Single-Pixel Eye Tracking via Patterned Contact Lenses: Design and Evaluation in HCI Domain

Ioannis Rigas
Texas State University
601 University Drive
San Marcos, TX 78666 USA
rigas@txstate.edu

Oleg Komogortsev
Texas State University
601 University Drive
San Marcos, TX 78666 USA
ok11@txstate.edu

CHI'15 Extended Abstracts, Apr 18-23, 2015, Seoul, Republic of Korea
ACM 978-1-4503-3146-3/15/04.
http://dx.doi.org/10.1145/2702613.2732745

Abstract
This paper presents a preliminary study of an eye tracking technique suitable for use in devices with low-power consumption demands, e.g. Google Glass. The method uses a patterned contact lens and a single-pixel imaging sensor. Its applicability is explored via a semi-simulated user study, where real eye movements from 50 subjects are used to animate a 3-D graphics replica of an eye wearing a patterned contact lens. An accurate single-pixel camera simulator is used to perform gaze estimation via capturing of the imprinted pattern. The results show the promising potential of the technique in the field of eye tracking and eye gesture recognition.

Author Keywords
Eye tracking techniques; patterned contact lenses; single-pixel imaging; eye gesture interfaces;

ACM Classification Keywords
H.5.2. [Information Interfaces and Presentation]: User Interfaces-Miscellaneous.

Introduction
The use of eye tracking interfaces in computing machinery is a topic of vigorous research, given the unique advantages provided by the use of gaze data

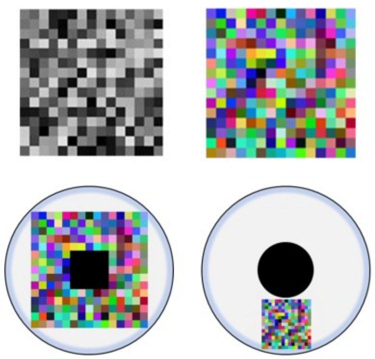

Figure 1. Different patterns and designs for the creation of a PCL.

both in cases where mouse/keyboard input is infeasible (e.g. due to physical disabilities), and as a fast pointing method [5]. In [1], gaze was explored as a mechanism for the smart management of graphic interfaces. Furthermore, several studies explored eye tracking for the dynamic interaction with screens, and the selection of screen items, e.g. [3], [12]. The progress of network multimedia communications induced some research efforts on gaze-based interfaces in that domain too, as demonstrated in [14]. The novel scheme presented in [15], employed eye movements in a hybrid technique that combined gaze and manual input. The technique, named Manual and Gaze Input Cascaded (MAGIC) pointing, showed better behavior in terms of accuracy, speed, and ease of use, compared to methods that use gaze alone. MAGIC pointing technique inspired other studies too, involving various devices and application environments, e.g. [11], [2]. Finally, it should be noted a growing research interest in the development of HCI interfaces driven by eye movement gestures [13], [9].

Eye Tracking Techniques

The research of eye movements, either as the means for understanding human behavior or as an HCI related tool, spurred the development of various eye tracking techniques. The characteristics that differentiate these techniques include, among others, the type of captured signal (optical, electric field), the technical complexity and size of the device, the invasiveness etc. In general, most eye tracking techniques can be classified into one of three classes: a) electrooculography (EOG) methods, where the recording is accomplished via measurements of changes in the electric potential induced by the eye movements. b) Methods based on eye attachments, where the recording is accomplished via an attachment to the eye surface, typically in the form of a customized contact lens. c) Optical tracking methods, where an optical capturing device (e.g. an image sensor) is used for recording eye images, usually in infrared spectrum. A more detailed overview of the characteristics of most common eye tracking techniques can be found in [6].

Motivation of the Proposed Technique

In this work, we introduce an eye tracking technique that can be used for the development of interaction interfaces in cases where reduced power consumption is of utmost importance. Such environments can be encountered in specialized applications (e.g. military and space), and also in wearable devices such as Google Glass. The technique is based on the following novel rationale: instead of employing images of pupil and corneal reflection, we propose a process where an image sensor captures pattern printed on a contact lens. This unique approach makes feasible the task of performing eye tracking by using only a single-pixel capturing element. In this case, power consumption of the pixel sensor can be reduced over 90% compared to other optical eye tracking methods, even those that employ low resolution images, e.g. (4x12 pix.) [7], (11x13 pix.) [10]. Also, since the method is void of searching and fitting for pupil/corneal reflection in an image, the energy footprint is reduced even further.

