

ELECTRONIC IMAGE STABILIZATION ALGORITHM BASED ON ADAPTIVE MOTION FILTER

¹ZHU JUANJUAN, ²GUO BAOLONG

¹Assoc. Prof., Department of Electromechanical Engineering, xidian Univ., China

²Prof., Department of Electromechanical Engineering, xidian Univ., China

E-mail: zhujoo@126.com

ABSTRACT

An electronic image stabilization algorithm is proposed based on improved Sage-Husa adaptive visual smoothing. It uses visual attention features to smooth inter-frame global motions, thus resulting in stabilized sequences. Firstly, corners are detected from reference frame as feature points. Through matching features in current frame, deleting mismatched points and solving motion equations, global motion vectors are generated. Then Sage-Husa adaptive filtering method is described in detail, which corrects the process and observation noise by estimating their statistical property in real time. So, it can improve the filter accuracy greatly and remove camera jitter but retain camera scan simultaneously. Finally each current frame is warped by fast motion compensation in consideration of the linear storage of the image. The experimental results illustrate that the proposed algorithm can stabilize the inter-frame jitter and track the real scene with smooth vision.

Keywords: *Electronic Image Stabilization, Adaptive Motion Filter, Motion Estimation, Fast Motion Compensation*

1. INTRODUCTION

Due to the duration of vision, the camera on moving vehicles will capture dithering or fuzzy images. Electronic Image Stabilization (EIS) technique aims to detect and compensate the inter-frame motion by image processing methods. Up to now, it has been widely used in moving object detection and tracking, walking robot, video compression^[1] and image mosaic.

The research mainly focuses on motion estimation and motion compensation. The former is to estimate inter-frame global motion vectors with high accuracy. The feature matching^[2] is the most popular method, which can deal with translation, rotation and zooming. Motion compensation accomplishes motion correction by analyzing the motion parameters to reduce dithering and contain scan. The already proposed algorithms include motion attenuation^[3], average filter^[4] and Kalman filter^[5-6]. The attenuation coefficient is set by experimental results, which is not applicable to all videos. The average filter has simple computation but the extra low frequency noise still exists. The Kalman filter is based on the assumption that the noise is given and obeys the Gauss distribution with zero mean value, which is impossible in real applications.

According to human visual attention mechanism, eyes can observe the stabilized videos if camera moves uniformly at low frequency speed. When the camera is mounted on moving vehicles such as planes, cars and ships, the irregular motions can be assumed as high-frequency random vibration. Based on the above analysis, the feature points are matched to get the global motion and the adaptive filter is proposed to compensate images. Firstly, we extract feature points with the improved Harris operator in regions. The points are validated using the statistical property of distance and the global motion is obtained by computing the motion model. In terms of motion compensation, the Adaptive Filter based on Sage-Husa filter is proposed to estimate process and observation noise of Kalman filter in real time. Finally, we use fast linear compensation method to ensure real-time performance. Experimental results show that the Adaptive Filter can get smoother video sequences comparing to Kalman filter, and it can track the camera scan with improved visual quality.

2. MOTION MODEL

It's supposed that the scene is a flat plane, and the optical axis of camera is perpendicular to this plane. At this time, the same zooming rate and translation of 2D images are caused by the camera

movements; and image rotation is caused by camera tilt across vertical axis. Thus, we establish affine transformation equation as follows:

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} \Delta X \\ \Delta Y \end{bmatrix} \quad (1)$$

Here, (x_{t-1}, y_{t-1}) and (x_t, y_t) are coordinates of reference frame $t-1$ and current frame t , θ is rotation angle, $(\Delta X, \Delta Y)$ is translation in horizontal and vertical direction. When the rotation angle is small (less than 5°), $\cos \theta \approx 1$, $\sin \theta \approx \theta$, and thus we linearize the original affine transformation equation (1) to get the similarity transformation equation(2).

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} 1 & -\theta \\ \theta & 1 \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} \Delta X \\ \Delta Y \end{bmatrix} \quad (2)$$

3. GLOBAL MOTION ESTIMATION

3.1 Feature Point Detection

Feature point detection is the first step in the global motion estimation. It is essential to extract feature points in reference frames quickly and accurately. Harris operator has excellent performance in consistency and effectiveness. Meanwhile, Harris operator is robust to noise and light variation [7]. So in this paper, the Harris corners are detected as feature points. The principle of Harris corner detection is as follows:

$$M = w \otimes \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \quad (3)$$

$$R = \lambda_1 \lambda_2 - k(\lambda_1 + \lambda_2)^2 \quad (4)$$

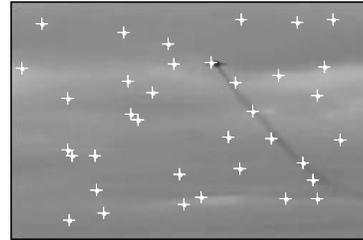
Here, w is the discrete 2-dimensional zero mean Gaussian function; I_x and I_y are image gradients in horizontal and vertical direction; $[\lambda_1, \lambda_2]$ is the eigenvalue of the auto-correlation matrix M . R is the interest value and the point with large R is taken as the feature point.

