A Comparative Study of Eye Tracking Techniques

R. Suppharachyothin[1,2] J. Nittayasak[2], T. Kondo[2], I. Nilkhamhang[2], Y. Koike[3], K. Tungpimolrut[4], T. Leelasawassuk[4]

[1]TAIST Tokyo Tech, ICTES Program
[2]School of Information, Computer and Communication Technology
Sirindhorn International Institute of Technology, Thailand
[3]Tokyo Institute of Technology, Japan
[4]National Electronics and Computer Technology Center, Thailand
E-mail: rawikarn@siit.tu.ac.th, tkondo@siit.tu.ac.th

Abstract - **This paper presents a comparison between three real-time image processing techniques for an eye tracking system. The techniques tested in this paper are a projection-data based method, the circular Hough transform (CHT), and the template matching method (TM). We reveal the strength and weakness of each technique. The position of the eye and pupil may be detected using both vertical and horizontal projections of image intensities in the projection-data based method. The computation of the method is simple and works well when the contrast of the pupil/iris is high. The CHT is based on the boundary edges of the pupil/iris. Thus, the CHT is well-suited to sharp images. Both the projection-data based and CHT methods can be applied to scale-varying images. In the TM, a target region to track is firstly cropped from a reference frame in an image sequence as a template. Thus, the technique can handle various types of eyes of different colors and shapes. Under the constant scaling condition, the TM outperforms the other two methods.**

Keywords: Eye tracking, pupil detection, projection-data based, circular Hough transform, template matching.

I. INTRODUCTION

Eye tracking finds various applications such as cognitive study, medical research, communication systems for disabled, improved image and video communications, and equipment for ophthalmic diagnosis [1]. We are concerned with tracking the human eye in real time for the purpose of assisting ophthalmologists to perform eye treatment and surgery. The localization of the eye region (esp. the pupil and iris) can be achieved through various image processing techniques. A gray prediction model developed by Tang and Zhang is used to find the eye's position in the subsequent frame [2]. Thresholding can be used under a well-controlled lighting condition [3] – [7]. These techniques are well suited to real-time applications. This paper presents a comparative study result of three common approaches: a projection-data based method [8], the circular Hough transform (CHT) [9] and template matching technique (TM) [10]. In Section II, we will explain those techniques. In Section III, we will present our simulation results on a real video sequence. Finally, we will conclude the paper in Section IV Conclusion.

II. METHODS

In this section, the process of the projection-data based, the CHT, and the TM techniques are explained.

A. Projection-data Based

The projection data based method locates the center of the eye using the average of both vertical and horizontal projections of image intensities, sum the intensity value of each row and column, then divided by the number of pixels in each row and column respectively [8]. A color input image is converted to a gray-scale image first. The image is then down-sampled to the one fifth of the original image. This step is effective for removing small unimportant details such as bright reflection spots as well as reducing the computational cost. We apply Median filtering next to further remove noise. Then vertical and horizontal projections of image intensities are computed. Since the eye region of interest (pupil/iris) is darker than its surrounding tissues, those projections data always show valleys that correspond to the eye (Fig. 1). In this way, the detection of the eye is reduced from 2-D search to the valley detection in two 1-D projection data graphs in this scheme. The process of the projection-data based method is summarized in the flow chart in Fig. 2.

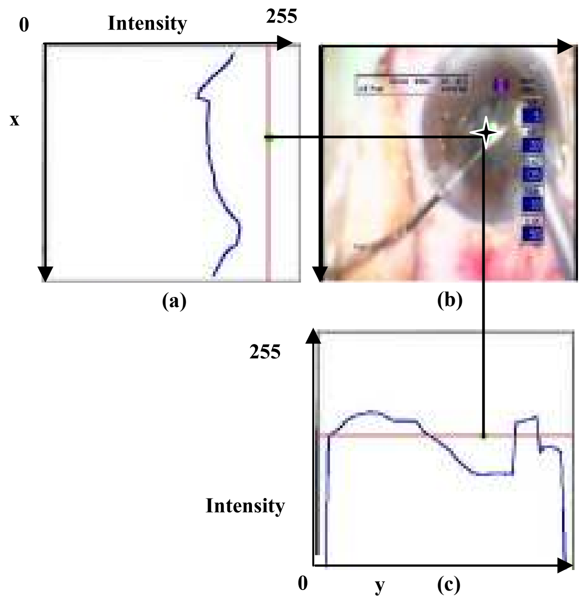

Figure 1. (a) Horizontal projection of the input image (b), and (c) Vertical projection of the input image (b)

