A Fast and Robust Method for Pupil Center Estimation Using Chi-square Significance Test of Measurement Residual

Ming Zhai*† and Jianmin Wu*†
*National Key Laboratory of Science and Technology on Avionics System Integration, Shanghai
† China National Aeronautical Radio Electronics Research Institute, Shanghai
E-mail: tootoogo@hotmail.com

Abstract—**In this paper a low cost pupil center locating system was introduced and a robust pupil center estimation method was presented. Differing from other methods, eye image was converted into the space of chi-square significance of measurement residual (CSMR) first in the proposed method. In CSMR space, pupil area was enhanced while others were depressed. And then, a threshold selected by the statistical method was used to binarize the image in CSMR space. Finally pupil center was obtained by calculating mass center of the binarized image. Experiments showed that the proposed method was effective, even if in sunlight environment, pupil was small and images of eye were noisy.**

I. Introduction

Eye Tracking played an important role in many applications including human computer interaction, virtual reality, driver assistance, and diagnosis or early screening of some health problems. Head mounted eye tracking systems were more accurate than remote video-based systems, electrooculography (EOG) system [1], cornea reflection systems[2], and flying-spot laser based systems [3]. Eye movement was fast, and therefore low sampling rate led to obviously discontinuous eye movement in video. So most system used high sampling rate, e.g., in SMI system [4], sample rate exceeded 400 Hz. But hardware with high sample rate was expensive. In order to set up a low cost eye tracking system, web camera was a good choice for low price. The sampling rate of web camera was low, so it was more feasible to detect pupil than track pupil. In order to detect pupil, some used user-set threshold [5] or threshold obtained by histogram based method to binarize images [6, 7]; some based on Hough transformation based method [8] or fitted iris with ellipse to estimate pupil center [18]; some based on statistical models, PCA and neural classifier based method [9,10]; some based on intensity difference of pupil in multiple light sources [17]. However, using user-set threshold was inconvenient. The threshold obtained by histogram based method was effective, but small pupil size would bring trouble to histogram based method. As to methods based on Hough transformation, these methods were often time consuming. PCA and neural classifier based methods were often not accurate. Methods based on intensity difference of pupil in multiple light sources were limited by environments and devices. Moreover most methods were suit for detecting pupil indoor, because in indoor environment pupil size was often larger which made detection easier. If in outdoor environment pupil size was small or eye was half-closed, it would bring more troubles.

In this paper a low cost pupil center location system was introduced and a pupil location estimation method was presented for both indoor and outdoor environment. Due to low cost, images obtained by our system were often noisy or even blurred. However, most of these problems could be solved by chi-square significance of measurement residual based detection technique. Firstly eye images were captured by the web camera. And then the images was converted into the space of chi-square significance of measurement residual (CSMR). In CSMR space, pupil area was enhanced while other areas were depressed and moreover threshold selected by a statistical method to binarize the image in CSMR space. At last pupil center was obtained by calculating the mass center of the binarized image. We also compared our method with the two methods presented by Chirayuth [6] and Cho [7] respectively. The merit of our method was the robustness in rigorous situation. Furthermore when eye was closed, nothing would be detected by our method, this property could be used to detect eye blinking.

This paper was organized as follows: Section I the introduction; in Section II some basic notions were reviewed and the implementation of our technique was introduced; Section III experimental results; Section IV conclusions.

II. Pupil Center Estimation

Our pupil center locating system was a low cost system. Due to low cost, images obtained by our system were noisy or even blurred. After eye images were obtained, the images were transformed in to CSMR space. Pupil center was obtained by calculating the mass center of the binarized image in CSMR.

A. Hardware Configuration

Our pupil center locating system was shown in Fig. 1. A glass with infrared reflective film was fixed near volunteer's eye. Most of visible light could go through the glass while most of infrared light could be reflected. An infrared camera which was a web camera fixed with an infrared filter to absorb visible light was fixed on volunteer's head to capture eye images of the volunteer. An infrared LEG light was fixed around the web camera to illuminate if in dark environment. Due to low cost, images obtained by our system were often noisy or even blurred, it would bring troubles to most methods mentioned in previous section, but our technique was suit for this condition.

