Proceedings of the 2018 IEEE
International Conference on Robotics and Biomimetics
December 12-15, 2018, Kuala Lumpur, Malaysia

Hybrid Image Stabilization of Robotic Bionic Eyes *

Xiaopeng Chen[1] Taoran Zhang ShaoWen Fu[2]
Changjing Wang Chenghao Hua Qiang Huang

***Abstract*— Traditional vision system mounted onto robots may have image blurring and shaking caused by unwanted robot motion, which makes it difficult to sense the environment precisely. This paper presents a real-time hybrid image stabilization method combining mechanical motion compensation and electronic compensation to solve the problem based on a 9-DOF vision platform called Robotic Bionic Eye developed by us. In the mechanical video stabilization stage, a distrubtance reduction control method is proposed to suppress the disturbtance. In the electronic video stabilization stage, 3D camera rotational motion is obtained from IMU and 2D image motion is detected by image motion detection algorithms. These motion is further filtered and image sequences are rewarped to remove rotational shake as well as translational shake for a second time. Experiments verified that our algorithm improves the inter-frame transformation fidelity while ensuring real-time performance.**

I. INTRODUCTION

Quality in robot vision systems is not given only by the quantitative features such as the resolution of the cameras, the frame rate or the sensor gain, but also by the qualitative features such as sequences free of unwanted movement, fast and good image pre-processing algorithms and real-time response[1]. Video image stabilization technology is to reduce or even eliminate the motion of video sequences caused by irregular motion of camera system, the main reason of this irregular movement is the attitude change of camera system at different times. For example, in the on-board camera system, the movement of the vehicle and the body vibration caused by the uneven road surface will make the image unstable, as well as on the long-range target reconnaissance and observation, because of the movement of ships and sea waves fluctuations, the image jitter may be large, resulting in poor imaging quality.

Image stabilization first hit the consumer market in 1995, with the introduction of a stabilized zoom lens by cannon. Since this introduction, image stabilization has appeared in an increasing number of products[2]. Video image stabilization technology can be divided into three categories: optical stabilization, mechanical image stabilization and digital image stabilization[3].The optical image stabilization adjusts the optical path adaptively through the active optical components, compensating the image motion caused by the shaking of the camera platform, so as to achieve the goal of image stabilization; Mechanical image stabilization is achieved by detecting the jitter of the camera platform through sensors such as gyroscopes and then adjusting the servo system; Digital image stabilization includes electronic stabilization and pure digital stabilization[4], The difference is that the electronic image stabilization technology using gyroscope or other sensor to detect the camera motion, while the pure digital stabilization technology estimates the camera shaking through the processing of the continuous video images, and finally, the video is stabilized by filtering and motion compensation of the original video frame image.

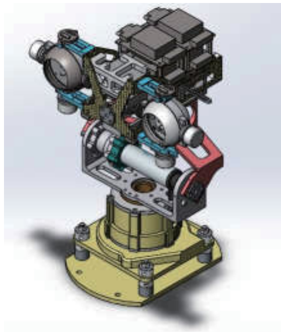

Fig. 1. Bionic Eye Vision Platform

The 9-DOF bionic eye vision platform as shown in Fig. 1 is designed with 3-DOF in the neck and 3-DOF in each eyeball[5]. The neck has three joints connected in series, while the range of yaw, pitch and roll angles can be up to 360°, ±60° and ±30°, which can achieve a wide range of perspective conversion. However, due to its relatively large moment of inertia, the maximum speed can only reach 2rad/s. The left and right eyeballs adopt the same structure of double sliding block + tandem mechanism. with the left eye as an example, the right eyeball also have three degrees of freedom: yaw, pitch and roll. Compared with the neck joints, the angular range of the eye's 3-DOF is small, the angles of yaw, pitch and roll are ±50°, ±50° and ±6.5°. But the eyeballs load is small, the maximum speed can be achieved as 5.5rad/s. This makes the eye movement more flexible than the neck. The mechanism is also designed to mimic the large, but relatively bulky, range of human neck rotation. However the rotation range of the eyeball is small, but the bionic design is very flexible, which makes the biomimetic eye can have a large range of motion and high flexibility.