Description of the Eye Tracking Technique

The Patterned Contact Lens (PCL)

The PCL is a contact lens with a printed pattern, e.g. a rectangle checkerboard. We assume that the PCL is a passive contact lens operating in the visible spectrum, but the technique can be generalized for active PCLs, or infrared spectrum PCLs. In Figure 1 (top), we present some examples of patterns that can be printed on top of a PCL, e.g. a gray-scale checkerboard, or a more

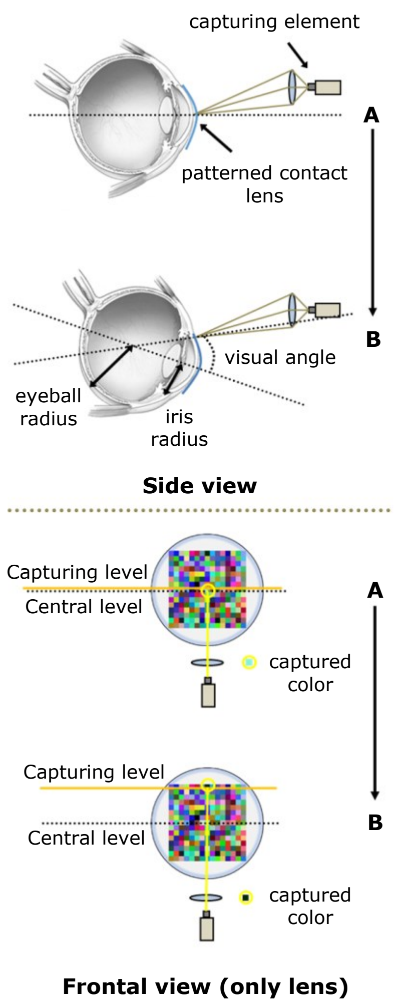

Figure 2. Eye tracking setting of the proposed technique.

robust color checkerboard (used in our experiments). Also, Figure 1 (bottom) shows two possible designs for embedding the pattern on a PCL. The contact lens can be covered either fully or partially. The first case (fully) requires a hole in the place of the eye pupil, whereas the second case (partially) is void of such an occlusion.

Analysis of the Eye Tracking Setting

The basic setting of the technique (Figure 2) consists of a PCL on the subject's eye, and a single-pixel capturing element, i.e. sensor and optics. Using Figure 2, we can proposed technique. Let us assume an eye with a PCL moving from Position A (primary eye position - looking straight forward) to Position B (eye looking downward). The single-pixel light-capturing element, in affixed position, captures a passing sequence of colors from the imprinted pattern. In order to translate the captured color sequence to the corresponding visual angles of eye movement, we need to obtain a mapping between these entities. Although the color values on the pattern are known by design, in practice, the lighting conditions during capturing can affect these values. Consequently, a color (or pattern) calibration procedure is required in order to have valid estimates of gaze. During this procedure, the real color value from every color 'bin' is captured and stored, resulting in a color calibration map that is later employed during gaze estimation. If the lighting conditions are relatively stable, calibration process needs to occur only once.

Detection Range and Precision

Using the eye tracking setting geometry (Side view), we can calculate the theoretical limitations of the technique in terms of detection range and precision. Calculations can be simplified by making the following assumptions: 1) position of the capturing element is fixed with regard to the user's head, 2) the lighting remains relatively steady, and 3) curvature of the contact lens does not distort the pattern. It should be emphasized that during our simulations we employed a 3-D model for the eye globe, therefore the effect of curvature is incorporated into the reported results.

Let us assume that the single-pixel sensor is positioned to capture the central pixel of the pattern at primary eye position. In the first PCL design (Figure 1 bottom-left), the pattern will cover approximately the whole iris. Using Figure 2, and by employing average values for the iris radius = 7 mm and eyeball radius = 12 mm [8], we can calculate the maximum theoretical range of eye movement (degrees of visual angle) that can be detected in each separate direction, as:

$$\text{maxAngle}_1 \cong \pm\text{atan}\left(\frac{\text{iris radius}}{\text{eyeball radius}}\right) \cong \pm 30^\circ$$

For the second PCL design (Figure 1 bottom-right), the pattern covers roughly one third of the iris diameter. In this case we have:

$$\text{maxAngle}_2 \cong \pm\text{atan}\left(\frac{\text{iris radius}/3}{\text{eyeball radius}}\right) \cong \pm 9^\circ$$

Note that although the first PCL design allows a larger traceable range, it requires a hole for the pupil, which disrupts eye tracking for movements performed near the primary eye position. Thus, we decided to adopt the second PCL design for our experiments. In this case the traceable range is narrower but still acceptable for a variety of applications. An example is eye movement gesture recognition in narrow display area devices such as the Google Glass, were the maximum viewing field is approximately 14°.

