USING STRUCTURED ILLUMINATION TO ENHANCE VIDEO-BASED EYE TRACKING

Feng Li[†]*, Susan Kolakowski, Jeff Pelz*[‡]

Chester F. Carlson Center for Imaging Science, Rochester Institute of Technology,
54 Lomb Memorial Drive, Rochester, NY 14623, USA
[†] lxf7267@cis.rit.edu [‡] pelz@cis.rit.edu

ABSTRACT

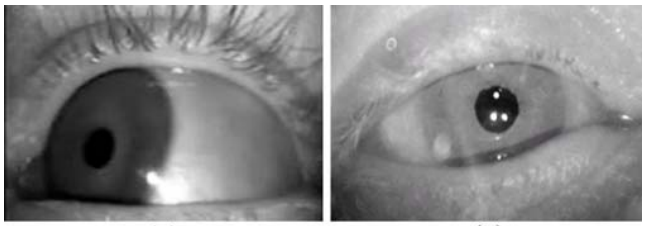
Video-based eye tracking commonly relies on the positional difference between the pupil center and the *first-surface corneal reflection of a light source (CR)* to map the line of sight. In this paper, a new structured illumination configuration, utilizing an array of IREDs to illuminate the eye, is proposed to address two common problems in current systems: 1) loss of CR during extreme eye movements, and 2) specular reflection artifacts being mistaken for the desired CR. The design provides reliable difference vectors even during extreme eye movements. The challenge of spurious specular reflections is solved by a two-stage processing approach. First, potential CRs and the pupil are isolated based on statistical information in an eye image. Second, genuine CRs are distinguished by a novel CR location prediction technique. A prototype implementation in the RIT Lightweight Eye Tracker [1] is described.

Index Terms — Video-based eye tracking, pupil, corneal reflection, eye movement

1. INTRODUCTION

Eye tracking techniques have been actively studied since the nineteenth century. The diverse demands of researchers have spawned many different techniques, such as electro-oculography, scleral search coils, dual-Purkinje trackers and video-based methods [2]. Because of its minimal obtrusiveness to observers, relatively easy set-up and reliance on rapidly developing optical and electronic imaging devices, video-based eye tracking has become one of the most popular and successful eye-tracking techniques.

In existing laboratory and commercial video-based eye trackers, track losses caused by extreme eye movements and false detection of the CR due to uncontrolled ambient illumination are significant problems. When the eye rotates to an extreme position, the CR may 'roll off' the spherical region of the cornea (Figure 1a) or appear on the eye where it has a low contrast with its surroundings. In this situation, CR detection becomes a non-trivial task and track loss often occurs. As such, video-based systems usually restrain measurement ranges to eye movements within approximately ±25° horizontally and ±20° vertically. Another challenge may occur when more than one specular reflection is a geometrically satisfactory CR (Figure 1b).



(a) (b)

Figure 1. Challenging eye images. (a) A CR rolls over to the aspherical part of the cornea; (b) Spurious specular reflections from metal frames in a lift truck (the bright spot below and to the left of the pupil center is the desired CR). Image (b) courtesy of D. Giguère at the IRSST - Safety-Ergonomics Research Program, Canada.

Under this circumstance, the bright spot closest to the pupil may be taken as the CR (e.g., ASL 6000 series [3]); this makes the system run smoothly but inevitably introduces error.

What we term "*structured illumination*" utilizes an array of infrared emitting diodes (IREDs) to produce multiple CRs[*]. This configuration ensures that at least one CR is trackable – reflected from the spherical surface of the cornea and superimposed on either the pupil or iris such that it has a high contrast with its surroundings – even with extremely eccentric gazes. One approach for using multiple illuminators in video-based eye tracking has been reported in [4, 5]: two sets of IREDs are mounted on-axis and off-axis with the eye camera and flash alternately to produce bright-pupil and dark-pupil images, and their difference image is used to improve pupil detection. This use of multiple sources is different from the Structured Illumination configuration in the sense that it does not aim to solve the problem of track loss or limited measurement range. Another approach, using a symmetric arrangement of four IREDs around the camera's optical axis to extend the tracking range, has been proposed recently [6]; in this approach, at least three out of four CRs need to be detectable for effective tracking. In this paper, the challenge of spurious specular reflections from external sources or scleral reflections is solved by a two-stage processing approach. First, global and local statistical information from an eye image are used to isolate and determine the pupil center and potential CRs. Second, genuine CRs are

[*] This Structured Illumination design was first presented at the OSA Rochester (NY) Section Meeting in January 2006.