*This work was supported by the National Key R&D Programm of China (Grant No.2017YFB1302200).

[1]Xiaopeng Chen is with the school of Mechatronical Engineering, Beijing Institute of Technology, and the Key Laboratory of Biomimetic Robots and Systems, Ministry of Education xpchen@bit.edu.cn

[2]ShaoWen Fu is with the National Key Laboratory of Science and Technology on Aerospace Intelligent Control, Beijing 100854, China

978-1-7281-0377-8/18/$31.00 © 2018 IEEE

II. RELATED WORK

Mechanical image stabilization is to use all kinds of machinery and equipment to ensure the stability of the camera platform, which has many applications in the video shooting. The principle is to detect the angle change between the axis and the angle of the base line in the horizontal or vertical direction of the platform through various sensor devices, and then change the angle movements of platform joints to an opposite direction through controller. In mechanic-based image stabilization the solutions is built around the camera module. One method that is purely mechanical is "steadicam", Another method is to build housing around the camera and compensate the vibration with the help of actuators[6]. Platform stability can be divided into primary stability and secondary stability[7]. The overall stability of the first level stabilization technology is to use frame system as photoelectric sensor platform. After amplifying the signal of the attitude angle change measured by gyroscope on the platform, the stability of the photoelectric sensor is maintained through driving the torque motor of the platform. The overall stability method can be divided into two-axis, three-axis and four-axis gyro stabilized platform, the main error is from torque error (such as friction moment, etc.) and sensor error (such as gyro drift, CCD visual axis installation error, etc.). There are some well-known mechanical stabilized platform such as Osmo in Dji and Karma Grip in GoPro.

Digital image stabilization system generally includes three main functional modules, namely motion estimation, motion correction and image compensation. Motion estimation is the first and most important step, through the analysis of the detected global motion vectors, the motion can be divided into random jitter or the normal scanning movement or both; Then through the motion correction of motion vector filter, we can extract unstable jitter (unintentional motion) and keep stable scanning component (intentional motion), then the compensating components are used to correct the image of each frame (warping); Finally, image compensation is used to compensate the undefined area caused by motion correction through interpolation, amplification, mosaicing and clipping, providing a complete full frame visual effect.

According to the complexity of the motion model based on the application scenario, video stabilization can be classified as 2D-based, 3D-based and 2.5D-based. The early video image stabilization technique uses the 2D image stabilization method, which adopts the 2D transform model (e.g., homography, affine) to represent the camera motion, and filtering through various low-pass filters. The 2d method is simple in calculation and robust in algorithm, but ignores the three-dimensional motion information of the camera itself, sometimes there may be parallel parallax problem, so the image stabilization effect is limited. The 3D image stabilization method can restore the 3D motion of the camera through 2D image information, Video rendering based on image-based rendering technology proposed by Buehler et al[8] is regarded as the earliest 3D rendering scheme, while LIU et al[9]. proposed the most classic 3D image stabilization scheme for preserving content, in which the traditional method of solving 3D motion estimation is motion to structure algorithm (structure from motion). In addition to SFM algorithm, Wang et al. [10]proposed a 3D approximate estimation method for handheld devices to stabilize images, which streamlines the estimation process of 3D motion of the camera to some extent, such as the characteristic trajectory method[11], multipath method [12]and epipolar constraint[13]. Since Hinton proposed the concept of deep learning in 2006, some people have tried to solve the problem of digital image stabilization by deep learning in recent years. Wang et al. [14]used the handheld stabilizer to collect data sets, and used the Siamese network to train and estimate frame to frame conversion parameters to solve the real-time image stabilization problem.

