

A Scale-variable Template Matching Technique for Eye Tracking

R. Suppharachyothin¹, J. Nittayasak¹, T. Kondo¹, I. Nilkhamhang¹, Y. Koike², K. Tungpimolrut³, T. Leelasawassuk³

¹School of Information, Computer and Communication Technology
Sirindhorn International Institute of Technology, Thailand

²Tokyo Institute of Technology, Japan

³National Electronics and Computer Technology Center, Thailand

E-mail: rawikarn@siit.tu.ac.th, tkondo@siit.tu.ac.th

Abstract - This paper presents a scale-variable template matching technique for real-time human eye tracking. The proposed technique is a combination of a conventional template matching technique and the circular Hough transform. Firstly, the template matching technique is used for tracking the patient's eye (the pupil and iris). The circular Hough transform is then applied to the best matching region to acquire the size of the eye. When the scale of the eye varies, the size of the template is equally varied. In this way, a scale-variable template matching technique is realized and implemented. Results show that the proposed technique can track a scale-changing eye in real time successfully.

Keywords: Eye tracking, pupil detection, template matching, cross correlation, circular Hough transform.

I. INTRODUCTION

There are many applications that an eye tracking technique can be applied such as cognitive study, medical research, communication systems for disabled, improved image and video communications, and equipment for ophthalmic diagnosis.

There are many existing methods for eye tracking such as eye tracking based on grey prediction [1], video-oculograph [2], the circular Hough transform (CHT) [3], ellipse fitting [4], thresholding [5] – [9], and edge based technique [10], [11]. Tang and Zhang [1] use gray prediction model to predict the position of moving eye in the next frame, and use that position as the reference point for the searching region of eye. Kim et al. [2] proposed pupil/iris detection algorithm for video-oculograph -based eye movement tracking, the center of pupil is hard to detect when it has a reflection and more than half of the pupil/iris is hidden by something. Toennies et al. [3] demonstrated Hough transform-based following the assumption that an average ratio between iris size and distance is known and the distance between iris and camera is limited. T. Funahashi et al. [4] show that ellipse fitting technique can detect an oval object, but not suitable for image of low quality (low contrast, noisy, blurred) as a precise edge information is necessary. Thresholding techniques are a well suited to real-time applications, but sensitive to the change of image intensities. These approaches can be used only under a well-controlled lighting condition [5] – [9]. Edge-based techniques are heavily dependent on the quality of input

image sequences, and thus can be used in combination with other techniques [10], [11].

We are concerned with a real-time human eye tracking system for assisting ophthalmologists who perform medical operations. At the moment, the ophthalmologists we work with need to move a microscope to follow the motion of patient's eye using a foot pedal while performing eye treatment and surgery. A smooth and adequate positioning of the microscope requires a highly sophisticated skill. Consequently, a long-term experience is necessary. The objective of this project is to automate the positioning of the microscope. This paper describes an image processing technique that can detect and track the human eye (both pupil and iris) in real time.

II. SYSTEM OVERVIEW

Fig. 1 shows an ophthalmic operating microscope. The ophthalmologist needs to control the position of the microscope following the patient's eye movement using a foot pedal while performing eye treatment and surgery. It is apparent that this manipulation requires a highly sophisticated skill. The ultimate goal of the project is to realize a fully automatic eye tracking system to assist the ophthalmologist to concentrate on surgical operations. For this purpose, we have been currently developing a real-time image processing technique to locate the position of the eye (pupil and iris). Subsequently, the detected eye position will be sent to a motor that controls the position of the microscope.

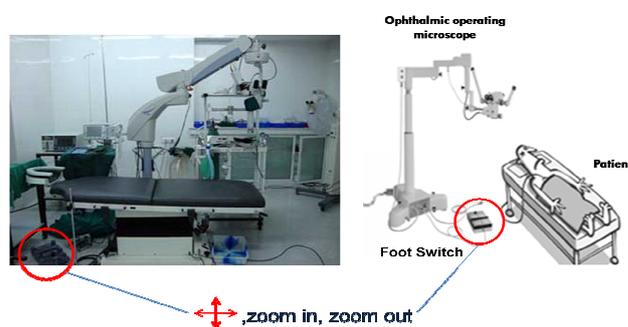


Figure 1. System overview of operation equipment.

III. METHOD

We aim to locate the eye (pupil/iris) in real time using two conventional image processing techniques; the template matching technique (TM) and the circular Hough transform (CHT). By combining these two, a scale-variable template matching technique has been devised.

A. Template Matching

Template matching is widely used in many applications of image and signal processing. It is a technique for determining whether the image contains a particular object or region in which we are interested in [12]. 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 global maximums after vector product is performed as in (1).

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

where R store the matching map, 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 template matching methods employ a fixed size of template. However, in our application, the ophthalmologists move the micro scope up and down to obtain an appropriate visibility of the eye during eye surgery. As a result, the scale of the eye varies, which causes a problem for template matching. Therefore, we monitor the size of the eye every frame and if the scale of the eye is changed, we adjust the size of the template adaptively. In real operation, the ophthalmic operating camera needs to zoom. When the camera zooms in and out, the size of pupil/iris in the video is bigger and smaller respectively. Therefore, the scale-variable template matching method is constructed. The technique for monitoring the size of the eye is described next.