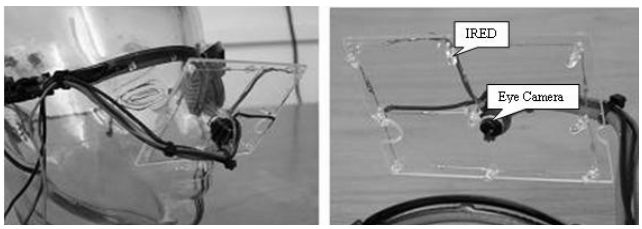


Figure 2. Prototype of the Structured Illumination

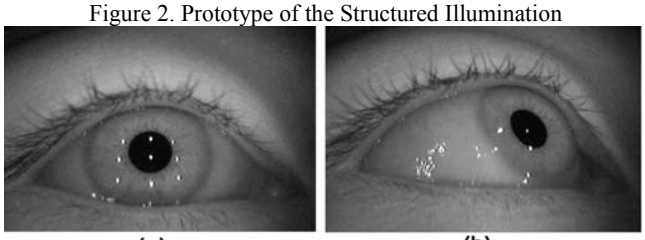


Figure 3. Eye images illuminated by the Structured Illumination. (a) The eye looking straight ahead; (b) the eye moved to an extreme position.

separated from false CRs by a novel physiological optics based technique – CR location prediction via a well-correlated pupil-CR offset ratio.

This paper first introduces the prototype implementation of the Structured Illumination configuration in the RIT Lightweight Eye Tracker (Section 2). Detection of potential CRs and the pupil using statistical information of an eye image are described in Sections 3 and 4, respectively. The ratio between the offset of the pupil center and the offset of the CR, and its application in discriminating genuine CRs are presented in Section 5. Finally, Section 6 consists of conclusions and possible future extensions of this work.

2. STRUCTURED ILLUMINATION

The new illumination configuration uses nine IREDs in a 3×3 array as opposed to many eye tracking systems that only use one. The illumination sources are mounted equally spaced on a transparent supporting frame with an eye camera placed to the lower-right of the center IRED (Figure 2).

Figure 3 shows two example images with multiple CRs from the new design. Note that reflections from some of the illuminators are visible even during extreme eye movements (Figure 3b). However, many artifacts are observed in the eye images (actually more artifacts may be seen in daylight) due to spurious reflections from multiple IREDs and/or extraneous ambient illumination. The aspherical nature of the cornea-scleral boundary is another source of irregular reflections.

In our initial testing, at least three CRs could be detected in each video frame when recording eye movements up to ±40° horizontally and ± 25° vertically.