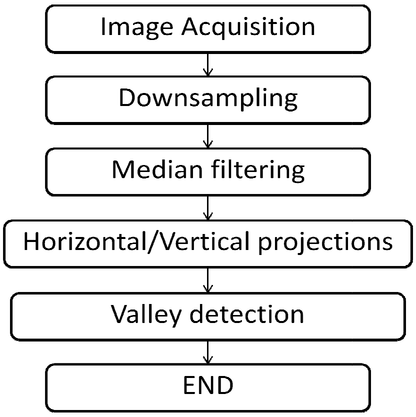

Figure 2. Steps of the projection-data based method

B. Circular Hough Transform

The process of the CHT is based on the boundary edges of the pupil/iris region in the image. Thus the method starts with edge detection. The CHT is then applied to the detected circumference. The procedure of CHT is shown in Fig. 3.

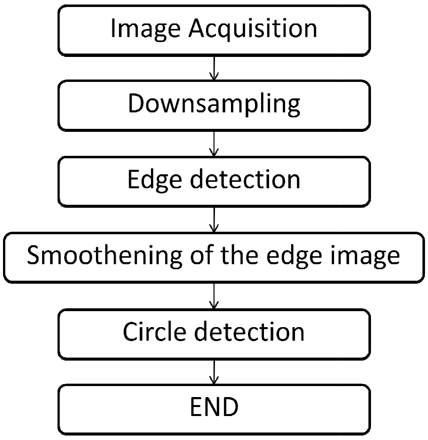

Figure 3. Steps of the circular Hough transform method

First step, downsampling is used to get a sharper image by compressing the resolution to 50% of the original image. The size of the pupil/iris is large, even if the image resolution is reduced, the diameter (or radius) of the pupil/iris can be determined reliably. In addition, downsampling helps to speed up the eye detection step. Edge detection heavily depends on the quality of input images. Thus, it is suited for only high quality image but it can combine with other techniques to get better results [11]. In this study, we extract edge information of the images by using the Canny edge detector. That is a first-derivative based edge detection method [12]. The procedure comprises smoothening an image with a circular 2-D Gaussian filter, computation of the gradient magnitude and direction, nonmaxima suppression, double thresholding, and connectivity analysis. Then, the edge image needs to be smoothened to remove noise and also widen the edges. A large mask can reduce more noise, but if it is too large, the small edge information will be lost. In this research, the image is convolved with a Gaussian mask of size 5x5 pixels. Lastly, the circumference and the center of the pupil/iris can be obtained by using the CHT technique. It finds the best position through voting in the parameter space [13]. In our approach, the minimum and maximum radii of the circle are set to reduce the chance of having false circle detection.

C. Template Matching

Template matching (TM) is widely used in many applications of image and signal processing. It is a technique for determining where the image contains a particular object or region in which we are interested in [10]. The computational cost of this technique grows with the number of pixels in an image [14]. There are many methods to match the template with the image such as sum of absolute differences and cross correlation method. In this research, the cross-correlation method is used. It finds the best matching position by determining the global maximum after vector product is performed as in (1).

$$\boldsymbol{R}(\boldsymbol{x},\boldsymbol{y}) = \boldsymbol{\Sigma}_{\boldsymbol{y}=\boldsymbol{1}}^{\boldsymbol{N}}\boldsymbol{\Sigma}_{\boldsymbol{x}=\boldsymbol{1}}^{\boldsymbol{M}}\boldsymbol{x}',\boldsymbol{y}'[\boldsymbol{T}(\boldsymbol{x}',\boldsymbol{y}')\cdot\boldsymbol{I}(\boldsymbol{x}+\boldsymbol{x}',\boldsymbol{y}+\boldsymbol{y}')] \quad (1)$$

where R store the cross-correlation coefficients, T is template, I is an input image, M and N are the x and y dimension of the image, respectively. The template used in this paper is defined in a reference frame by a user manually. Traditional TM methods employ a fixed size of template.