The feature points that detected by Harris operator are too dense in edge or corner region, as shown in Fig. 1(a). The points are mostly detected in the moving object, and these local motions will interfere with global motion. Furthermore, the number of points is uncertain and it depends on the threshold. To solve the above problems, the improved sub-area Harris operator is proposed. The reference image is divided into $r \times s$ non-overlapped regions, and in each region the point

with the maximal R is selected as the feature point. Thus the points distribute uniformly and these independent points can stand for the global motion.



(A) Result Of Harris Operator



(B) Result Of Sub-Area Harris Operator

Fig.1 Feature Points Detection Results

3.2 Feature Points Matching And Verification

After extracting feature points in reference frame, we use the approach of template matching to get the feature points in current frame. The feature point in reference is set as the center of each template. Then, the sequential similarity detection algorithm (SSDA) is used to find the best matched block and its center is the corresponding point in current frame. Due to local motion or texture repetition, there are unreliable matching points. So we have to verify the matching points and reserve the right ones. In this paper, we use distance criterion [8] to verify the feature points.

The Euclidean distance $d_i = (\Delta x_i)^2 + (\Delta y_i)^2$ is defined, in which Δx_i and Δy_i are coordinates' difference in horizontal and vertical direction between points' pair i in adjacent frames respectively. According to the statistic of experimental data, d obeys normal distribution approximately:

$$P(d) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(d-\mu)^2}{2\sigma^2}\right) \quad (5)$$

Here, μ and σ are the mean and standard deviation of d . According to the "3 σ criterion" of normal distribution, the data in interval of $[\mu - 3\sigma, \mu + 3\sigma]$ takes up 99.7% of total data, as

shown in Fig.2. When $|d_i - \mu| > 3\sigma$, the point is considered to be the mismatch and rejected.

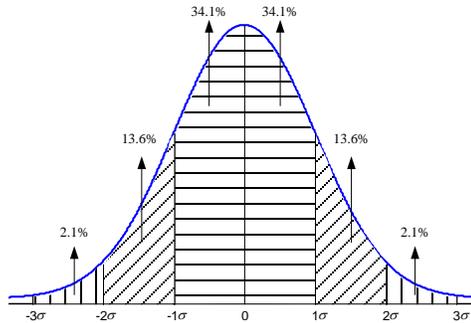


Fig.2 3σ Criterion

3.3 Global Motion Computation

All the verified feature points are then brought into equation (2) to get 2N-linear equations containing three unknown parameters. The final function $B = Am$ is in form of matrix, as shown in Equation (6). The motion parameter is defined as $m = [\theta, \Delta X, \Delta Y]^T$.

$$B = \begin{bmatrix} x_{t,1} - x_{t-1,1} \\ y_{t,1} - y_{t-1,1} \\ \vdots \\ x_{t,N} - x_{t-1,N} \\ y_{t,N} - y_{t-1,N} \end{bmatrix}, \quad A = \begin{bmatrix} -y_{t-1,1} & 1 & 0 \\ x_{t-1,1} & 0 & 1 \\ \vdots & & \\ -y_{t-1,N} & 1 & 0 \\ x_{t-1,N} & 0 & 1 \end{bmatrix} \quad (6)$$

The least-square solution of the over determined linear equation is defined as $m = (A^T A)^{-1} A^T B$.

4. MOTION COMPENSATION

Camera motion is composed of intentional scan and undesired jitter. The basic hypothesis of Sage-Husa filter is that the scan motion is intended to move smoothly at low speed towards one direction, while the jitter is rather random in altitude and direction. So, the intentional scan is of low frequency while undesired jitter is of higher frequency. We can get smooth motion vector through filter (MV_{filter}), and then the jitter is the difference between original motion vector and smooth vector, $MV_{jitter} = MV_{filter} - MV_{raw}$.