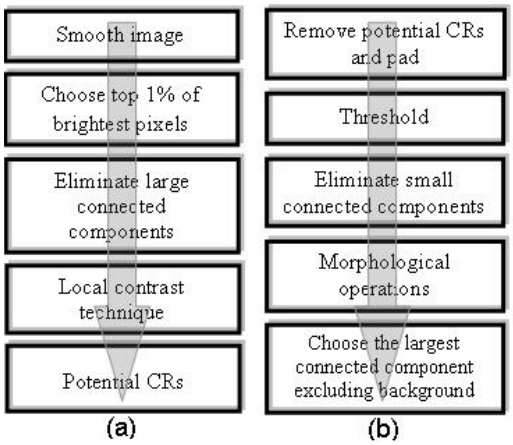


Figure 4. Flowcharts of CR (a) and pupil (b) detection

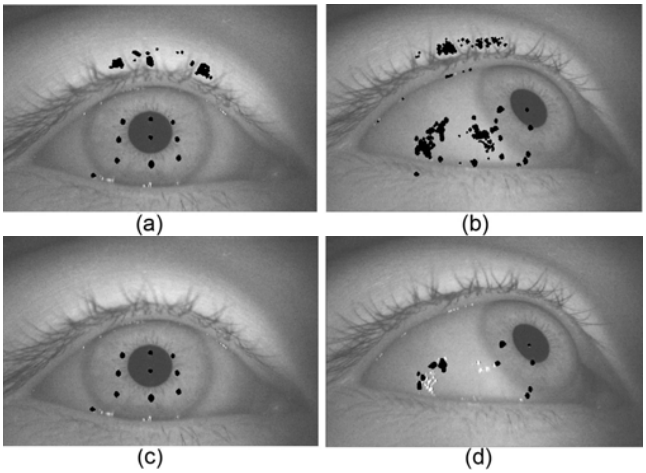


Figure 5. Eye images in CR detection. (a)(b) Connected components with pixel values in top 1% after eliminating components larger than 0.15% of image size; (c)(d) Potential CRs after the local contrast technique. Superimposed original eye images are washed out for display purposes.

3. POTENTIAL CR DETECTION

Detection of potential CRs consists of several steps (flowchart in Figure 4a). An eye image is first smoothed by a 5×5 Gaussian filter with a standard deviation of 2 pixels to reduce noise effects. The top 1% of the brightest pixels in the image is then selected; CR pixels are among these brightest pixels. After region-labeling, connected components larger than 0.15% of the image size are eliminated (Figure 5a and b). Additional spurious specular reflections are removed by a local contrast technique, which only accepts potential CRs that have high contrast with a small set of their neighboring pixels. In other words, for each connected component (with pixel values in the top 1% of pixel values in the image), the grayscale difference between its mean and that of its surroundings must be larger than a threshold (50 in our implementation) to be selected as

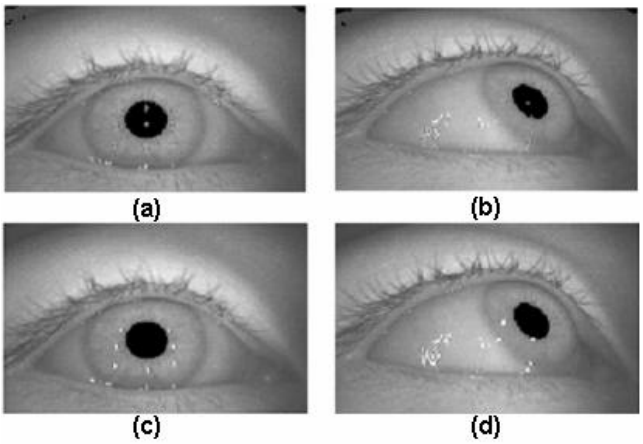


Figure 6. Eye images in pupil detection. (a)(b) Images thresholded with the grayscale of the first valley in the histogram; connected components smaller than 1% of image size are eliminated; (c)(d) Highlighted pupil regions (black ellipses). Superimposed original eye images are washed out for display purposes.

a *potential CR* (Figure 5c and d). The surroundings are defined as eight neighboring rectangular regions, each with width and height equal to the width and height of the connected component.

4. PUPIL DETECTION

In order to detect the pupil in an eye image, the potential CRs are first removed and their remaining holes are padded with the mean gray values of their eight rectangular surrounding areas. The basic steps for detecting the pupil are shown in the flowchart in Figure 4b.

The eye images after CR removal are blurred by a Gaussian filter with a large standard deviation of 20 pixels. This step aims to smooth sharp edges caused by padding CR holes and to reduce the influence of small occlusions (such as those due to eyelashes). The first valley in the histogram of the blurred image represents the pupil grayscale limit; a threshold operation is thus applied to select the pupil pixels as those whose values fall below this grayscale value. Connected components smaller than 1% of the image size are eliminated (Figure 6a and b). Note that some artifacts in the upper corners and holes inside the pupil area can still be seen. Morphological closing and filling operations are then used to join boundaries and remove holes inside all connected components. The pupil detection process is finalized by choosing the largest connected component in the binary image, excluding the background (Figure 6c and d).