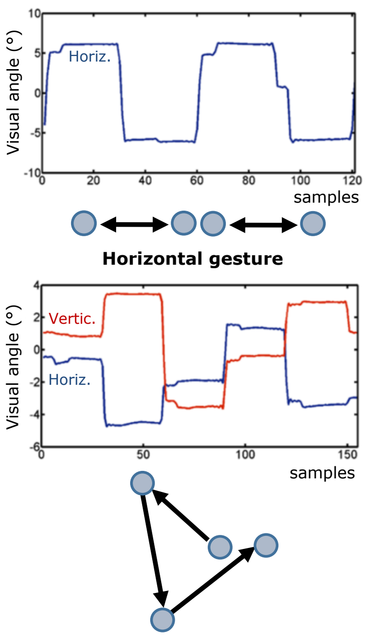

Figure 3. Examples of recorded signals and the eye gestures they represent.

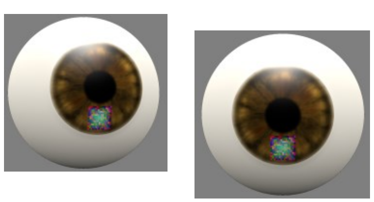

Figure 4. Different views of the 3-D graphics replica, created for the simulation of an eye with a PCL.

The finite number of color 'bins' on a PCL implies certain limitations in the eye tracking *precision*, i.e. minimum amount of detected eye rotation change. The rectangular pattern used in our work has 15 x 15 'bins' (225 levels), corresponding to a 'bin' size of 1 mm^2 for the fully covered PCL, and 0.09 mm^2 for the partially covered PCL. The totally traceable range in a direction (horizontal/vertical) is twice the maxAngle. Thus, the theoretical *precision* due to color level quantization for the fully covered PCL design can be calculated as:

$$\text{precision} = \frac{2^{*}\text{maxAngle}_1}{\text{levels in a direction}} \cong \frac{60°}{15} = 4°$$

For the case of the partially covered PCL design (used in our experiments), we can calculate:

$$\text{precision} \frac{2^{*}\text{maxAngle}_2}{\text{levels in a direction}} \cong \frac{18°}{15} = 1.2°$$

It is noteworthy that the narrower viewing field traced by the partially covered PCL design results in theoretically better precision, since the same number of 'bins' is used to cover a smaller range of visual movement. In this case, though, the employed optics must be able to focus on a smaller area on the PCL.

Experimental Setup

We performed a semi-simulated user study for the evaluation of the proposed technique. Initially, real eye movement data were recorded with a reference eye tracker from a total of 50 subjects (16 male/34 female), ages 18-44 (M = 22, STD = 4.96). The used device was an *EyeLink 1000* eye tracker (sampling rate 1000 Hz, spatial accuracy 0.5°). The recordings were captured with an average calibration accuracy of 0.49° (STD = 0.18°) and an average data validity of 97.1% (STD = 3.96%). Visual stimulus was presented on a computer screen at a distance of 550 mm from the subject's head (dimensions 474 x 297 mm, resolution 1680 x 1050 pix.). The subject's head was stabilized with the use of a chin-rest to ensure high data quality.

Each subject performed three eye movement tasks: horizontal saccades, random oblique saccades, and reading text. From each recording we extracted a signal of 10 seconds and sub-sampled it from 1000 Hz to 30 Hz, in order to represent an eye gesture, i.e. a horizontal gesture, a complex path gesture, and a reading gesture. Some examples of the recorded signals and the eye gestures they represent, are shown in Figure 3. The formed database of eye gestures consisted of a total of 150 unique signals.

The formed eye movement gesture signals were used to simulate the movement of a 3-D graphics functional model of a human eye with a PCL. For the creation of a high quality 3-D graphics replica of the eye (Figure 4) we utilized the freely available 3-D graphics software package *Blender*. The eye model was carefully designed using the exact human eye specifications [8], and a colored PCL image was positioned on the eye surface.

