with good visual quality in large area surveillance system which used static camera. we assume that camera angle is not changed frequently so that we can choose key frame as reference and global motion is reasonably small. It enables us to set the motion model as X and Y axis translation to attain real-time stabilization. Once key frame is obtained, we adaptively select feature extraction region and then estimate global motion from the selected region by probabilistic voting. After estimating global motion, motion compensated video sequence is generated.

2. Proposed Approach

In this paper, we propose stabilization method for real-time stabilization with visually good quality. First, we can set the key frame D_k, as reference to current frame to estimate the global motion without calculating interframe transformation or summation of pixel value. In order to select key frame with low computational cost, we define a measure P_k, average distance from each feature to the center of a frame that describes distribution of features.

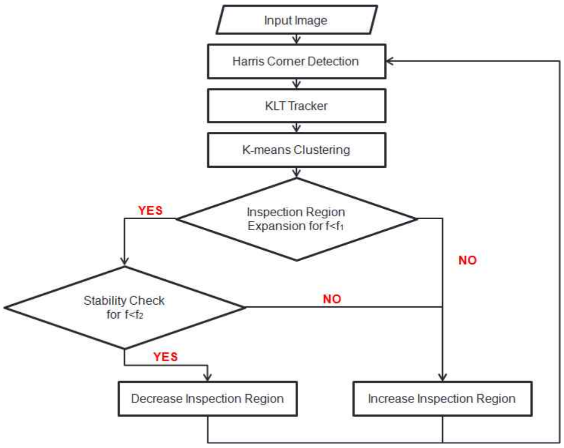

Fig.2 Flowchart of Adaptive ROI selection

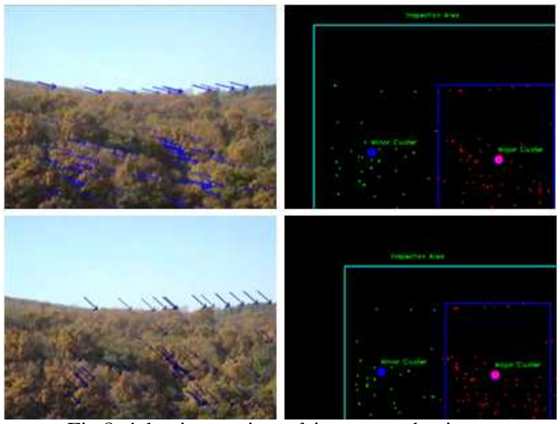

Fig.3 Adaptive region of interest selection
(Blue: pre-inspection region, White: Inspection region)

If the average distance P_k, from N frame is calculated, the frame with the smallest SAD(summation of absolute difference) would be chosen as key frame compared to the neighboring frame. By denoting average distance of a frame by P_k, key frame, D_k is defined by

$$D_k = \arg_k \min\left(\sum_{n=1}^{N} |P_k - P_n|\right) \qquad (1)$$

$$k = \{x|1 < x < N\} \;\; N = Nmber\,of\,frames$$

Fig.1 show the result of key frame selection based on (1).

After selecting key frame, the next step is to adaptively select feature extraction region based on feature distribution. Computational cost can be reduced by minimizing feature extraction region. In Fig.2 features are extracted by using Harris corner detector[4] and classified by k-means clustering. K-means clustering algorithm which shows fast and good classification performance can be fitted on real-time performance. Features are clustered as major and minor cluster and then pre-inspection region starts to expand from the cluster center until it covers whole major cluster features. Fig.3 shows that inspection region (white rectangle) is made enough to cover the trace of pre-inspection region (blue rectangle) after its expansion for user defined frame time f_1. After expansion of inspection region, we test stability of the inspection region. We calculate the ratio between the number of major features from previous and current frame for user defined frame time f_2. If stability check is fine, inspection region starts to shrink until it reaches to certain range from pre-inspection region. If stability check fails, inspection region become frame size. Fig.3 shows that feature extraction region is changed according to the algorithm above.

Next step is to estimate global motion based on the features on the inspection region. We used KLT feature tracker to find feature motion[6][7]. Fig.4 shows histogram of horizontal and vertical disparity of a frame. Global motion is selected in probabilistic voting from each histogram then motion compensated video sequence is generated based on global motion.

3. Experiment Result

To evaluate the performance of the proposed method, we have captured 4 video clips (Sony MHS-FS1, 640 x 480 resolutions, 3 minutes each) to consider different types of scene. We measure mean color difference of initial 400 frames.

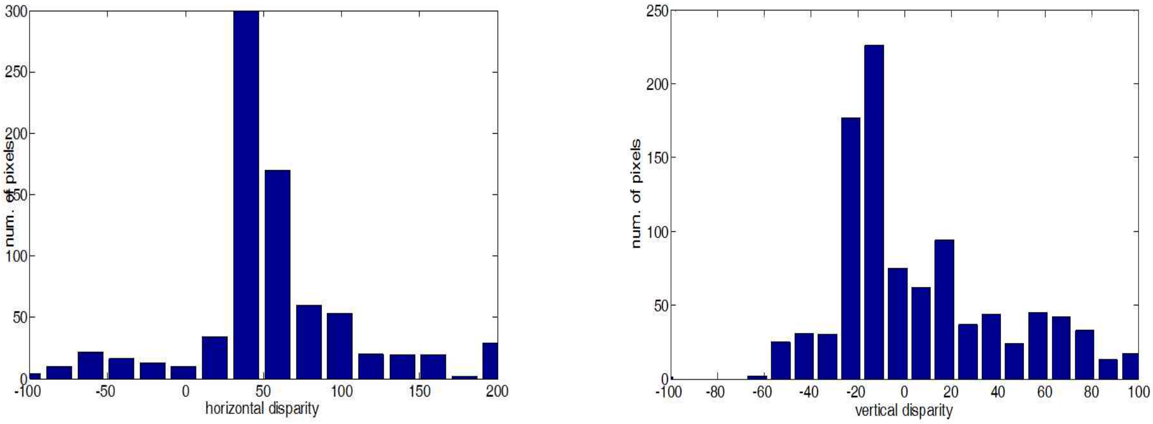

Fig.4 Histogram of horizontal and vertical disparity

Fig.5 Stabilization results Top row : original frames Bottom row : stabilized frames

