On Gaze Estimation Using Integral Projection of Eye Images

Lan-Rong Dung[1(✉)], Yu-Cheng Lee[1], and Yin-Yi Wu[2]

[1] National Chiao Tung University, Hsinchu 30010, Taiwan
`lennon@nctu.edu.tw`
[2] National Chung-Shan Institute of Science and Technology, Taoyuan, Taiwan

Abstract. This paper presents a gaze estimation algorithm using integral projection of eye images with advantage of low additional hardware requirement and low computational power. The algorithm needs only a webcam under nature light source and captured eye images in a non-intrusive way. Before integral projection, we used binarization process to eliminate the non-related image information to gaze position. Projected on binary eye images with projection adjustment method to avoid eye tilt makes projection error and defined the accurate integral range of eye ROI images to achieve robust gaze estimation. We analyzed the projection diagram with skewness to describe the variation of different gaze position. In skewness calculation, the pixel coordinate of eye ROI images has been normalized to avoid head moved back and forth makes the size of ROI changed. In horizontal direction, the error angle of our algorithm is 2.29°, maximum error angle is 4.8° and the resolution we defined is 7.5. Because our algorithm is inaccurate in vertical, we could only estimate gaze direction, but to estimate precise angle. The computational power of our algorithm is low, the average execution time of each frame is only 0.01652 s, only 24% of opponent.

Keywords: Gaze estimation · Integral projection · Skewness

1 Introduction

Gaze estimation is an important technique in human-computer interaction area. The gaze information reveals what we are paying attention to and our mental state. This technique could apply in several areas in our daily life such as assistive techniques, automotive, learning ability research and advertisement research.

The gaze estimation and eye tracking techniques have developed several decades. It could divide into four categories, Electro-Oculography, Scleral Search Coils, Infrared-Oculography and Video-Oculography [8]. In these four categories, video-based method is the most widely used method, because it could get and develop easier than other three categories and it won't causes potential danger to our eye. The method we proposed is also video-based method, so we will focus on this method in the following introduction. The video-based method can subdivided into two categories, feature-based method and appearance-based method. Feature-based method using characteristics on our eyes like iris, pupil, corneal reflections and eye corner to estimates eye

movement. In Yang's [1] paper, they proposed a gaze estimation algorithm with near-infrared light sources. With precise detecting the pupil center and using the position between pupil center and corneal reflections, it could estimate the eye gaze on the screen. Appearance-based method using template matching and training large samples to estimate eye gaze directions. In Raudonis's [2] paper, they proposed an algorithm using PCA method to find the six components of eye images first and using ANN (Artificial Neural Network) to classify the position of pupil, when calibration procedure they could get the training samples to matching the gaze direction.

No matter feature-based method or appearance-based method have some defects. Feature-based method needs additional hardware like special camera and near-infrared light sources. Precise detecting pupil center costs lots of computational power. Appearance-based method needs lots of samples to training. Because of these problems this paper proposed a gaze estimation algorithm using integral projection of eye images. With advantages of low additional hardware requirement and computational power, estimating eye gaze in a non-intrusive way. The horizontal direction average error angle of our algorithm is 2.29°, the maximum error angle is 4.8°. The resolution we defined is 7.5, it means in the case of zero error angel, the horizontal direction of screen could divided into 7.5 sections. Because of the limitation of our algorithm, in the vertical direction the algorithm we proposed could only estimate directions but to estimate precise gaze angle. We will introduce the algorithm details in the following sections. Video object tracking is an important topic within the field of computer vision. It has a wide range of applications such as human-vehicle navigation, computer interaction, etc. Various approaches for object tracking have been proposed. [1] proposed a tracking method based on mean shift. It maximizes the similarity iteratively by comparing the color histogram of the object. The advantage is the elimination of a brute force search and low computation. [2] extended to 3D domain, combines color and spatial information to solve the problems of orientation changing and small scale changing. [3] used stochastic meta-descent optimization method. It can track fast moving objects with significant scale change in a low-frame-rate video.

2 Related Work

The face and eye detection are using the haar-cascade classifiers. Because the eye position on face will not change, we use this characteristic to set the ROI of eye images and don't need to re-detect eye position every frame.

In mathematics and statistics, moment is a kind of indicator to measure the morphological characteristics of a point set, includes first moment, second moment, third moment and fourth moment [3]. The first moment means mean of distribution, second moment means standard deviation, third moment means skewness and fourth moment means kurtosis. In our algorithm, we will use skewness to estimating eye gaze.