Our work: Based on the 9-DOF bionic eye vision platform, we propose a real-time image stabilization technique using mechanical and electronic hybrid fast motion compensation. First, the acceleration in the X, Y and Z direction of the platform is obtained by using the gyroscope and then converted to the rotation angle in the world coordinate system, which as the motion feedback to the image stabilizing decision unit. The first level of mechanical image stabilization is achieved by compensating the angular motion of the neck joint. Then, the motion matching estimation is carried out by the adjacent frame images after the first level stabilization, The gyro data is used again, and the jitter component between adjacent images is detected by pure digital image stabilization, and finally the second level of electronic stabilization is realized by filtering and motion compensation.

III. MECHANICAL IMAGE STABILIZATION

The bionic eye vision platform can be regarded as an inertial stabilization platform composed of gyroscope as the main device, and is essentially a speed servo control system[15]. The position regulator of traditional position closed-loop servo system, which is based on the input position and position feedback, generates the required speed instruction to the speed regulator, and finally controls the output speed of the motor and transmits it to the load side, and the feedback control based on the error is realized through the classical PID[16]. In general, this method is very useful. However, in case of disturbance, in order to improve the performance of interference suppression, we should give priority to the interference of system output[17]. However, when the velocity loop is disturbed, the instruction of the regulator output of the position loop should be the speed instruction after the disturbance, but not the motor side instruction before the disturbance, otherwise the influence of the disturbance will always exist, leading to insufficient control effect.

When external disturbances occur:

$$V_{gyro} = V_{motor} + V_{disturbance} \tag{1}$$

while the speed of the motor V_{motor} is the speed of the stable phase platform.

Mechanical stabilization based on the joints of the neck can obtain the speed instruction compensation by measuring the disturbance velocity. The corrected compensation is added to the position adjuster's instruction value and used as the instruction input of the speed adjuster. Thus the input of the drive's velocity loop becomes the difference between the position loop output and the gyroscope speed feedback, and the motor is more sensitive to the change of the disturbance signal.

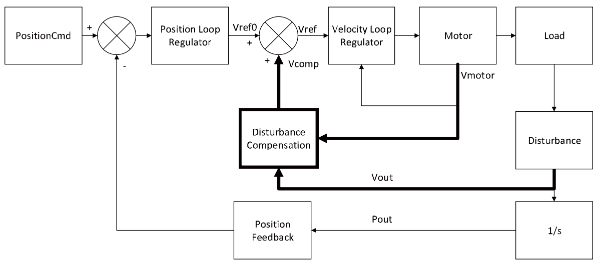

Fig. 2. Control Model

Fig. 2 shows the disturbance compensation link in the whole position closed loop. There are two inputs of the disturbance compensation link, One is the speed output of the motor side Vmotor, another is the final velocity output after the disturbance is added Vout. The output of the disturbance compensation link is Vcomp, the velocity adjuster's instruction compensation quantity. After increasing the disturbance compensation, the instruction input of the velocity regulator is changed from Vref0 to Vref, while Vref = Vref0 + Vcomp.

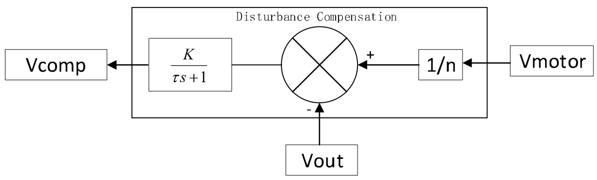

Fig. 3. Disturbance Compensation Model

Fig. 3 shows the internal composition of the disturbance compensation link. The input of the compensation Vcomp is the speed output of the motor side Vmotor and the final output speed vout. These two velocities are measurable. Assuming the motor's deceleration ratio is n, we first get their difference (Vmotor/n-Vout). Because of the measurement noise, we adopt the first order inertial link filter $\frac{1}{\tau s+1}$ to do smoothing, while τ is the time constant of the first order inertial link. We also add a gain K to that link, while K is a constant coefficient that takes the range of values as 0 1. When K = 0, the compensation is completely canceled. When K = 1, the complete compensation is achieved. K can be adjusted according to experience, the selection of its value is related to the system: when the external disturbance velocity is higher than the motor itself or the external disturbance has obvious disturbance to the system, then the weight is selected relatively large, and the value of K is close to 1. When the external disturbances are lower than one order of magnitude below compared with the motor feedback speed, the K value can be taken relatively small, such as 0.1; the specific K value is adjusted according to the actual effect.