Fig.5 shows that stabilization result using of proposed method compared with original frames and stabilized frames. Most previous methods used affine motion model to estimate global motion that can consider scale and rotation variation. However, these methods are relatively slow due to pixel interpolation. In order to attain real-time performance, we assume simple translation motion model that approximate the conventional affine motion mode based on these observations. First, rolling motion of camera is rarely occurred. Second, scenes are far from the camera. Even though we used simplest motion model, resulting videos are visually well aligned compared to key frame. Quantitative evaluation of the quality of resulting video is decribed in Fig.6 and Table 1. We measures mean color difference between before and after stabilization. By denoting kth pixel from frame I,S i_k, s_k ,mean color difference M defined by

$$M = \sum_{k=0}^{N} \sqrt{\frac{1}{3}(Ri_k - Rs_k)^2 + (Gi_k - Gs_k)^2 + (Bi_k - Bs_k)^2}$$

$$k = \{x|1 < x < N\},\ N = total\,element\,of\,a\,frame \quad (2)$$

Even though global motion compensation is almost similar to key frame, most mean color difference in stabilized frame is still located around 25 due to the model approximation and noise. However, resulting measures in stabilized frame are generally 30% lower compared to unstabilized case. In the case of consecutive stabilized frames, it shows smaller mean color difference compared to previous case due to small global motion difference between frames. The proposed method shows quantitatively good performance compared to unstabilized case.

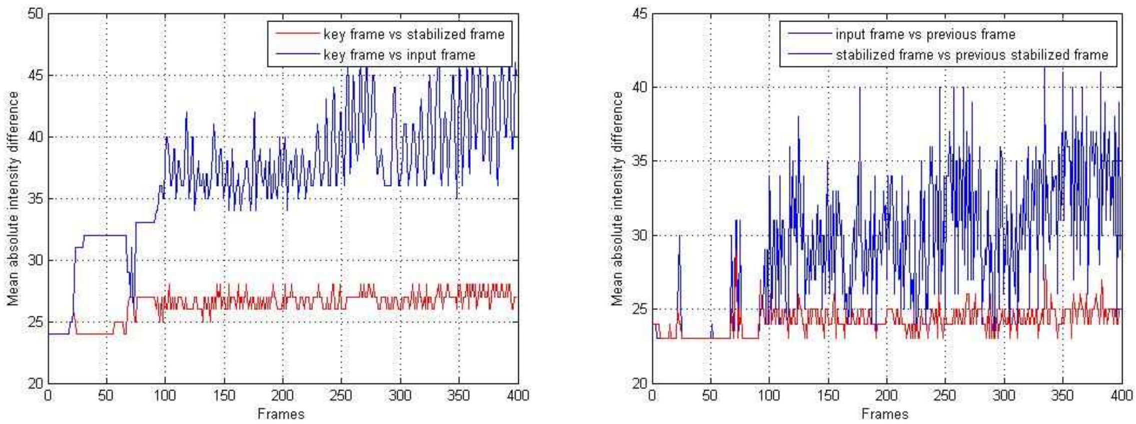

Fig.6 Mean color difference against key frame and for consecutive frames

TABLE 1 : Comparison with the unstabilized case

Case	Mean color difference
key frame vs current frame	36.89
key frame vs stabilized frame	26.34
current frame vs previous frame	29.06
current stabilized frame vs previous stabilized frame	24.17

4. Conclusion

We propose video stabilization method based on key frame selection and adaptive region selection of feature extraction. Frame that has smallest summation of absolute difference from the center is selected as key frame and feature extraction region is adaptively assigned according to the distribution of features. Global motion is estimated based the features from selected region then motion compensated video sequence is generated. We can achieve real-time performance by using proposed method.

5. Reference

[1]C. Morimoto and R. Chellappa. Evaluation of image stabilization algorithms. In Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing 1998, pages V.2789 - 2792, 1998.

[2] H.-Y. Shum and R. Szeliski, "Construction of Panoramic Mosaics with Global and Local Alignment, "Int'l J. Computer Vision and Pattern Recognition, vol.2, no.3, pp.101-130,2000

[3] Matsushita, Y.; Ofek, E.; Weina Ge; Xiaoou Tang; Heung-Yeung Shum; , "Full-frame video stabilization with motion inpainting," Pattern Analysis and Machine Intelligence, IEEE Transactions on , vol.28, no.7, pp.1150-1163, July 2006

[4] C. Harris and M.J. Stephens. A combined corner and edge detector. In Alvey Vision Conference, pages 147 - 152, 1988.

[5]J. Bergen, P. Anandan, K. Hanna, and R. Hingorani, “"Hierarchical Model-Based Motion Estimation,”" Proc. Second European Conf. Computer Vision, pp. 237-252, 1992.

[6] Bruce D. Lucas and Takeo Kanade. An Iterative Image Registration Technique with an Application to Stereo Vision. International Joint Conference on Artificial Intelligence, pages 674 - 679, 1981.

[7] Jianbo Shi; Tomasi, C.; , "Good features to track," Computer Vision and Pattern Recognition, 1994. Proceedings CVPR '94., 1994 IEEE Computer Society Conference on , vol., no., pp.593-600, 21-23 Jun 1994