III. RESULTS AND DISCUSSION

The results of the projection-data based, the circular Hough transform (CHT), and template matching technique (TM) are shown in Figs. 4 - 6, respectively.

We evaluate experimental results by means of the ratio of the estimation error and the diameter of the pupil/iris as shown in (2). The estimation error is a deviation between the manually selected center of the pupil/iris (C_{tx},C_{ty})and that of the estimated by the proposed method (C_{px},C_{py}). The video sequence used in this experiment contains 676 frames. The size of each input image is 320 × 240 pixels. The template is cropped only the area of the pupil/iris in square shape. In order to simplify the computation, we randomly select 100 frames out of 676 frames to evaluate the result.

$$Error(\%) = \frac{\sqrt{(C_{tx}-C_{px})^2+(C_{ty}-C_{py})^2}}{Pupil\ diameter} \times 100 \qquad (2)$$

where C_t and C_p is center of template and center of pupil respectively.

A. Projection-data Based

The projection-data based method can detect the position in every frame but it has the lowest accuracy among these three methods. It also obtains the least success detection. The limitation of this method is its heavy dependence on the quality of input image sequences. In general, the applicable cases of this approach may be limited because it is based on the assumption that the pupil/iris is darker than its surrounding tissues. This may not be always the case, for example, if a patient has cataract or eyes of a light color. However, this method can handle scale-varying images. In Fig. 4, a diamond shaped marker denotes where the center of the pupil is, which is calculated by using the projection-data based method.

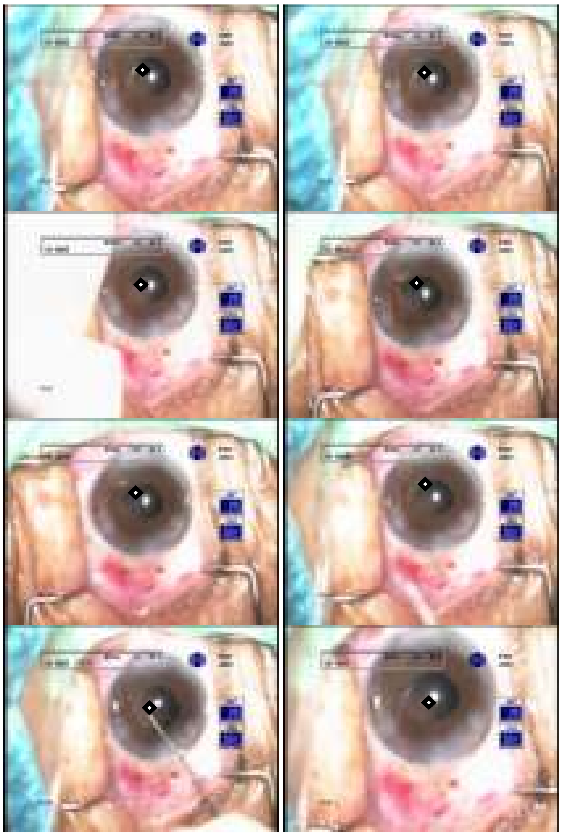

Figure 4. Results of the projection-data based method

B. *Circular Hough Transform*

The circular Hough Transform (CHT) seems to work well in most frames. However, when input images are not very sharp, the CHT fails to detect a circle. In Fig. 5, a dark point indicates the center of the pupil/iris detected by the CHT. The CHT is dependent on the sharpness of input image sequences because the technique is dependent on edge information. When an input image is blurred, the threshold values for edge detection need to be lowered adaptively, which is not always easy to perform. An advantage of the CHT is that it can also handle scale-varying images.

C. *Template Matching*

For template matching (TM), the template image cropped out from the pupil in the reference frame of the video sequences. If the scale of the image is does not vary much, this method works best. Fig. 6 shows the best matching position determined by the TM.

