Feature-based Video Stabilization for Vehicular Applications

Keng-Yen Huang, Yi-Min Tsai, Chih-Chung Tsai, and Liang-Gee Chen
DSP/IC Design Lab, Graduate Institute of Electronics Engineering
National Taiwan University, Taipei, Taiwan
{kyhuang,ymtsai,cctsai,lgchen}@video.ee.ntu.edu.tw

***Abstract*—This paper describes a method to stabilize video for vehicular applications based on Harris features and adaptive resolution. Lucas-Kanade method is applied to match feature points of consecutive frames and construct the feature motion flow. A damping filer is utilized to model the unwanted motion and global motion is separated by extracting oscillation. 92% correct rate with 0.54 second per frame is achieved. The provided benchmark shows outperformance of the proposed method.**

I. Introduction

As vehicular safety systems become more and more popular these years, many applications such as advanced driver assistance system (ADAS) are proposed to protect drivers from car accidents. Vehicular video-based processing (VVP) is to provide a more intelligent assistance and reduce the cost of the whole system. However, severe oscillation may cause blur problems and result in low accuracy of VVP . To eliminate the oscillation effects, video stabilization is adopted as a preprocessing stage.

Video stabilization mainly relies on global motion estimation (GME). GME attempts to estimate the global motion and separate the motion into intentional motion (IM) and unwanted motion (UM). By subtracting the UM, a stabilized video can be obtained from a shaky condition.

Conventionally, video stabilization is used to solve jitter in hand-held devices. In [1], camera motions are transformed into frequency domain. UM belongs to higher frequency which can be removed through frequency decomposition. Extracting the edge information [2] and projecting the frame information from 2D to 1D [3] are proposed in literatures for high-speed processing. The motion vector field is utilized to estimate global motion through Newton-Raphson's method [4]. However, these algorithms are degraded in accuracy and are time-consuming while applying in vehicular applications.

For this purpose, lane line monitoring [5] utilizes lane lines and the vanishing point to stabilize video. However, it is vulnerable to the clearance of lane lines. In this paper, a technique based on feature flow analysis is proposed to estimate global motion under variant driving conditions.

II. Proposed Processing

The system block diagram is shown in Fig. 1. Firstly, current resolution is downsampled to desired one by resolution adaptation. Secondly, Harris feature points are extracted from the downsampled frame after RGB-to-Gray color conversion. Feature points are matched and feature flow is constructed by Lucas-Kanade method. Lastly, through a damping filter, UM is subtracted from calculated global motion and the stabilized result is produced.

A. Resolution adaption

Frames are processed under two resolutions, original full resolution (R_1) and subsampled lower resolution (R_2). Every frame is first processed under R_2 and the primary estimation is produced. The primary estimation is then compared with the prediction generated form motion prediction module. If both results are similar, the primary estimation is passed to motion prediction stage as the final

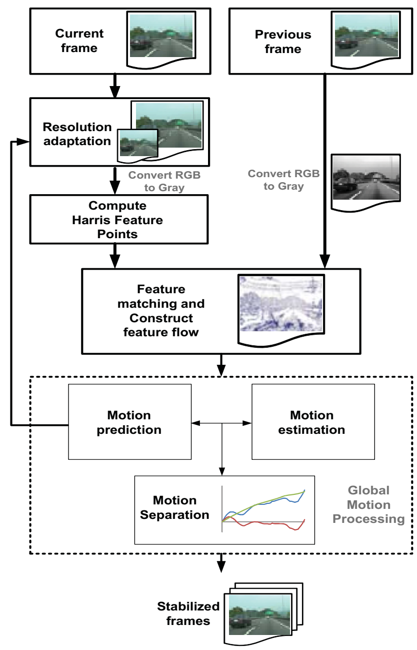

Fig. 1. Proposed video stabilization processing flow

estimation, otherwise, it reprocesses the frame under full resolution R_1.

B. Feature extraction and matching

Points are extracted as feature points by applying Harris matrix A shown in Eq.1.

$$A = \begin{bmatrix} I_x^2 & I_xI_y \\ I_xI_y & I_y^2 \end{bmatrix} \quad (1)$$

where I_x and I_y are denoted as partial derivatives. If both eigenvalues λ_1 and λ_2 of A are large, this point is claimed as a feature. M_c in Eq. 2 is used to determine the values of eigenvalues.

$$\begin{aligned} M_c &= \lambda_1\lambda_2 - \kappa(\lambda_1+\lambda_2)^2 \\ &= det(A) - \kappa trace^2(A) \end{aligned} \quad (2)$$

where κ is a tunable parameter ranging from 0.04 to 0.15. If M_c is large, it implies λ_1 and λ_2 are large and features are decided accordingly.

Lucas-Kanade method [6] is applied to match each point to its corresponding location in previous frame after feature extraction. The feature flow is constructed when all feature points are matched.