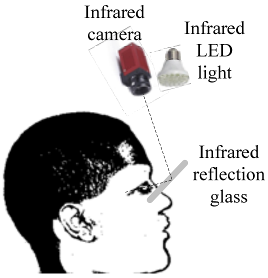


Fig. 1. Configuration of pupil center locating system.

B. Kalman filter and Chi-square Significance Test of Measurement Residual

The Kalman filter was a linear estimator based on the minimum mean squared error (MMSE) [11, 12], for the sake of less computation and better performance it had been widely used since it was introduced. Here was a brief review of Kalman filter.

Given a system equation and measurement equation as follows:

$$\begin{aligned} x(k+1) &= F(k)x(k)+G(k)u(k)+v(k) \\ z(k+1) &= H(k+1)x(k+1)+w(k+1) \end{aligned} \tag{1}$$

where x and z were state vector and measurement vector respectively; v and w were zero mean Gaussian noise with covariance Q and R; u was known input vector; F and H were state matrix and measurement matrix. F, G, H, Q and R were assumed known and possibly time-varying. The two noise sequences and the initial state were assumed mutually independent. The above constituted the linear Gaussian (LG) assumption.

The main equations of Kalman filter were as follows:

$$\begin{aligned} &\overline{x}(k+1|k) = F(k)\overline{x}(k|k)+G(k)u(k) \\ &\overline{z}(k+1) = H(k+1)\overline{x}(k+1|k) \\ &\mu(k+1) = z(k+1)-\overline{z}(k+1|k) \\ &P(k+1|k) = F(k)P(k|k)F(k)'+Q(k) \\ &S(k+1) = H(k+1)P(k+1|k)H(k+1)' \\ &+R(k+1) \\ &\overline{x}(k+1|k+1) = \overline{x}(k+1|k) \\ &+W(k+1)\mu(k+1) \\ &W(k+1) = P(k+1|k)H(k+1)'S(k+1)^{-1} \\ &P(k+1|k+1) = P(k+1|k) \\ &-W(k+1)S(k+1)W(k+1)' \end{aligned} \tag{2}$$

where the expression $\overline{A}(k+1|k)$ denoted the estimate of A at time k+1 estimated at time k; P and S were state covariance and innovation covariance; W and μ were filter gain and measurement residual (MR).

In Kalman filter, some kinetic models were often used to describe the kinetic property of the target, such as the constant velocity (CV) model and the constant acceleration (CA) model [12]. Consider a target moving with constant velocity, CV model was suitable to describe the kinetic property of the target. In CV model, state vector was $x=(X,X')^T$, X was looked on as displacement, X' was the differential of X, i.e. velocity. If measurement was displacement then measurement matrix was $H=(1\quad 0)$. Otherwise if measurement was velocity then measurement matrix was $H=(0\quad 1)$. In CV model state matrix was $F=\begin{pmatrix}1 & T\\ 0 & 1\end{pmatrix}$, T was a time interval, and $u(k)$=0.

The measurement residual based chi-square test was often used to detect target maneuver in maneuver target tracking (MTT) applications [13-16]. With linear-Gaussian assumption, the measurement residuals of a Kalman Filter were zero mean, Gaussian distributed and white; i.e., $\mu(k)\sim \mathrm{N}(0,S(k))$ and

$$\begin{aligned} S(k) &= \mathrm{cov}(\mu(k)) \\ &= R(k)+H(k)P(k|k-1)H(k)' \end{aligned} \tag{3}$$

Furthermore

$$\xi(k)=\mu^T(k)S^{-1}(K)\mu\ (k) \tag{4}$$

was a chi-square distributed variable, which was $\xi(k)\sim\chi^2_{n_z}$ where $n_z=\dim(\mu)$ was the dimension of vector μ. If

$$\xi(k)>\chi^2_{n_z}(\alpha) \tag{5}$$

then decided a maneuver occurs, where $1-\alpha$ was confidence level. The decision which was based on a single sampling time in Equation (4), could be replaced by a moving average (or moving sum) of the normalized innovations squared over a sliding window of s sampling times

$$\xi^s(k)=\sum_{j=k-s+1}^{k}\xi(j) \tag{6}$$

The above was chi-square distributed with sn_z degrees of freedom. Alternatively, a fading memory average (also called exponentially discounted average)

$$\xi^\beta(k)=\beta\xi^\beta(k-1)+\xi(k) \tag{7}$$

where $0<\beta<1$ and with initial condition $\xi^\beta(0)=0$, could be used. The variable ξ^β was approximately distributed as $\frac{1}{1+\beta}\chi^2_{n_\beta}$ with $n_\beta=\frac{1+\beta}{1-\beta}n_z$. Its effective window length was $\frac{1}{1-\beta}$, and e.g. for $\beta=0.95$ one had s=20.