Scale-variable template matching is based on template matching and the size of template can vary according to pupil/iris diameter. It consists of 3 main steps, *first* is find the diameter of pupil/iris, *second* is resize the template to the size of pupil/iris's diameter, and *third* is template matching.

First step, we find the diameters of pupil/iris by using CHT, the details of CHT are explained in next section. Second step, resize the template according to diameter of pupil/iris and set the limit follow (2) to stabilize the view and to avoid track extremely large change.

$$0.7 \times D_O < D_N < 1.3 \times D_O \quad (2)$$

where D_O and D_N is previous and new diameter respectively. If value of new diameter exceeds this range, it becomes the value of diameter of previous frame. Third step, after the template is resized, we do template matching by using cross correlation method to find the best matching position using (1).

B. Circular Hough Transform

The process of the CHT is used to locate the pupil/iris region of the image. This method starts after template matching in the first frame is finished. Then, CHT is applied to the area where the template matching's result is. This method starts from extract each frames from the video file and then downsampling it follows by detecting the edge and smoothing the image. After that, CHT is applied to detect the circumference and the center of the pupil/iris. From this point, the diameter of the pupil/iris is used for determining and resizing the template for current frame. The procedure of CHT is shown in Fig. 2.

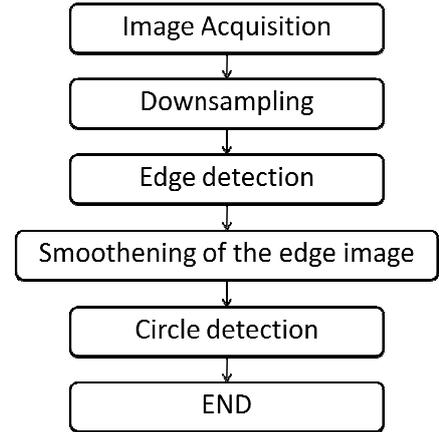


Figure 2. Process of Circular Hough Transform method.

a) Downsampling

Downsampling is use to reduce sampling rate. It is used to get the sharper image by compress the resolution to 50% of original image. Due to the size of the iris is large, when the image resolution is reduced, it will not be affected to the process to find the diameter of the iris. This step also helps to speed up the eye detection step.

b) Edge detection

Edge detection is very useful to extract the edge information of the images but it is heavily dependent on the quality of input image sequences. It is suited for only high quality image but it can combine with other techniques to get better result [10]. In this study, we extract edge information of the images by using the canny edge detector. The canny edge detection is a first-derivative based edge detection method [13]. The procedure is smooth an image with a circular 2-D Gaussian filter, compute the gradient magnitude and direction, nonmaxima suppression to the gradient magnitude, double thresholding and connectivity analysis respectively.

c) Smooth the image

The image needs to smoothen to remove the noise and also wider the edge. Size of mask is very important, big 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.

d) *Circle detection*

The circumference and the center of the pupil/iris are detected in this step by using CHT technique follow (3).

$$(x - c_1)^2 + (y - c_2)^2 = c_3^2 \tag{3}$$

which c_1, c_2, c_3 , determine a 3-D parameter space where a circle is detected by a voting process. It finds the best position through voting in parameter space [14]. The procedure is to increment c_1 and c_2 to solve for c_3 that satisfies (3), and update the accumulator corresponding to the cell associated with these three parameters. Moreover, the minimum and maximum radius of circle is set to reduce the false detection.

IV. RESULTS AND DISCUSSION

A. *Circular Hough Transform*

The circles in Fig. 4 are circumferences of pupil/iris and the point is the center of that circle. The CHT cannot detect every frame due to changing lighting conditions and image quality. If an image is too blurred, a fixed threshold will fail to detect edges that are necessary for the CHT. Therefore, some frames cannot find the suitable circular edge.

B. *Template Matching*

Pupil/iris area cropped from the first frame is used as the template for this video sequence, shown in Fig. 3. In this method, the size of template is fixed. In case the size of pupil/iris from video sequences is varied too much, the result will cause an error. Most existing object recognition systems suffer from poor scaling with the number of recognizable objects [15].

C. *Scale-Variable Template Matching*

The square is also the best matching position and the size of template can vary according to the pupil/iris's diameter which can be derived from the CHT. The result is shown in Fig. 4.

We evaluate experimental results by means of the ratio of the estimation error and the diameter of the pupil/iris as shown in (4). The estimation error is a deviation between the true 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 is in the length of 500 frames. The size of each input image is 320×240 pixels. The template is cropped only the area of pupil/iris in square shape. In order to simplify the computation, we randomly select 200 frames out of 500 frames in a video sequence to evaluate the result.

$$Error(\%) = \frac{\sqrt{(C_{tx}-C_{px})^2+(C_{ty}-C_{py})^2}}{Pupil\ diameter} \times 100 \tag{4}$$

where C_t and C_p is the center of the template and the center of the pupil/iris respectively.