C. Global motion processing

The GME is calculated by averaging the sampled feature motion vectors (FMV) in constructed feature flow field.

$$GME_x = \frac{1}{N}\sum_{i=1}^{N} FMVx_i \ , \ GME_y = \frac{1}{N}\sum_{i=1}^{N} FMVy_i \quad (3)$$

978-1-4244-6673-3/10/$26.00 ©2010 IEEE

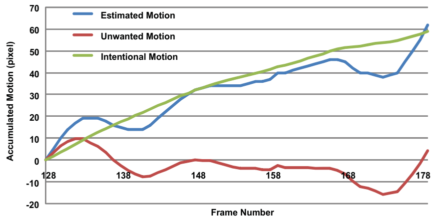

Fig. 2. Motion separation

	M1	M2	M3	M4	Proposed
Highway(%)	97.4	89.74	94.85	95.4	96.3
Tunnel(%)	53.71	46.3	52.26	N/A	92.82
City street(%)	92.65	40.32	91.26	N/A	94.2

TABLE I
CORRECT RATE OF ALGORITHMS IN DIFFERENT TESTING CONDITIONS

where N is the total number of FMV in the feature flow.

The calculated global motion is composed of IM and UM. A damping filter is proposed to model the UM as a damping spring described in Eq. 4.

$$UM(f) = O_A exp(\frac{2\xi\pi f}{T})\cos(\frac{2\pi f}{T} + \psi) \quad (4)$$

where UM(f) is denoted as the function of unwanted oscillation of a certain frame f, O_A is the maximum oscillating amplitude appeared in image plane of an oscillation, ξ is the damping coefficient, ψ is the initial condition of each oscillation, and T is denoted as the damping period in terms of frame number. The proposed algorithm generates these parameters automatically after experiencing an initial oscillation. Through modeling the UM, global motion (GM) is separated and refined to only intentional fraction, which suggests a stabilized video.

III. EXPERIMENTAL RESULTS

The program is run on a Pentium IV 2.8GHz PC. The original resolution of tested sequences is 1280×960 of different environments, including highway, tunnel, and complicated city streets. There are more than 500 tested frames in each case. Processing resolution R_1 is 1280×960, while R_2 is 320×240. Correct estimation is defined as at most 1 pixel difference between estimation and ground truth.

An example of motion separation is shown in Fig. 2. Estimated motion (blue line) is divided into IM (green line) and UM (red line). Accuracy under different resolutions is shown in Fig. 3. With resolution adaptation, computation time is reduced more than 60% while only 1.4% drop in accuracy compared to full-frame processing.

Four state-of-the-arts are compared and listed as follows.

M1:Pixel-based diamond search
M2:Curve warping [3]
M3:Motion vector field with Newton-Raphson's method [4]
M4:Lane line and vanishing point stabilization [5]

In Table. I, all algorithms claim about 90% correct rate in highway. However, some shortages appeared while applied in other conditions. In complicated environments, curve warping (M2) is too sensitive to changes of local objects. Large homogeneous area like tunnel degrades the performance of pixel-based diamond search (M1) and motion vector field (M3).

The proposed method attains 0.54 second to process a frame in Table. II. Although curve warping (M2) is faster by processing frame in 1D, only 50% in accuracy is achieved. Example of the proposed video stabilization is shown in Fig. 4.

IV. CONCLUSION

We present an efficient video stabilization method targeted at vehicle applications. Resolution adaptation greatly accelerates the stabilization procedure. A damping filter is utilized to remove unwanted oscillation calculated from feature motions. Results show the identical accuracy under different conditions and report decent performance in computation time.

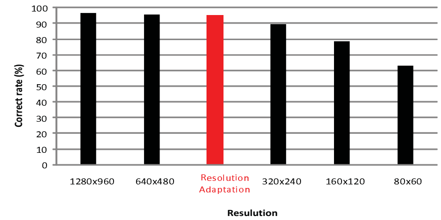

Fig. 3. Correct rate under different resolutions

	M1	M2	M3	M4	Proposed
Time(s/frame)	8.27	0.0322	4.69	0.8	0.5393

TABLE II
COMPUTATION TIME COMPARISON

REFERENCES

[1] Etay Mar and Dmitry Pundik, "Hand motion and image stabilization in hand-held devices," *IEEE Transactions on Consumer Electronics*, vol. 53, pp. 1508-1512, November 2007.

[2] Ohyun Kwon, Jeongho Shin, and Joonki Park, "Edge based adaptive kalman filtering for real-time video stabilization," *IEEE International Conference on Consumer Electronics (ICCE)*, pp. 75-76, January 2006.

[3] Angelo Bosco, Arcangelo Bruna, Sebastiano Battiato, Giuseppe Bella, and Giovanni Puglisi, "Digital video vtabilization through curve warping techniques," *IEEE Transactions on Consumer Electronics*, vol. 54, pp. 220-224, May 2008.

[4] Yeping Su, Ming-Ting Sun, and Vincent Hsu, "Global motion estimation from coarsely sampled motion vector field and the applications," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 53, pp. 1636-1647, November 2004.

[5] Yu-Ming Liang, Hsiao-Rong Tyan, Shyang-Lih Chang, Hong-Yuang Mark Liao, and Sei-Wang Chen, "Video stabilization for a camcorder mounted on a moving vehicle," *IEEE Transactions on Vehicular Technology*, vol. 53, pp. 1636-1647, November 2004.

[6] Alex Rav-Acha and Shmuel Peleg, "Lucas-Kanade without iterative warping," *IEEE Proc. of International Conference on Image Processing (ICIP)*, pp. 1097-1100, 2006.

Fig. 4. Example for video stabilization. Left parts are original four consecutive frames during oscillation. Right parts are stabilized results.