Skewness is a measurement of asymmetry of the probability distribution. According to Pearson's moment coefficient of skewness, the skewness γ defines in (1).

$$\gamma = E\left[\left(\frac{X - \mu}{\sigma}\right)^3\right] = \frac{E[X^3] - 3\mu\sigma^2 - \mu^3}{\sigma^3} \tag{1}$$

σ is standard deviation, X is a random variable, E is the expect value operator and μ is mean.

3 Proposed Method

In this paper, we proposed a gaze estimation algorithm using integral projection of eye images. In Fig. 1, it shows the structure of our algorithm. It could divided into three main processing stages, pre-processing and integral projection, integral projection adjustment and projection diagram analysis.
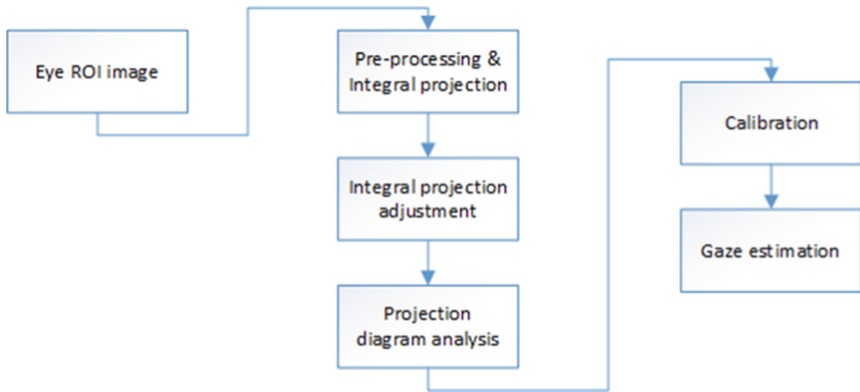
Fig. 1. The flow chart of proposed algorithm.

Integral projection function is proposed by Zhou [4], the mathematic function shows in (2) and (3).

$$IPF_v = \int_{y1}^{y2} I(x, y)dy \tag{2}$$

$$IPF_h = \int_{x1}^{x2} I(x, y)dx \tag{3}$$

IPF_v is vertical projection function and IPF_h is horizontal projection funtion. $I(x, y)$ is the gray scale value on image coordinate (x, y).

Figure 2 shows the horizontal and vertical projection of eye image. In pre-processing stage, we will use binarization process. We found that using integral projection based on gray scale image may have some problems. There is some image information which is not related to eye gaze like skin. If we projected using gray scale image, it may causes error in gaze estimation stage. So before integral projection, we binarized the image by Otsu's thresholding [5] method and projecting only black pixels, the following projection diagram analysis will focused on binary image. The projection result shows in Fig. 3.
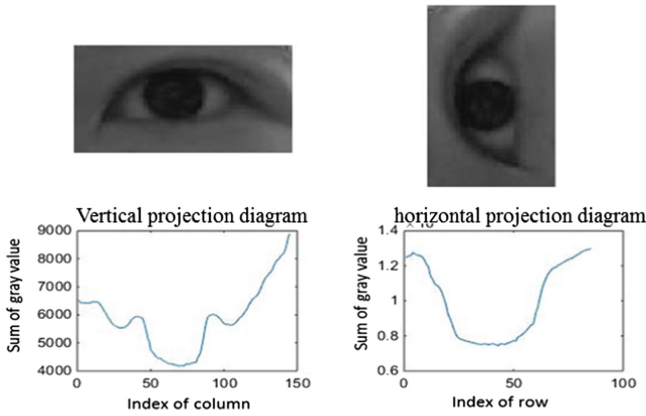

Fig. 2. The horizontal and vertical projection on gray scale image.
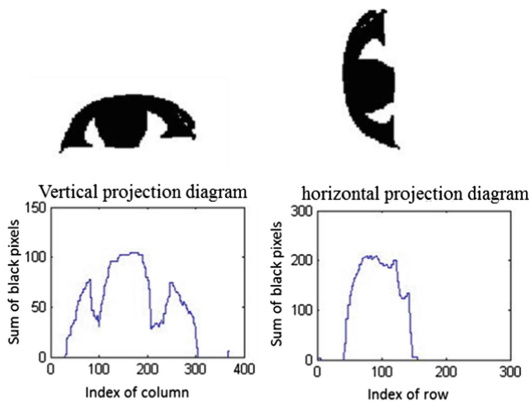

**Fig. 3.** The horizontal and vertical projection on binary image.

The projection is still on whole ROI image of eye now, but the head's slightly movement may cause eye tilting and makes projection error. Because of these problems we will need to adjust projection surface by using features on our eyes.