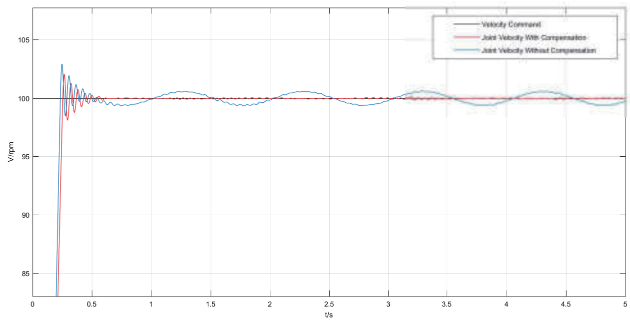

Fig. 4. Simulation Results In Matlab

The simulation model is built in the MATLAB environment, Add a sinusoidal disturbance to the input and the final result is shown in the Fig. 4, in which the black curve indicates the given velocity command, the red one represents the actual joint velocity after the disturbance compensation link is increased, and the blue curve represents the actual output joint speed without the disturbance compensation. Compared with the curves in Fig. 4, it can be seen that the joint velocity can be stabilized soon even after being disturbed, so as to realize the stability of the whole platform after the disturbance compensation link is added as feedforward.

IV. ELECTRONIC IMAGE STABILIZATION

Through the first level stabilization we has eliminated the big degree jitter of bionic eye vision platform, next we will eliminate the unintentional motion between sequence frame through the second level stabilization.

Based on the method of combining gyro data with image frame information, the paper first utilizes gyro data to obtain platform rotation information, and then detects the translation vector between adjacent frames by image processing, finally uses two-dimensional affine transformation model to eliminate image jitter.

A. Information Synchronization

In order to ensure the accuracy of the model transformation parameters, the image information and gyro data must be obtained at the same time, while the visual information and gyro information of the bionic eye platform are obtained by FPGA and MEMS gyroscope respectively. Therefore, the algorithm adopts quaternion spherical linear interpolation to realize the synchronization of image information and gyro data.

Quaternion[18] is a simple super plural, in computer graphics and computational physics applications, we use the unit quaternion to represent the rotation of three-dimensional space and directional. SLERP has a geometric equation that is independent of the dimensions of the space between the quaternions and the arc.

$$slerp(p0, p1; t) = (1 - t)p0 + t * p1 \tag{2}$$

Through the data collected by the gyroscope, the following sketch map can be obtained by interpolation analysis in MATLAB, so that the SLERP can realize the smooth interpolation transition of gyro data. Where red represents the original data point, the green triangle indicates the point after the interpolation.

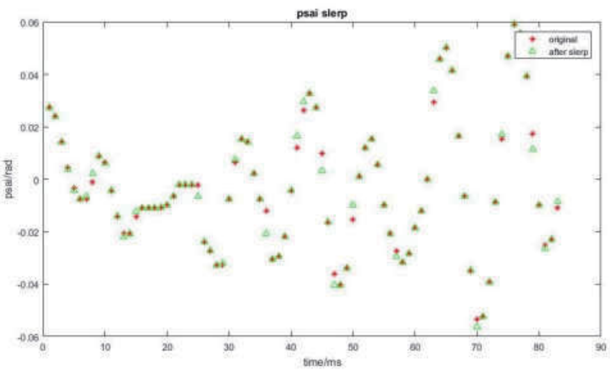

Fig. 5. Angle Interpolation with Slerp

B. Gyro Data and Model Parameters

After SLERP interpolation, the angle acquired by the gyroscope obtains the ϕ, θ and ψ at the same time as the video information, which represents the rotation angle of the axis Z, Y and X respectively, thus the rotational matrix of the current moment can be obtained:

$$\begin{bmatrix} cos\phi cos\theta & cos\phi sin\theta sin\psi - sin\phi cos\psi & cos\phi sin\theta cos\psi + sin\phi sin\psi \\ sin\phi cos\theta & sin\phi sin\theta sin\psi - cos\phi cos\psi & cos\phi sin\theta cos\psi + cos\phi sin\psi \\ -sin\theta & cos\theta sin\psi & cos\theta cos\psi \end{bmatrix} \tag{3}$$

The matrix R calculated at this time is the rotation matrix of the gyroscope coordinate system relative to the world coordinate system which can be written as ${}^{w}_{g}R$, while we need the rotation matrix of the world coordinate system relative to the camera coordinate system ${}^{c}_{w}R$, so we establish the geometrical model of bionic eye vision platform(Fig.6).

The rotation matrix of gyroscope coordinate system relative to camera coordinate system can be obtained according to the establishment of coordinate system:

$${}^{c}_{g}R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \tag{4}$$

The external parameter matrix of the camera can finally be obtained:

$${}^{c}_{w}R = {}^{c}_{g}R \ * \ {}^{w}_{g}R^{-1} \tag{5}$$

The conversion formula of world coordinate system and image coordinate system is as follows:

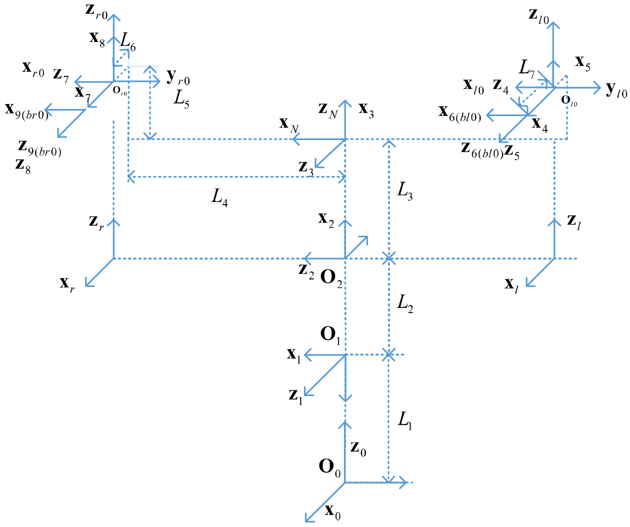

Fig. 6. D-H Model

$$x \ = \ K[R \ t]X \tag{6}$$

where x is the image coordinates, K and [R T] are the internal and external parameters of the camera respectively, and X is the world coordinate.

If the effect of translation vector is not taken into account, the corresponding image coordinates of original frame image I and stable frame image I' can be expressed as:

$$x \ = \ KRX \tag{7}$$

$$x' \ = \ KR'X \tag{8}$$

The transition relationship between the original frame I and the stable frame I' of the same image is[19]:

$$x' \ = \ KR' * K^{-1}R^{-1} * x \tag{9}$$

In the case of ignoring the translation, the transformation matrix $W = KR' * K^{-1}R^{-1}$ is a perspective transformation matrix, and the coordinate transformation between the original frame and the stable frame can be realized by using the perspective transformation model to eliminate the image jitter caused by the platform rotation.