5. CR PREDICTION

A functional relationship between CR displacements and pupil displacements during an eye rotational movement and its application in discriminating genuine CRs are presented in this section. Some parameters used in the following

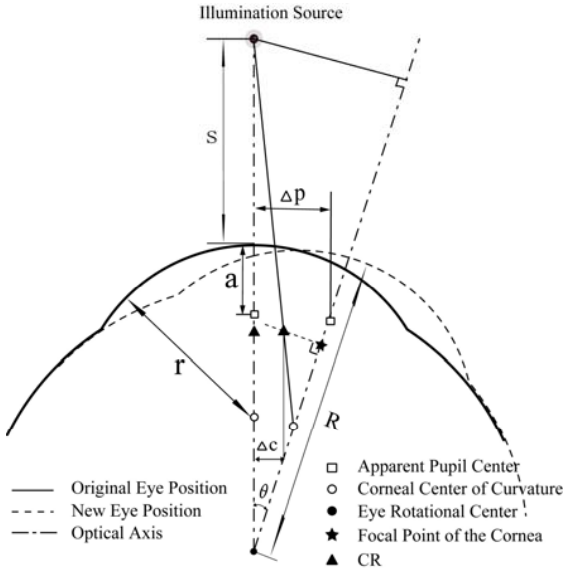


Figure 7. Eye before (solid) and after (dashed) an eye movement

calculations are from Gullstrand's Simplified Eye [7]: the apparent pupil center sits a=3.05mm behind the corneal vertex and the radius of the cornea is r=7.8mm. The rotational center of the eye is not provided in this model, but is reported in Cornsweet and Crane's paper [8] as R=13.5mm from the corneal vertex.

The relationship between CR displacements and pupil displacements can be applied to an illumination source any distance away from the eye given two assumptions:

(1) The CR is located in the focal plane of the cornea;

(2) Small rotational eye movements (θ within ±10˚).

The first assumption is only true for a distant illumination source. However, according to the setup of the Structured Illumination configuration or a similar source mounted about 30-40mm away from the eye, this approximation is reasonable. Based on this assumption, a CR is estimated as the intersection point of the ray passing through the corneal center of curvature with the corneal focal plane. The second assumption is used to simplify the final equation describing the relationship, introducing an error of 1% or less if θ is 10˚ or less [9].

The ratio between the offset of the CR and that of the pupil is given in the following equation (accompanying diagram in Figure 7, detailed derivation in [10]).

$$\frac{\Delta c}{\Delta p} = \frac{R - r/2}{R - a} - \frac{r(s+R)}{2(R-a)(s+r)} \approx 0.5 \qquad (1)$$

where s, Δc and Δp are the distance of the illuminator to the cornea, the offset of the CR and the offset of the pupil center, respectively, in millimeters. When s is equal to 30mm and 40mm (approximate limits of s for the Structured Illumination design), the ratio is equal to 0.49 and 0.51, respectively. These values match the experimental results in [11]. As such, when the eye rotates, the CR moves only about half the distance as does the pupil.

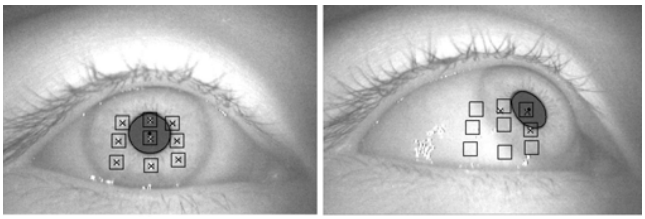


Figure 8. Genuine CR discrimination. Squares represent predicted CR regions, crosses represent detected genuine CRs, ellipses represent the pupil boundaries, and dots represent the pupil centroids. Superimposed original eye images are washed out for display purposes.

After rearranging and expanding Equation 1, we obtain:

$$CR_{current} - CR_{previous} = \Delta c \approx 0.5\Delta p \quad (2)$$

$$CR_{current} \approx 0.5\Delta p + CR_{previous} \quad (3)$$

This last equation shows that a CR position can be predicted if the pupil offset between the current and previous frame and the CR position in the previous frame are known.