C. Implementation

Assuming image I with M row and N column was a frame obtained by our system, e.g. Fig. 2 B. The histogram of image I was shown in Fig. 3 which showed that it was difficult to binarize image I with histogram based threshold method. Before implementation, image I was smoothed by a Gaussian mask to reduce noise first. The intensity of every column (row) of image I constituted a Gaussian sequence and was looked on as measurement sequence of velocity in CV model. Under these conditions, $H=(0\quad 1)$, state matrix was

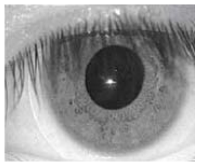
$F = \begin{pmatrix} 1 & T \\ 0 & 1 \end{pmatrix}$, T was a time interval, $u(k)$ =0, and variance R should be set a bit larger than the variance of the intensity of image I. Every column was filtered by a Kalman filter from top to bottom, for the N columns, the N Kalman filters were used, and meanwhile the measurement residual (MR) of every position from the Kalman filters were obtained to calculate the chi-square significance of the MR using Equation 7 and the result was stored into the matrix Σ. The matrix Σ was the CSMR space of image I. The CSMR space of Fig 2B was shown in Fig 4. Binarize Σ by threshold selected by Equation 5, we obtained pupil mask Σ_{mask} in which pupil area was labeled with 1, otherwise labeled with 0. Due to the delay reaction of chi-square significance of MR, that was, the change of chi-square significance of MR was often behind the change of image I. It would result in imprecise location estimation. Pupil was often symmetrical, so we filtered image I from the 4 directions: from top to bottom; from bottom to top; from left to right; from right to left, that meant $2M+2N$ filters were used. Denoted Σ_{up}, Σ_{down}, Σ_{left}, Σ_{right} as pupil mask of image I come from the four filtering directions respectively. Computed the logical OR of the four mask

$$\Sigma_{mask}(i,j) = \Sigma_{up}(i,j) \,|\, \Sigma_{down}(i,j) \,|\, \Sigma_{left}(i,j) \,|\, \Sigma_{right}(i,j) \qquad (8)$$

where the operator " | " was logical OR, $\Sigma(i,j)$ was the (i,j)th element of Σ. Pupil mask Σ_{mask} was the final mask and pupil area of image I was labeled in Σ_{mask}. Although pupil area labeled in Σ_{mask} was a bit larger than the real pupil area but they had the same center.

Due to mirror reflection, sometimes there was a light spot near pupil, e.g., the white spots in the pupils of Fig. 1A and Fig. 1B, and the spot often was labeled as pupil area. Before pupil center estimation, the spot should be removed from the pupil mask first. The intensity of the spot area was high so labeled the (i,j)th element of the pupil mask with zero if the intensity of the (i,j)th element of image I was higher than the average intensity of image I. After that pupil center was obtained by calculating the mass center of pupil mask.

Kalman filter needed initialization, that was, determined the first state vector $x(0)$ and the first state covariance $P(0|0)$. Here set $x(0) = (0, X'_0)^T$, where X'_0 was the average intensity of image I. It was an alternative choice to set X'_0 with the intensity of the pixel which the filter started at. Set the first state covariance $P(0|0)$= $\begin{pmatrix} R & R/T \\ R/T & 2R/T^2 \end{pmatrix}$. It was an alternative choice to set $P(0|0)$=$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.