Figure 3. Pupil and iris are used as a template.

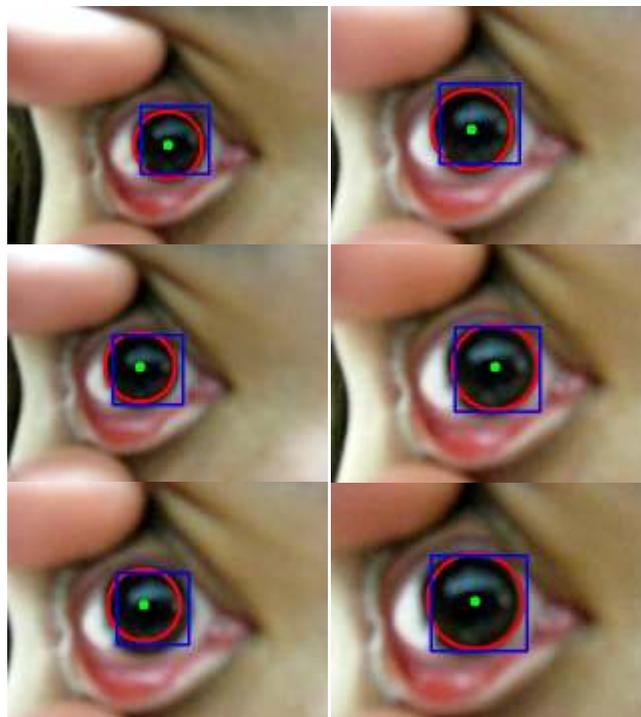


Figure 4. Results of scale-variable template matching method.

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, mean error and standard deviation of the error is shown in Table1.

Table1. Mean error and standard deviation.

Success Rates	83%
Mean Error	8.28%
Standard Deviation	9.37%

V. CONCLUSION

This paper proposes a novel approach for eye tracking. The proposed technique is based on a template matching technique as it is well-suited to the detection of a fixed pattern from a frame to another. One limitation of the template matching technique is that it cannot detect a scale-varying object because of the fixed size of the template. For this, we introduced the CHT to obtain the size of the pupil/iris. The CHT is applied only to a limited small region determined by the proceeding template matching technique. In this way, the

computational cost of the CHT is reduced, and more importantly, false circle detection may be avoided. With the aid of the CHT, we can monitor the size of the eye. Therefore, it is possible to adjust the size of the template immediately when a scale change occurs.

Experimental results show a satisfactory performance when the size of the pupil/iris or iris varies. As future work, we plan to further evaluate the performance of the proposed technique on more video sequences. Subsequently, we will aim to develop an eye-tracking method that is robust to varying sharpness of images. Eventually, we will implement a motor control part of the microscope so that it follows patient's eye automatically in real time.

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] Jianxiong Tang and Jianxin Zhang, "Eye Tracking Based on Grey Prediction", *1st International Workshop on Education Technology and Computer Science*, pp.861-864, 2009.
- [2] S. I. Kim, J. M. Cho, J. Y. Jung, S. H. Kim, J. H. Lim, T. W. Nam, J. H. Kim, "A Fast Center of Pupil Detection Algorithm for VOG-based Eye Movement Tracking", *IEEE Engineering in Medicine and Biology 27th Annual Conference*, pp.3188-3191, September 2005.
- [3] 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.
- [4] J. Zhu and J. Yang, "Subpixel Eye Gaze Tracking", *5th IEEE International Conference on Automatic Face and Gesture Recognition*, 2002.
- [5] 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.
- [6] 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.
- [7] 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.
- [8] 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.
- [9] A. Villanueva, R. Cabeza, and S. Porta, "Eye tracking: Pupil Orientation Geometrical Modeling," *Image and Vision Computing*, pp.663-679, 2006.
- [10] S. Dey and D. Samanta, "An Efficient Approach for Pupil Detection in Iris Images", *15th International Conference in Advanced Computing and Communications*, 2007.
- [11] A. Discant and S. Demea, "An Algorithm for Extraction of Iris Information", *International Conference in Automation, Quality, and Testing, Robotics*, vol. 2, pp. 445-450, May 2006.
- [12] Rafael C. Gonzalez, Richard E. Woods, "Digital Image Processing", Second edition, pp 206.
- [13] 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, pp.1485-1490, September 2005.
- [14] Qi-Chuan Tian, Quan Pan, Yong-Mei Cheng, and Quan-Xue Gao, "Fast Algorithm and Application of Hough Transform in Iris Segmentation.", *3th International Conference on Machine Learning and Cybernetics*, Shanghai, pp. 3977-3980, August 2004.
- [15] Ardeshir Goshtasby, "Template Matching in Rotated Images", *IEEE Transactions on Pattern analysis and machine intelligence*, Vol. PAMI-7, No. 3, pp. 338-344, May 1985.