In order to model the single-pixel capturing element of the eye tracking setting, we employed the accurate digital camera simulator provided by *ISET* software package [4]. For the case of color capturing, the single-pixel element was instantiated as a sensor array of 2 x 2 elements, with 8-bit quantization for each element.

Results

Eye Tracking Accuracy

In Figure 5, we present some representative examples of reference eye movement signals (directly recorded

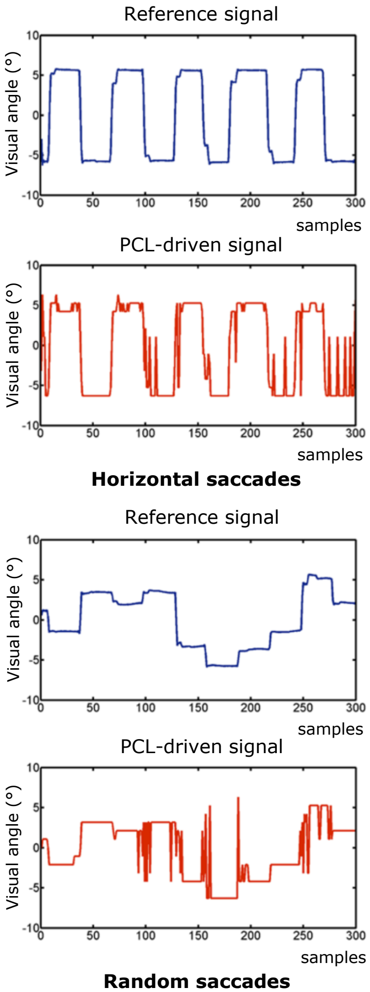

Figure 5. Reference signals and the corresponding PCL-driven signals.

from real subjects), and the signals reconstructed via the proposed eye tracking technique (using the described semi-simulated scenario). It can be visually verified that the PCL-driven (i.e. reconstructed) signals follow closely the shape of the original signals. There are, however, tracking inaccuracies in the form of spikes that deform the original shape.

In order to quantify the eye tracking accuracy of the proposed technique we will use the Mean Absolute Error (MAE) between the reference signals and the PCL-driven signals:

$$MAE = \sum_n \left| x_{reference}(n) - x_{PCL-driven}(n) \right| / N$$

(n the sample index, and N the number of samples). In Table 1, we show the average reconstruction MAE calculated over all 50 subjects for each one of the visual tasks (horizontal, random, and reading). The calculated average errors range from 0.98^o -1.90^o, with the horizontal saccade signals presenting the higher reconstruction error. This effect should be attributed to the larger eye movement amplitudes implicated in this specific task, which bring capturing close to the limits of the pattern. The lowest reconstruction error was observed in the signals from the reading task.

Recognition of PCL-driven Gestures

In order to assess if the signal reconstruction accuracy of the proposed technique allows for the development of a gesture recognition interface, we performed experiments and report the recognition rates in the following scenario: for every subject, we assume the reference gesture as the gallery sample, and the PCL-driven gesture as the probe sample. Then, we calculate the accuracy errors of each specific PCL-driven gesture (e.g. horizontal) with regard to the three reference gestures (horizontal, random, text) recorded from the same subject. Using these errors as dissimilarity measures we classify the specific PCL-driven gesture to one of the three classes of reference gestures. By repeating this process for all subjects we can calculate the respective recognition rates for the three categories of gestures. Table 2 shows the calculated gesture recognition rates, which are close to 100% in all cases.

We also explore a more diverse scenario, where instead of comparing the PCL-driven and the reference signals for every subject individually, we assess the between-subject gesture recognition accuracy. In this specific scenario we use only the horizontal and text reading gestures. The random gestures cannot be included, since the presented random pattern was different for each subject. The procedure for the calculation of the gesture recognition rates is analogous to the one described in the previous case. In the current case, though, we need to compute the errors between the PCL-driven signals among all subjects. Table 3 presents the calculated between-subject gesture recognition rates for the two types of gestures. As in the previous scenario, the gesture separation is very high, portrayed by the recognition rates of 100% in both cases. In Figure 6, we present the complete confusion matrix for all the signals from the two eye gesture classes.

We need to point out that the obtained results should be considered under the prism of certain limitations. The influence of large deviations from the previously stated assumptions, as well as the exact performance of the technique in case of a larger variety of users and target eye movement gestures, are still open fields of research and the main subjects during the extension of our current study.