A

B

Fig. 2. samples of eye images,
A . sample in CASIA-IrisV3-Interva B. sample obtained by our low cost system

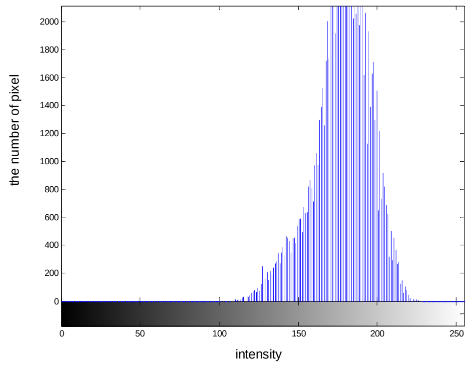


Fig. 3. The histogram of Fig. 2 B.

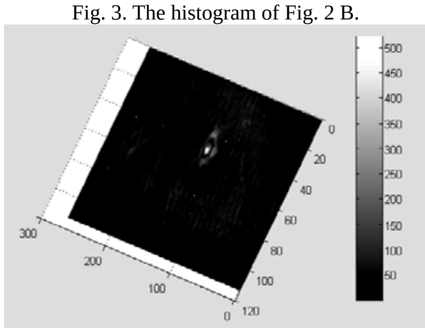


Fig. 4 CSMR space of of Fig. 2 B

III. Experimental Results

First we compared our method with the two methods presented by Chirayuth [6] and Cho [7] respectively with the samples from the database of CASIA-IrisV3-Interval and CASIA-IrisV3-Twins. One sample of CASIA was show in Fig. 2A. CASIA-IrisV3-Interval was created in an indoor environment while CASIA-IrisV3-Twins was created in outdoor environment. Cho used histogram based binarization and Pythagorean Theorem to estimate the center of pupil. Chirayuth used illumination normalization first and then binarized the image, and finally pupil center was estimated by elliptical model fitting. The performance comparison was shown in Table 1. In Table 1, E_x and E_y were the average error in x-direction and y direction respectively, which were computed by

$$E_x = \frac{1}{n}\sum_{i=1}^{n} | x_d^i - x_m^i | \text{ and } E_y = \frac{1}{n}\sum_{i=1}^{n} | y_d^i - y_m^i | \qquad (9)$$

where (x_d^i , y_d^i)pupil center coordinates in the *i*th samples and (x_m^i , y_m^i) was pupil center obtained manually which was considered to be right and precise. Of the three methods, Chirayuth's methods seemed to perform best in precision and the results of our method and Cho's method were similar. Chirayuth's method produced failures in both database while Cho's method produced failures only in outdoor database. Our method did not produce failures.

We also compared the three methods in the sample images obtained by our low cost system and the performance comparison was shown in Table 2. Due to low cost, sample images obtained by our system were noisy comparing with samples in database CASIA. The performance of three methods in precision was similar to the former experiments. Chirayuth's method and Cho's method produced many failures in the samples while our method produced only a few failures. It showed that the merit of our method was the robustness in rigorous situation.

database	Our method			Cho's method			Chirayuth's method		
	E_x	E_y	Failure(%)	E_x	E_y	Failure(%)	E_x	E_y	Failure(%)
CASIA-IrisV3-Interval (indoor)	0.29	0 .30	0	0.21	0.23	0	0.15	0.19	0.6
CASIA-IrisV3-Twins (outdoor)	0.41	0.37	0	0.33	0.35	0.3	0.19	0.22	5.6

Table 1 The performance comparison in CASIA-IrisV3-Interval and CASIA-IrisV3-Twins

database	Our method			Cho's method			Chirayuth's method		
	E_x	E_y	Failure(%)	E_x	E_y	Failure(%)	E_x	E_y	Failure(%)
Samples captured by our low cost system (outdoor)	0.45	0 .50	3.3	0.42	0.39	13.3	0.25	0.29	20.6

Table 2 The performance comparison in Samples obtained by our low cost system

IV. Conclusions

The proposed method converted eye image into the space of chi-square significance of measurement residual (CSMR). Threshold selected by a statistical method was used to binarize eye image in CSMR space. At last pupil center was obtained by calculating the mass center of the binarized image. When image was converted into CSMR space, pupil area was enhanced while other areas were depressed. So it was much more effective to binarize the image in CSMR space than in the origin image. The proposed method was robust for many scenarios especially when in the sunlight the pupil was small. The merit of our method was the robustness in rigorous situation. Furthermore when eye was closed, nothing would be detected by our method, this property could be used to detect eye blinking. Moreover in the proposed method the four filtering processes of the four directions were mutually independent, so the four processes could be carried out in parallel. If carried out in parallel, the process time could reduce three quarters and therefore it made our technique practicable for real-time applications. For example, on a PC with Intel Q9400 CPU (quad core, 2.66 GHz), it took less 30ms to deal with an image with size of 1024x768, if in parallel mode.