The threshold of the error in each frame is set to determine the accuracy. If the error is less than 10%, it will consider as successful detection. The success rates and mean error of the projection-data based, the CHT, and the TM method is shown in Table1.

Table 1. Comparison of the accuracy

Comparison of the accuracy		
Methods	Mean Error	Success Rates
Projection-data based	9.62%	62%
Circular Hough Transform	6.94%	72%
Template Matching	4.12%	98%

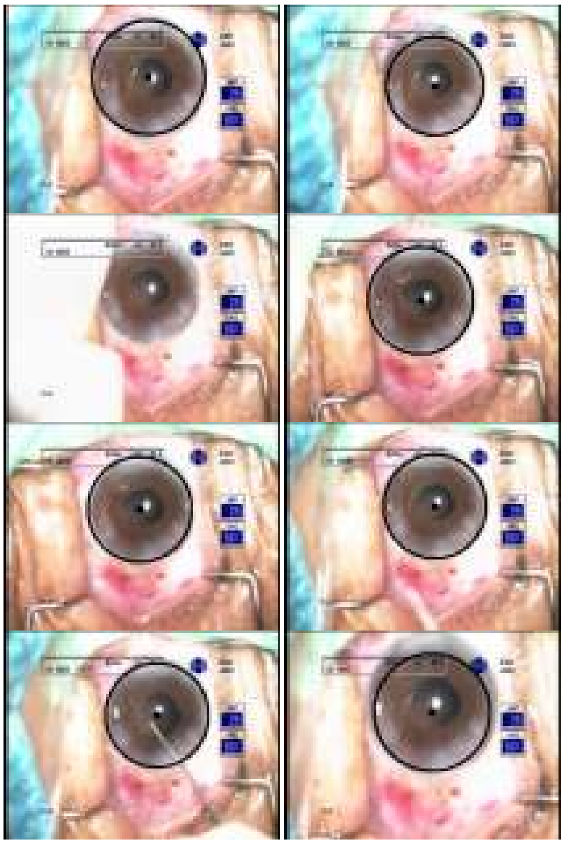

Figure 5. Results of the circular Hough transform method

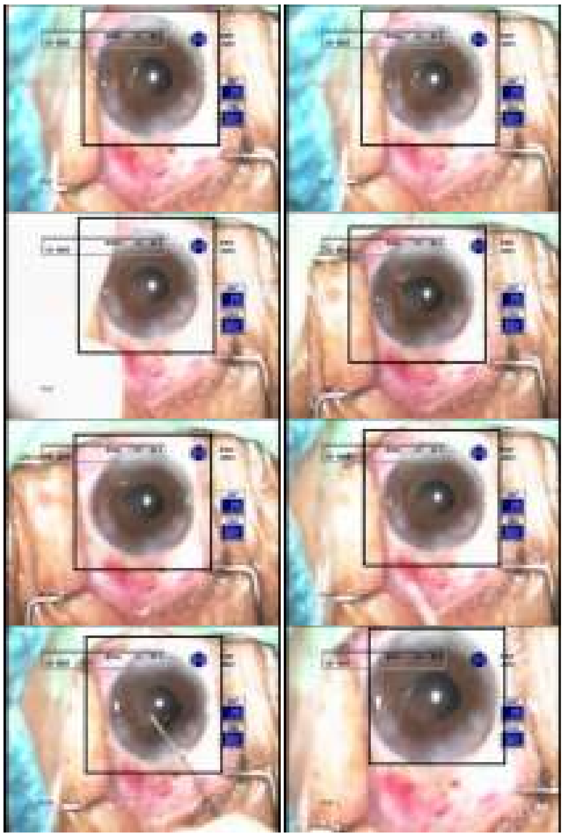

Figure 6. Results of the template matching method

IV. CONCLUSION

This paper describes three real-time image processing methods that can be used for eye tracking. Our simulation results show that the template matching technique is the best among the three, whereas the projection-data based method is the worst. The projection-data based method works well when the pupil/iris region is much darker than its surrounding tissues. This implies that the technique is weak if the eye region is not dark enough, for example, with cataract. In addition, the technique will not be suitable for the patients with eyes of light colors. The circular Hough transform (CHT) works well when input images are on-focus and sharp because the method is based on edge information. This means that the CHT will not suitable for the eyes with some anomaly, such as cataract, that obscures the eye region. Finally, the TM technique works the best in all circumstances, excluding variations in scale. Since our objective is to assist the ophthalmologist to perform surgery on the eye with some anomaly. The TM should be the most suitable technique for the purpose. One limitation of the TM, however, is that the technique is sensitive to variations in scale. Currently, we are working on this problem with some promising results.

ACKNOWLEDGMENT

This research is financially supported by Thailand Advanced Institute of Science and Technology - Tokyo Institute of Technology (TAIST-Tokyo Tech), National Science and Technology Development Agency (NSTDA), Tokyo Institute of Technology (Tokyo Tech), and Sirindhorn International Institute of Technology (SIIT), Thammasat University (TU).

REFERENCES

[1] http://en.wikipedia.org/wiki/Eye_tracking