In integral projection adjustment stage, we proposed two adjustment methods, one is ellipse fitting method. Human's eye could approach to an ellipse, using ellipse fitting [9] to fit the best ellipse of eye. After get the ellipse, we could reset the ROI image size of eye, the long axis and short axis of ellipse becomes the integral projection surface and integral range of horizontal and vertical direction. We calculate the angle θ between ellipse long axis and horizontal plane and the pixel coordinate (x, y) on original ROI image could transform to the new coordinate (x′, y′). The relationship between (x, y) and (x′, y′) shows in (4).

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \tag{4}$$

After pixel coordinate transformation, we could do the integral projection on the eye ROI image. Another adjustment method is canthus line, human's canthus is a good reference point, so we detecting canthus by FAST-corner [6] detection. It used a 7×7 mask scanning on image by comparing the gray scale value of center pixel and other sixteen pixels around it.

The detection result shows in Fig. 4(a), the green points are possible canthus position. Because our canthus may locate on the furthest position of eye, it means it will locate on the endpoints of ellipse long axis. So, we using the ellipse fitting result shows in Fig. 4(b) to get the final canthus points, it shows in Fig. 4(c). After we get the canthus, the canthus line and its orthogonal line becomes horizontal and vertical projection range and reset the ROI of eye image size. We calculated the angle θ between canthus line and horizontal plane, and do the pixels coordinate transformation which we have mentioned in ellipse fitting method. In the following sections we will compare two different projection adjustment methods and choose the better one to become a part of our algorithm.

(a)   (b)   (c)
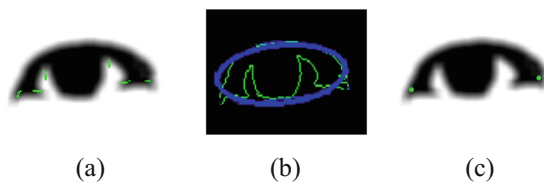
Fig. 4. The corner detection by FAST algorithm and the canthus point selection by ellipse fitting.

After projection adjustment, we could get the projection diagram. The projection diagram reveals the gaze information, Fig. 5 shows horizontal and vertical projection diagram when gaze direction is changed. In the projection diagram analysis stage we used skewness to describe the projection diagram.
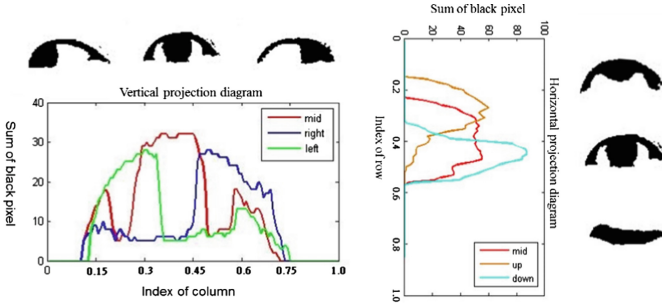
**Fig. 5.** The horizontal and vertical projection diagram under different gaze direction.

Before calculating skewness, we need to calculate the mean position of the projection diagram first. One important thing is, when head moved back and forth in front of camera, the width and height of eye ROI images will changed, it may cause error if we calculated mean of projection diagram by exact index value. To avoid this problem, the pixel coordinate of eye ROI images have been normalized. According to the formula we mentioned above, (5) and (6) show the calculation of mean values.

$$\mu_v = E[x] = \sum xP(x) = \frac{\sum (x_i/width) \cdot y_i}{\sum y_i} \tag{5}$$

$$\mu_h = E[y] = \sum yP(y) = \frac{\sum (y_i/height) \cdot x_i}{\sum x_i} \tag{6}$$

μ_v is the mean of vertical projection diagram, μ_h is the mean of horizontal projection diagram, width is the width of eye ROI image, height is the height of eye ROI image. After we have mean, we could calculated the skewness of projection diagram to describe the degree of skew when people's gaze point is changing and estimated gaze position by skewness value. According to the formula we mentioned above, (2) shows the calculation formula of skewness γ.

We have introduced integral projection and two projection adjustment methods. To evaluate which adjustment method is better, we conducted an experiment. There are ten volunteers, each person measured three times, collected total thirty samples to evaluate the result. The distance between screen to volunteers' eye is 40 cm, every gaze points on screen shows in Fig. 6, the angle between each gaze point is 3°. The red calibration point is used to normalized the data set, we will discussed how does the mean and skewness we measured under the regular gaze angle and evaluated the effect of pre-processing method and projection adjustment methods.
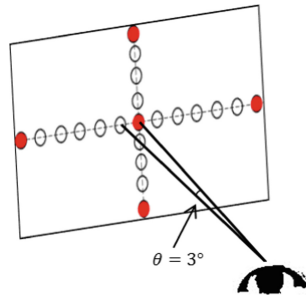
Fig. 6. The schematic diagram of gaze points on screen (red points are calibration points) (Color figure online)