Task	Hor. MAE	Vert. MAE
Horizontal	1.90^o	1.54^o
Random	1.43^o	1.27^o
Reading	0.98^o	1.12^o

Table 1. Reconstruction Errors.

Gesture type	Recognition rate
Horizontal	100%
Random	100%
Reading	98%

Table 2. PCL-driven gesture recognition.

Gesture type	Recognition rate
Horizontal	100%
Reading	100%

Table 3. Between-subject gesture recognition.

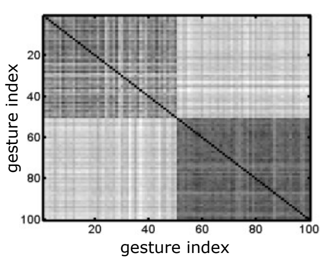

Figure 6. Confusion matrix for horizontal and reading gestures.

Conclusion

In this work, we performed a semi-simulated user study in order to assess the viability of a low-power eye tracking technique based on a patterned contact lens and a single-pixel capturing element. The experimental results show that, despite the observed eye tracking errors, the PCL-driven signals retain the shape of the original gestures sufficiently. Thus, the proposed technique can be considered suitable for the construction of eye gesture interfaces in low-power demanding applications.

Acknowledgement

This work is supported in part by Google Research Award #2014_R1_308 and NSF CAREER Grant #CNS-1250718. Special gratitude is expressed to Dr. E. Abdulin, T. Miller, and Ch. Heinich for proctoring eye movement recordings.

References

[1] Bolt, R.A. Gaze-orchestrated dynamic windows. SIGGRAPH Comput. Graph. 15, (1981), 109-119.

[2] Fares, R., Fang, S. and Komogortsev, O.V. Can we beat the mouse with MAGIC?. In Proc. CHI 2013, ACM Press (2013), 1387-1390.

[3] Hutchinson, T.E., White, K.P. Jr., Martin, W.N., Reichert, K.C. and Frey, L.A. Human-computer interaction using eye-gaze input. IEEE Transactions on Systems, Man and Cybernetics, 19, (1989), 1527-1534.

[4] ISET Software: http://imageval.com/831-2/

[5] Komogortsev, O.V., Ryu, Y.S., Do, H.K. and Gowda, S.A.M. Instantaneous Saccade Driven Eye Gaze Interaction. In Proc. International Conf. on Advances in Computer Entertainment Technology, ACM Press (2009), 1-8.

[6] Morimoto, C.H. and Mimica, M.R.M. Eye gaze tracking techniques for interactive applications. Computer Vision and Image Understanding 98, (2005), 4-24.

[7] Ono, Y., Okabe, T. and Sato, Y. Gaze Estimation from Low Resolution Images. Advances in Image and Video Technology 4319, Chang, L.-W. and Lie, W.-N., Eds., Springer-Verlag (2006), 178-188.

[8] Riordan-Eva, P. and Cunningham, E. Vaughan & Asbury's General Ophthalmology, 18th Edition, McGraw-Hill Education, 2011.

[9] Rozado, D., Agustin, J.S., Rodriguez, F.B. and Varona, P. Gliding and saccadic gaze gesture recognition in real time. ACM Trans. Interact. Intell. Syst. 1, (2012), 1-27.

[10] Sewell, W. and Komogortsev, O.V. Real-time eye gaze tracking with an unmodified commodity webcam employing a neural network. Ext. Abstr. CHI 2010, ACM Press (2010), 3739-3744.

[11] Stellmach, S. and Dachselt, R. Look & touch: gaze-supported target acquisition. In Proc. CHI 2012, ACM Press (2012), 2981-2990.

[12] Ware, C. and Mikaelian, H.H. An evaluation of an eye tracker as a device for computer input. In Proc. CHI 1987, ACM Press (1987), 183-188.

[13] Wobbrock, J.O., Rubinstein, J., Sawyer, M.W. and Duchowski, A.T. Longitudinal evaluation of discrete consecutive gaze gestures for text entry. In Proc. ETRA 2008, ACM Press (2008), 11-18.

[14] Vertegaal, R. The GAZE groupware system: mediating joint attention in multiparty communication and collaboration. In Proc. CHI 1999, ACM Press (1999), 294-301.

[15] Zhai, S., Morimoto, C. and Ihde, S. Manual and gaze input cascaded (MAGIC) pointing. In Proc. CHI 1999, ACM Press (1999), 246-253.