4.1 Sage-Husa Filter

In applications, the distribution of process and observation noise is unknown, so the accuracy of Kalman filter reduces and even diffuses. In the process of Sage-Husa filter approach, observation

data estimate the predictive value automatically and correct the process noise and observation noise at the same time, which can reduce the model error and improve filtering accuracy. Considering that the rotation has little effect on scanning movement, we construct linear discrete system model as follows:

$$S(k) = F \cdot S(k-1) + w \quad (7)$$

$$Z(k) = H \cdot S(k) + v \quad (8)$$

$$F = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

where F is state transition matrix, and H is observation matrix.

In this paper, the state vector $S(k) = [x(k), y(k), dx(k), dy(k)]^T$ is horizontal and vertical displacement and their instantaneous velocity; $Z(k) = [x(k), y(k)]^T$ is observation vector. $w \sim N(0, Q)$ and $v \sim N(0, R)$ represents the process noise and observation noise respectively; Q and R is variance matrix of process noise and observation noise.

The process of Sage-Husa state prediction and update is as Equation (9)^[9]:

$$S(k | k-1) = F \cdot S(k-1)$$

$$P(k | k-1) = F \cdot P(k-1)F^T + \hat{Q}(k-1)$$

$$S(k | k) = S(k | k-1) + K(k) \cdot \varepsilon(k)$$

$$\varepsilon(k) = Z(k) - H \cdot S(k | k-1)$$

$$K(k) = P(k | k-1) \cdot H^T [H \cdot P(k | k-1) \cdot H^T + \hat{R}(k)]^{-1}$$

$$P(k | k) = [I - K(k) \cdot H] \cdot P(k | k-1) \quad (9)$$

Where, $P(k | k)$ is filter error matrix and $P(k | k-1)$ is prediction error matrix; K is Kalman gain, $\varepsilon(k)$ is innovation sequence.

$$\hat{R}(k) = (1 - d_k) \hat{R}(k-1) + d_k [\varepsilon(k) \cdot \varepsilon(k)^T - H \cdot P(k) \cdot H^T] \quad (10)$$

$$\begin{aligned} \hat{Q}(k) &= [1-d(k)]\hat{Q}(k-1) \\ &+ d(k)[K(k) \cdot \varepsilon(k) \cdot \varepsilon(k)^T K(k)^T] \quad (11) \\ &+ d(k)[P(k|k) - F \cdot P(k-1|k-1) \cdot F^T] \end{aligned}$$

Here $d(k) = (1-b)/(1-b^k)$; b is forgetting factor and $0 < b < 1$.

In the above process, we can not estimate Q and R when they are unknown^[10]. And the stability and convergence are likely to decline because of losing positive definiteness. So, it is necessary to prevent the diffusion of filter.

4.2 Improved Adaptive Filter

The adaptive filter (AF) is proposed to correct the prediction error matrix $P(k|k-1)$ to avoid filter divergence. We determine whether the filter is divergent using the property of innovation sequence $\varepsilon(k)$.

$$\varepsilon(k)^T \cdot \varepsilon(k) \leq \gamma \cdot \text{Trace}[H \cdot P(k|k-1) \cdot H^T + \hat{R}(k)] \quad (12)$$

Here, γ is adjustable coefficient and $\gamma > 1$. When formula (12) holds, the filter is in normal working state to give the optimal estimate value of current state. Otherwise, the actual error is γ times more than theoretical estimate value, and the filter will diverge. At this time, $P(k|k-1)$ should be corrected by weighted coefficient $C(k)$ ^[11] as follows.

$$P(k|k-1) = C(k) \cdot F \cdot P(k-1) \cdot F^T + \hat{Q}(k) \quad (13)$$

$$C(k) = \frac{\varepsilon(k)^T \cdot \varepsilon(k) - \text{Trace}[H \cdot Q(k) \cdot H^T + R(k)]}{\text{Trace}[H \cdot F \cdot P(k) \cdot F^T \cdot H^T]} \quad (14)$$

Finally we can get smooth motion through updating $P(k|k-1)$.

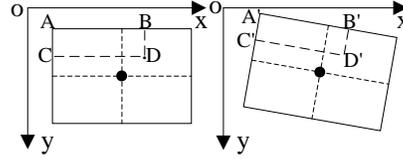
4.3 Fast Motion Compensation

After motion smooth, we can get jitter component as the difference between original motion vector and smooth vector. The horizontal and vertical jitter is computed as equation (15) shows.

$$\begin{aligned} \Delta x_{jitter} &= \Delta x_{filter} - \Delta x_{raw} \\ \Delta y_{jitter} &= \Delta y_{filter} - \Delta y_{raw} \end{aligned} \quad (15)$$

So, the compensating component as $m' = [\theta, \Delta x_{jitter}, \Delta y_{jitter}]^T$ is taken into similarity

transformation model (2) to compute new coordinates of each pixel at current image. Based on the linear storage structure of image, the relative position between pixels in an image does not change after rotation.