[2] Jianxiong Tang and Jianxin Zhang, "Eye Tracking Based on Grey Prediction", *1st International Workshop on Education Technology and Computer Science*, pp.861-864, 2009.

[3] K. S. Jin, S. G. Cho, J. S. Lee, and J. J. Hwang, "Real-time Pupil Detection Based on Three-step Hierarchy", *7th International Conference on Signal Processing*, 2004.

[4] J. M. Cho, S. J. Lee, J. K. Kim, H. H. Choi, O. S. Kwon, and J. W. Kwon, "A Pupil Center Detection Algorithms for Partially-covered Eye Image", *TENCON2004*, 2004.

[5] S. I. Kim, J. M. Cho, J. Y. Jung, S. H. Kim, J. H. Lim, T. W. Nam, and J. H. Kim, "A Fast Center of Pupil Detection Algorithm for VOG-based Eye Movement Tracking", *27th IEEE Annual Conference on Engineering in Medicine and Biology Society*, September 2005.

[6] S. Goni, J. Echeto, A. Villanueva, and R. Cabeza, "Robust Algorithm for Pupil-glint Vector Detection in A Video-oculography Eyetracking System", *17th IEEE International Conference in Pattern Recognition*, 2004.

[7] A. Villanueva, R. Cabeza, and S. Porta, "Eye tracking: Pupil Orientation Geometrical Modeling," *Image and Vision Computing*, pp.663-679, 2006.

[8] S. Keerativittayanun, K. Rakjaeng, T. Kondo, W. Kong-prawechnon, K. Tungpimolrut, and T. Leelasawassuk, "Eye Tracking System for Ophthalmic Operating Microscope", *ICROS-SICE International Joint Conference 2009*, pp 653-656.

[9] Klaus Toennies, Frank Behrens, and Melanie Aurnhamer, "Feasibility of Hough-transform-based Iris Localisation for Real-time-application", *16th International Conference on Pattern Recognition*, pp.1053-1056, 2002.

[10] Rafael C. Gonzalez, Richard E. Woods, "Digital Image Processing", Second edition, pp 206.

[11] S. Dey, and D. Samanta, "An Efficient Approach for Pupil Detection in Iris Images," *15th International Conference on Advanced Computing and Communications*.

[12] Paul Bao, Lei Zhang, and Xiaolin Wu, "Canny Edge Detection Enhancement by Scale Multiplication", *IEEE Transactions on Pattern analysis and machine intelligence*, Vol. 27, No. 9, September 2005, pp.1485-1490.

[13] Qi-Chuan Tian, Quan Pan, Yong-Mei Cheng, and Quan-Xue Gao, "Fast Algorithm and Application of Hough Transform in Iris Segmentation."*Proceeding 3th International Conference on Machine Learning and Cybernetics*, Shanghai, pp. 3977-3980, Aug. 2004.

[14] Luke Cole, David Austin, Lance Cole, "Visual Object Recognition using Template Matching", *National ICT Australia*, Locked Bag 8001, Canberra, ACT 2601.