In our implementation, a user selects nine CR positions in the first frame. Nine predicted CR locations for each following frame are then calculated using Equation 3. When a potential CR falls inside a predefined square (length of side = 50 pixels) centered on a predicted CR location, it is identified as a genuine CR. This square region tolerates a small amount of translational eye movement (e.g., eye camera movement with respect to the head) and individual variations between observers, such that the genuine CRs will still fit within these regions. If a potential CR does not fall inside this square, it is considered to be an artifact and excluded from the final genuine CR list (Figure 8). The circumferences and centers of the pupil and CRs are calculated by a direct least-squares ellipse fitting technique introduced in [12].

The CR prediction technique robustly distinguishes genuine CRs even when there are spurious specular reflections in the eye image. When a CR - such as the one in Figure 3b - falls close to the limbus (boundary between the iris and sclera), the prediction for this CR may fail. Another limitation occurs when the brightness of a CR, due to motion smear during a rapid eye movement, decreases below a significant level (upper-right CR in Figure 9). However, in both cases, some of the CRs can still be identified successfully.

6. CONCLUSIONS AND FUTURE WORK

Combining the Structured Illumination with the novel CR prediction technique, the new eye tracking system prototype is capable of measuring a wide range of eye movements and effectively locating the pupil and CRs. The noise level of a computed CR may be reduced dramatically by averaging or median filtering all trackable CR locations. In the current approach, a spherical corneal model is used in calculating the pupil-CR offset ratio. The prediction routine could be

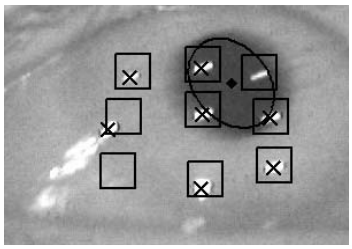


Figure 9. CR prediction for a blurred eye image (due to rapid eye movement). Note that the upper-right CR has not been detected. Superimposed original eye image (cropped) is washed out for display purposes.

improved by using an elliptical model. This would generate dynamic ratios according to eye position instead of one fixed value (~0.5) as in the current implementation. Additionally, a more detailed relationship between pupil and CR positions may be used (instead of the difference vector) to better account for movements of the eye camera with respect to the head [11].

7. REFERENCES

[1] Babcock, J. and J. Pelz. *Building a lightweight eyetracking headgear*. in Proceedings of the symposium on eye tracking research & applications. San Antonio, Texas: ACM Press. 2004.

[2] Duchowski, A.T., *Eye Tracking Methodology: Theory and Practice.*: Springer. 2002.

[3] ASL, *Eye Tracking System Instructions, ASL Eye-Trac 6000, Head Mounted Optics.*, Applied Science Laboratories. 2005.

[4] Morimoto, C.H., et al., *Pupil detection and tracking using multiple light sources.* Image and Vision Computing. 18(4): p. 331-335 2000.

[5] Haro, A., M. Flickner, and I. Essa. *Detecting and tracking eyes by using their physiological properties, dynamics, and appearance*. in Computer Vision and Pattern Recognition. 2000.

[6] Hua, H., P. Krishnaswamy, and J.P. Rolland, *Video-based eyetracking methods and algorithms in head-mounted displays.* Optics Express. 14(10): p. 4328-4350 2006.

[7] Smith, G. and D.A. Atchison, *Eye and Visual Optical Instruments* Cambridge University Press, page 788. 1997.

[8] Cornsweet, T.N. and H.D. Crane, *Accurate two-dimensional eye tracker using first and fourth Purkinje images.* J Opt Soc Am. 63(8): p. 921-8. 1973.

[9] Pedrotti, L.S. and F.L. Pedrotti, *Optics and Vision*: Prentice-Hall, Inc. p. 51. 1998.

[10] Li, F., S. Kolakowski, and J. Pelz, *A model-based approach to video eye tracking.* submitted to Journal of Modern Optics. 2007.

[11] Kolakowski, S. and J. Pelz. *Compensating for eye tracker camera movement*. in Proceedings of the symposium on eye tracking research & applications. San Diego, California: ACM Press. 2006.

[12] Fitzgibbon, A., M. Pilu, and R.B. Fisher, *Direct least square fitting of ellipses.* Pattern Analysis and Machine Intelligence, IEEE Transactions on. 21(5): p. 476-480. 1999.