(A) Before Rotation (B) After Rotation

Fig.3 Rotation-Invariant Structure Of Image

As it is shown in Fig.3(a), $A(x_A, y_A)$, B , C and D are four vertices of the rectangle. Their relative position does not change after the linear transformation. The new coordinates of D' in Fig.3(b) can be calculated as:

$$\begin{aligned} x_{D'} &= x_B' + (x_C' - x_A') \\ y_{D'} &= y_B' + (y_C' - y_A') \end{aligned} \quad (16)$$

The steps of fast linear compensation method are as follows.

Step1: Similarity transformation is made at each pixel of first row and first column with formula (2) to get its new coordinates;

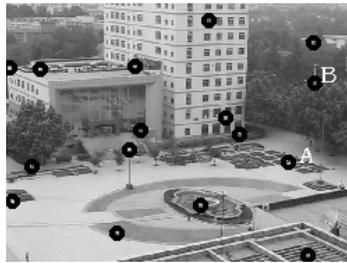
Step2: Coordinates addition is made at all the other pixels with formula (16).

The fast computation can avoid matrix multiplication at all the pixels in the whole image and save the computation time effectively.

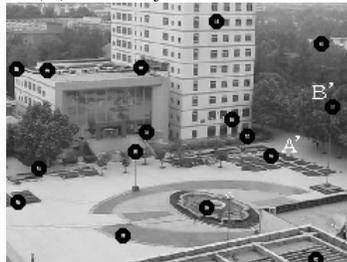
5. EXPERIMENTAL RESULTS AND ANALYSIS

We select 80 adjacent frames with the size of 640×480 from the test video on a moving camera. The camera scans slowly in horizontal direction and the platform dithers in horizontal and vertical direction.

5.1 Feature Points Matching And Verification



(A) Result Of Point Selection



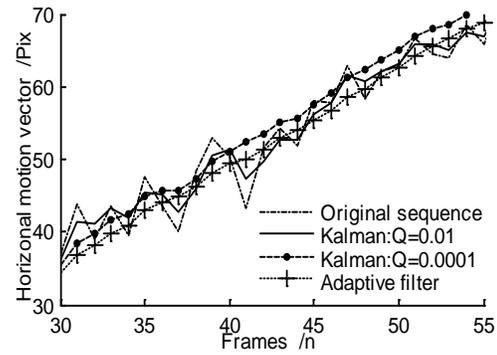
(B) Result Of Point Matching And Validation

Fig.4 Feature Points Selection And Validation

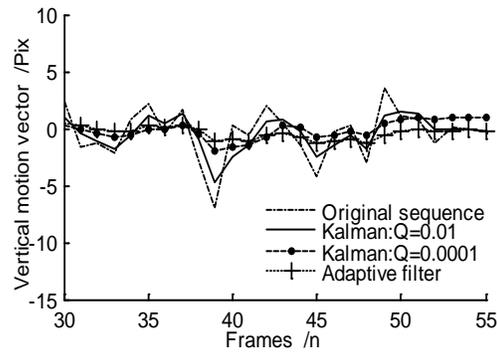
Fig.4 is the result of point selection, matching and validation. We detect 16 feature points in reference frame, and the matching results of feature points in current frame are shown in Fig.4(b). The points “A” and “B” are validated as mismatching points through the distance criterion. After deleting the mismatched points, the global motion estimation is improved.

5.2 Results Of Motion Smooth

Fig.5 is the comparison of the proposed AF (Adaptive filter) and Kalman filter. The original horizontal motion curve 5(a) accumulates at a steady increase because of the scanning. Curve 5(b) fluctuates in the zero-position due to jitter in vertical direction. As it is shown in Fig.5, different process noise Q in Kalman filter affects the smooth result obviously. When Q is large, the filter curve is approximate to the original curve without obvious smooth effect. While small Q leads to over smooth curve deviating from the real movement, and even results in filter divergence. It can be seen that the adaptive filter can smooth jitter component well and track camera scan effectively.