Figures 7 and 8 show the results of skewness analysis. The horizontal axis of plot is gaze angle, positive is left direction gaze points, negative is right direction gaze points and zero is midpoint. The vertical axis of plot is normalized mean and the black line is the ideal standard line.
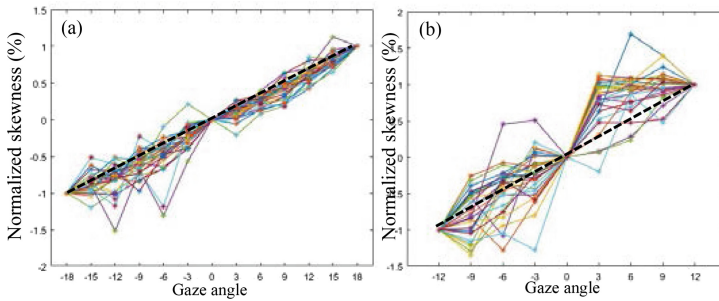

Fig. 7. The skewness analysis of (a) horizontal (b) vertical direction with projection adjustment of ellipse fitting method.
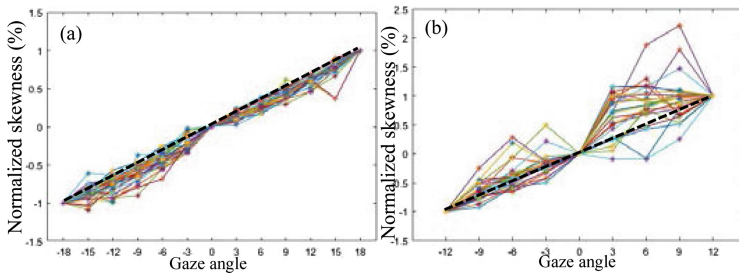
Fig. 8. The skewness analysis of (a) horizontal (b) vertical direction with projection adjustment of canthus line methods.

We estimated the execution time every frame and compared to another eye tracker proposed by Ferhat [7], under the same hardware and environment, CPU: Intel Core i5-4570; Ram: 8 GB; OS: Ubuntu 12.04. The comparison of execution time shows in Fig. 9.
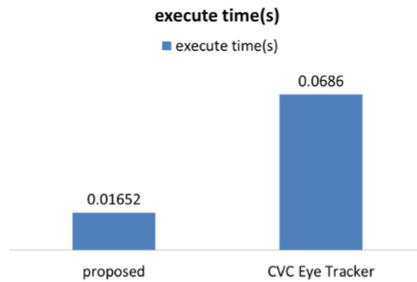
Fig. 9. Comparison of execute time.

4 Conclusion

This paper presents a gaze estimation algorithm with only a webcam under nature light sources and using integral projection of binary eye images with projection adjustment by canthus line method to achieve robust gaze estimation. We analyzed the projection diagram by skewness to describe the diagram characteristics. The average error angle in horizontal is 2.29° and the resolution we defined is 7.5. In vertical direction, because of the limitation of our algorithm, it could only estimate gaze directions but couldn't estimate precise gaze angle. The computational power of our algorithm is low, the average execution time of each frame is 0.01652 s, only 24% of opponent.

References

1. Yang, X., Sun, J., Liu, J., Chu, J., Liu, W., Gao, Y.: Agaze tracking scheme for eye-based intelligent control. In: 8th WCICA, pp. 50–55 (2010)
2. Raudonis, V., Simutis, R., Narvydas, G.: Discrete eye tracking for medical applications. In: 2nd ISABEL, pp. 1–6 (2009)
3. Joanes, D.N., Gill, C.A.: Comparing measures of sample skewness and kurtosis. J. R. Stat. Soc. **47**, 183–189 (1998)
4. Zhou, Z.H., Geng, X.: Projection functions for eye detection. Pattern Recogn. **37**, 1049–1056 (2004)
5. Otsu, N.: A threshold selection method from gray-level histograms. IEEE Trans. Syst. Man Cybern. **9**(1), 62–66 (1979)
6. Rosten, E., Drummond, T.: Fusing points and lines for high performance tracking. IEEE Int. Conf. Comput. Vis. **2**, 1508–1515 (2005)

7. Ferhat, O., Vilariño, F.: Low cost eye tracking: the current panorama. Comput. Intell. Neurosci. (2016)
8. Chennamma, O.H.R., Yuan, X.H.: A survey on eye-gaze tracking techniques. Indian J. Comput. Sci. Eng. **4**(5), 388–393 (2013)
9. Fitzgibbon, A.W., Fisher, R.B.: A buyer's guide to conic fitting. In: Proceedings of the 6th British Conference on Machine Vision, vol. 2, pp. 513–522 (1995)