C. Eliminate Image Shifting

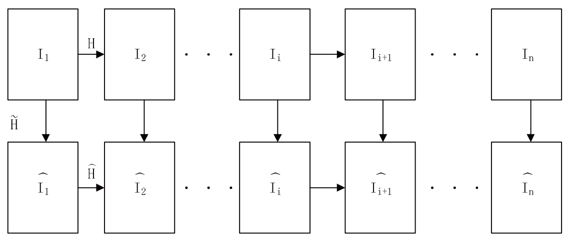

Fig. 8. Transformation Matrix

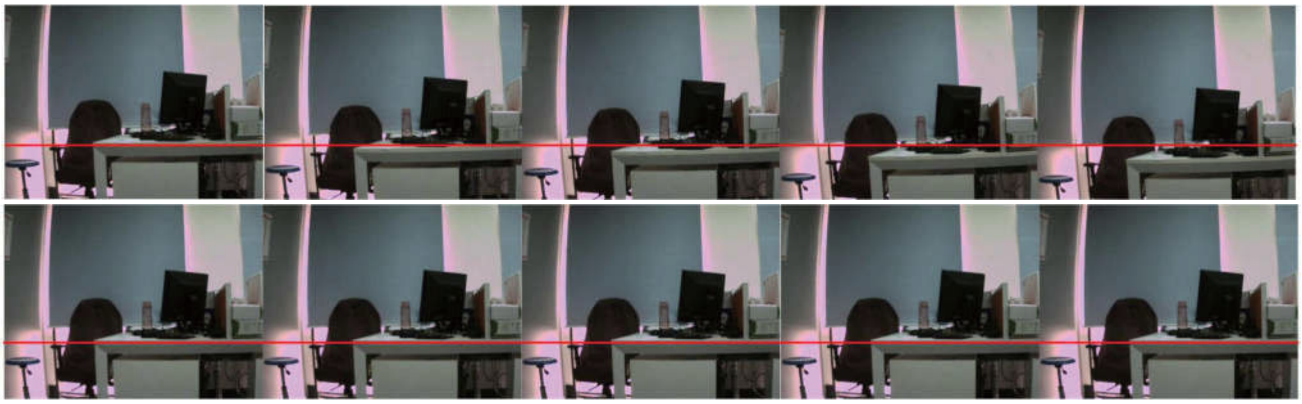

Fig. 7. Image Sequence (Top: original video. Down: stablized video)

The key points are detected by Shi-tomasi Corner detection, then the Pyramid Lucas-kanande algorithm is used to track the motion between adjacent frames, and the RANSAC algorithm is adopted to eliminate the key points without matching, thus the final affine transformation model parameter H is obtained. As shown in Fig. 8, the original video sequence is I, the transformation matrix between the adjacent image frames is H, the video sequence obtained after filtering is $\widehat{I}$, the transformation matrix between the adjacent image frames is $\widehat{H}$, and the transformation matrix between the original frame and the stable frame is $\widetilde{H}$.

$$I_{i+1} = {}^{i+1}_{i}H * I_i \tag{10}$$

$$\widehat{I}_{i+1} = {}^{i+1}_{i}\widehat{H} * \widehat{I}_i \tag{11}$$

$$\widehat{I}_{i+1} = {}^{i+1}_{i+1}\widetilde{H} * I_{i+1} \tag{12}$$

The transformation matrix between the original frame and the stable frame can be obtained by mathematical deduction:

$${}^{i+1}_{i+1}\widetilde{H} = {}^{i+1}_{i}\widehat{H} * {}^{i}_{i}\widetilde{H} * {}^{i+1}_{i}H^{-1} \tag{13}$$

The above formula can calculate the final transformation matrix $\widetilde{H}$, but at the same time because of the error accumulation, when the camera encounters large instantaneous migration, the matrix $\widetilde{H}$ will get the wrong data, therefore we adopt the scheme proposed by[20], which judging the situation of a large area of foreground motion and set the parameter threshold.