Reference

[1] A. Duchowski, "Eye-Based Interaction in Graphical Systems: Theory and Practice" available at http://vret.ces.clemson.edu/sigcourse/notes/c05-TOC.pdf, Clemson University

[2] S. Shih, Y. Wu, and J. Liu, "A calibration-free gaze tracking technique", in Proceedings of the 15th international conference on pattern recognition, Barcelona, Spain, vol. 4, pp. 4201-4205, 2000.

[3] K. Irie, B. A. Wilson, R. D. Jones, P. J. Bones, and T. J. Anderson, "laser-based eye-tracking system", in Behavior Research Methods, Instruments, & Computers, vol. 34, no. 4, pp. 561-572, Nov., 2002.

[4] Sensor Motoric Instruments, "Fastest Video-Based Eye Tracking", available at http://www.smi.de/.

[5] Xindian Long, "A High Speed Eye Tracking System with Robust Pupil Center Estimation Algorithm", Engineering in Medicine and Biology Society, 2007 Annual International Conference of the IEEE.

[6] Chirayuth Sreecholpech, "Circular and Elliptical Modeling for Pupil Boundary in Closed-up Human Eye Images", Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology, ECTI-CON 2008. 5th International Conference, Volume 1, May 2008 Page(s):445 - 448

[7] J. M. Cho, "A pupil center detection algorithms for partially-covered eye image" TENCON 2004, 2004 IEEE Region 10 Conference, Volume A, 21-24 Nov. 2004 Page(s):183 - 186 Vol. 1

[8] Zhifei Xu, "A Robust and Accurate Method for Pupil Features Extraction" Pattern Recognition, 2006, ICPR 2006 18th International Conference, Volume 1, Page(s):437 - 440

[9] Ioana Bacivarov, "Statistical Models of Appearance for Eye Tracking and Eye-Blink Detection and Measurement", Consumer Electronics, IEEE Transactions, Volume 54, Issue 3, August 2008 Page(s):1312 - 1320

[10] Vidas Raudonis, "Discrete Eye Tracking Based on PCA and Neural Classifier", Human System Interactions, 2008 Conference, May 2008 Page(s):237 - 241

[11] Kalman, R.E., A New Approach to Linear Filtering and Prediction Problems. Transaction of the SME—Journal of Basic Engineering, 1960: p. 35-45.

[12] Yaakov Bar-Shalom, X.-r.L., Estimation with Applications To Tracking and Navigation. 2001, New York: A Wiley-Interscience Publication.

[13] Y. Bar-Shalom and K. Birmiwal. "Variable Dimension Filter for Maneuvering Target Tracking". IEEE Trans. Aerospace and Electronic Systems, AES-18(5):621–629, Sept. 1982.

[14] S. S. Blackman. Multiple Target Tracking with Radar Applications. Artech House, Norwood, MA, 1986.

[15] J. R. Cloutier, C. F. Lin, and C. Yang. "Maneuvering Target Tracking via Smoothing and Filtering Through Measurement

Concatenationn". AIAA Journal of Guidance, Control, and Dynamics, 16(2):377–384, March-Apr. 1993.
[16] X. R. Li and V. P. Jilkov. "A Survey of Maneuvering Target Tracking—Part IV: Decision-Based Methods". In Proc. 2002 SPIE Conf. on Signal and Data Processing of Small Targets, vol. 4728, Orlando, Florida, USA, April 2002
[17] C. H. Morimoto, D. Koons, "Pupil detection and tracking using multiple light sources", Image and Vision Computing, Volume 18, Issue 4, 1 March 2000, Pages 331-335.
[18] Jian-Gang Wang, Eric Sung, "Gaze determination via images of irises", Image and Vision Computing, Volume 19, Issue 12, 1 October 2001, Pages 891-911