(A) Result Of Horizontal Motion Filter



(B) Result Of Vertical Motion Filter

Fig.5 Comparison Of Motion Filtering Result

5.3 Comparison Of Motion Compensation

The global motion of test images of Fig.(4) is $m = [0.42, 10.40, -4.25]^T$ and the dithering motion by equation (15) is $m' = [0.13, 3.52, -0.91]^T$. The standard deviation of motion curves before and after compensation is $[\sigma_x, \sigma_y] = [11.1, 6.0]$ and $[\sigma'_x, \sigma'_y] = [6.2, 4.3]$ respectively. The lower the standard deviation the smoother the filtering result is. Furthermore, the fast compensation time is 4.025ms per-frame, which reduces 76.2% comparing with 17.633mspf by tradition compensation method.

The $PSNR$ ^[12] (Peak Signal to Noise Ratio) is used to test the global fidelity of inter-frames. From Fig.6, we can see that the $PSNR$ after stabilization is increased greatly, which means the difference between frames is reduced and the video is stabilized.

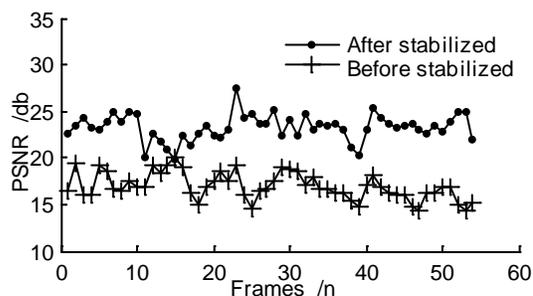


Fig.6 Comparison Of PSNR

6. CONCLUSIONS

A fast video stabilization algorithm is proposed based on adaptive motion filter, which is independent of the noise variance matrix and has fast convergence. The filter can track scan and reduce jitter simultaneously, thus improve the smoothness of videos. The real time processing is obtained with the fast linear compensation. Experimental results show that the algorithm is efficient and robust to noise. And, for more complicated real-time applications, other parametric motion model and motion filter are required to make further investigation.

ACKNOWLEDGEMENT

This work is supported by the National Science Foundation of China (61003196) and Fundamental Research Funds for the Central Universities (K5051004004).

REFERENCES:

- [1] Kaliki Sri Harsha Reddy, 2R.Saravanan. Efficient design of low area based H.264 compressor with H.264 integer transform. JATIT, 38(2):131-135, 2012.
- [2] A Aysun Yeni, Sarp Ertfürk. Fast digital image stabilization using one bit transform based sub-image motion estimation [J]. IEEE Transactions on Consumer Electronics.2005, 51(3):917-921.
- [3] Sung-Jea Ko, Sung-Hee Lee, Seung-Won Jeon et al.. Fast digital image stabilizer based on gray-coded bit-plane matching [J]. IEEE Transactions on Consumer Electronics.1999, 45(3):598-603.
- [4] ZHU Juan-juan, GUO Bao-long. Adaptive video stabilization system for camera on moving vehicle [J]. Journal of Optoelectronics-Laser. 2007, 18(1): 108-112.
- [5] Sndrew Litvin, janusz Konrad, William C Karl. Probabilistic video stabilization using Kalman filtering and mosaicking[A]. IS&T/SPIE[C]. 2003, 5022:.663-674.
- [6] ZHU Juan-juan, GUO Bao-long. Electronic image stabilization system based on global feature tracking[J]. Journal of Systems Engineering and Electronics.2008, 19(2):228-233.
- [7] Dong-O Kim, Rae-Hong Park. New image quality metric using the Harris response [J]. IEEE Signal Processing Letters.2009, 16(7):616-619.
- [8] P.Tissainayagam, D.Suter. Assessing the performance of corner detectors for point feature tracking applications [J]. Image and Vision Computing.2004, 22(8):663-679.
- [9] ZHANG Chang-yun. Approach to adaptive filtering algorithm [J].Acta Aeronautica et Astronautica Sinica.1998, 19(7):96-99.
- [10] R.Lynn Kirlin, Alireza Monghaddamjoo. Robust adaptive Kalman Filtering for systems with unknown step inputs and non-Gaussian measurement errors [J].IEEE Transactions on Acoustics.1986, 34(2):252-263.
- [11] XU Jing-shuo, QIN Yong-yuan and PENG Rong. New method for selecting adaptive Kalman filter fading factor [J].System Engineering and Electronics. 2004, 26(11):1152-1154.
- [12] J. Shanmuga Prathipa, T. R. Lakshminarayanan. PSNR based clustering and indexing for fast access. JATIT, 38(2): 170-176, 2012.