V. RESULTS AND DISCUSSION

A. Experiments Platform

The experiments on the image stabilization of the bionic eye Vision platform with nine degrees of freedom were carried out. The system functions can be separated into information processing and motion control. Information processing includes gyro data and image processing. The motion control part includes nine axis motion control, uses CANopen communication, constructs the ROS environment to be convenient to control the bionic eye platform. The hardware collects the image through the FPGA, uses the NVIDIA TX2 as the control machine. In running the image stabilization function, our mechanical image stabilization part only has the disturbance suppression to three joints of the neck, so the other 6 degrees of freedom of eyeball are in deadlock state.

B. Results and Analysis

In the case of good illumination conditions, we have done experiments in indoor and outdoor. We use the inter-frame transformation fidelity to evaluate the final video stability[21],a widely used evaluation metric of effectiveness and performance.

$$ITF = \frac{1}{N-1}\sum_{k=1}^{N-1} PSNR(k) \tag{14}$$

$$PSNR(k) = 10\,log_{10}\frac{I_{MAX}^2}{MSE(k)} \tag{15}$$

where N is the number of video frames, I_{MAX} is the maximum pixel intensity in the frame and MSE is the mean square error:

$$MSE(k) = \frac{1}{M*N}\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\left[I_k(i,j) - I'_k(i,j)\right]^2 \tag{16}$$

Where M*N represents the image size.

Table 1 ITF Value.

Video	Evaluation	Input	OurAlgorithm	Subspace
video1	ITF(dB)	14.8942	15.6485	15.6236
video2	ITF(dB)	13.0310	15.1381	13.7695

The top in Fig. 7 is the original input frame, the bottom is the final stabilized image frame, by comparing the table lines, we can see that the original images have obvious pixel movement, while the table line after image stabilization is basically in the same line. Video information saved after

(a) video 1

(b) video 2

Fig. 9. Video Background

real-time processing was detected in MATLAB, Its known that the calculation of the final ITF is 15.6485dB which is improved compare with the original input video, and we also use SubspaceStab to stabilized our video while the video1's ITF is also improved but less increase compared with out algorithm, especially even decreased in the case of poor background texture.

We take the video 1 in to analyse. The input original video is first in the state of no external fluctuation, and then we add disturbance to the platform to make the video frame unstable, and finally stops the jitter. The red and blue curve in Fig. 10 respectively corresponding to the original and stable frame pixel displacement in the direction of X, the following green and black curve corresponding to the original and stable frame pixel displacement in the direction of Y. The comparison shows that the real-time video stabilization processing will reduce the random jitter of video sequences caused by unintentional motion, the average pixel displacement in X and Y of this stabilized video are 0.0354 and -0.1327, the moving average of pixels is all within one pixel. Where camera image sensitive unit is 4.8um, focal length is 3.7mm, video frame rate is 30fps.

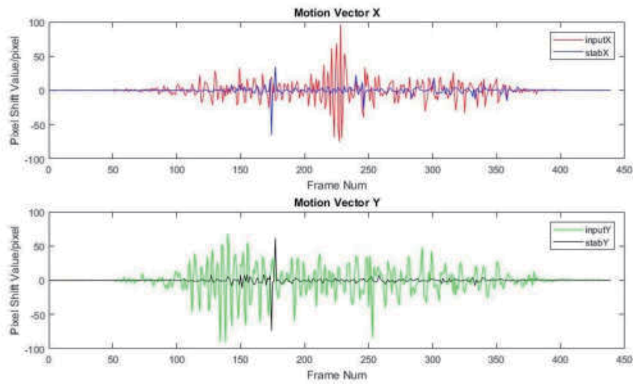

Fig. 10. Movement In The Direction Of X And Y

VI. CONCLUSIONS

We proposed a combined mechanical and electronic video stabilization method in real time for a robotic bionic eye. First, the mechanical stabilization is achieved by compensating motion disturbance of the neck joint. We measured the motion disturbance by analysis of the over all motion feedack and the dedicated motion applied to motor joints. Second, the 3D rotational motion of the camera as well as 2D translation motion in images are estimated and filtered to realize electronic image stabilization.

Experiments show that the two video's ITF of our method was increased by 5% and 15.9%, which is higher than the method of SubspaceStab, especially in the case of poor background texture. The frame rate after image stabilization can reach up to $30fps$, and the average image vibration the X and Y axes is about one tenth of one pixel. The experiments show that our algorithm performs well in image stabilization.

REFERENCES

[1] A, Gasteratos A, Kaburlasos V G, et al. Image Stabilization in Active Robot Vision[M]// Robot Vision. 2010:261-274.

[2] Sachs D, Nasiri S, Goehl D. Image Stabilization Technology Overview[J]. Invensense Whitepaper, 2006.

[3] Rawat P, Singhai J. Review of Motion Estimation and Video Stabilization techniques For hand held mobile video[J]. Signal & Image Processing, 2011, 2(2).

[4] Shanshan Wei, Wei Xie, Zhiqiang He. Overview of camera motion estimation techniques in emerging digital video image stabilization[J]. Application Research Of Computers, 2017, 34(2):321-327.

[5] Fan D, Chen X, Zhang T, et al. Design of anthropomorphic robot bionic eyes[C]// IEEE International Conference on Robotics and Biomimetics. IEEE, 2017:2050-2056.

[6] Gustavi, M. and L. Andersson (2017). Implementation of control algorithm for mechanical image stabilization.

[7] Yongyang Wang. Research on high precision optic axis stabilization technology based on fast mirror[D]. Graduate School of Chinese Academy of Sciences (Changchun Institute of Optics and Fine mechanics and physics), 2016.

[8] Buehler C, Bosse M, Mcmillan L. Non-metric image-based rendering for video stabilization[C]// Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on. IEEE, 2003:609.

[9] Liu F, Gleicher M, Jin H, et al. Content-preserving warps for 3D video stabilization[C]// ACM SIGGRAPH. ACM, 2009:44.

[10] Wang J M, Chou H P, Chen S W, et al. Video stabilization for a hand-held camera based on 3D motion model[C]// IEEE International Conference on Image Processing. IEEE, 2009:3477-3480.

[11] Liu F, Gleicher M, Wang J, et al. Subspace video stabilization[J]. Acm Transactions on Graphics, 2011, 30(1):4.

[12] Liu S, Yuan L, Tan P, et al. Bundled camera paths for video stabilization[J]. Acm Transactions on Graphics, 2013, 32(4):1-10.

[13] Goldstein A, Fattal R. Video stabilization using epipolar geometry[J]. Acm Transactions on Graphics, 2012, 31(5):1-10.

[14] Wang M, Yang G Y, Lin J K, et al. Deep Online Video Stabilization[J]. 2018

[15] Zhang T, Chen X, Owais H M, et al. Multi-loop stabilization control of a robotic bionic eyes[C]// IEEE International Conference on Cyborg and Bionic Systems. IEEE, 2017:87-90.

[16] Han J. From PID to Active Disturbance Rejection Control[J]. IEEE Transactions on Industrial Electronics, 2009, 56(3):900-906.

[17] Alagoz B B, Deniz F N, Keles C, et al. Disturbance rejection performance analyses of closed loop control systems by reference to disturbance ratio[J]. Isa Transactions, 2015, 55:63-71.

[18] Dam E B, Koch M, Lillholm M. Quaternions, Interpolation and Animation[J]. 1998.

[19] Karpenko, A., et al. (2011). Digital Video Stabilization and Rolling Shutter Correction using Gyroscopes.

[20] Yajin Xie, Huajun Feng, Zhihai Xu, Qi Li, Jiazi Huang, Yueting Chen, "A fast approach against large foreground motion in real-time image stabilization," Proc. SPIE 10462, AOPC 2017: Optical Sensing and Imaging Technology and Applications, 104620J (24 October 2017).

[21] Xu J, Chang H W, Yang S, et al. Fast feature-based video stabilization without accumulative global motion estimation[J]. IEEE Transactions on Consumer Electronics, 2012, 58(3):993-999.