Eyeflect - Low-Cost Multi-Flash Gaze Estimation

Doug Fritz
MIT Media Laboratory
20 Ames Street, 02139 Cambridge, MA
doug@media.mit.edu

Richard The
MIT Media Laboratory
20 Ames Street, 02139 Cambridge, MA
rthe@media.mit.edu

ABSTRACT
Controlling the high speed flashing and reading of multiple infrared LED/phototransistors pairs we create a structure for lightweight camera-less gaze estimation in a mobile form factor. We show that it is possible to do gaze estimation of the human eye using only four leds and phototransistors, by a trick in the sequence and processing of flashing. Thus enabling cheap, and effective gaze estimation for a variety of HCI applications.

ACM Classification Keywords
Eyetracking, HCI, gaze-estimation, multi-flash, computational camera.

INTRODUCTION

Motivation
The cost of professional gaze tracking systems remains high, roughly $10,000 dollars, and almost every modern solution relies on bulky cameras for tracking. Even the DIY and opensource solutions require multiple cameras and while they are significantly cheaper remain intrusive and cumbersome. The only exception being electrooculogram (EOG) techniques which uses the electric field potential of the eye. This method however, while good at capturing saccades does not capture fixations because the electrical potential of the eye only varies during movement. This situation lead to a growing need for something that was optically based but cheaper and faster than a full camera system.

Contributions
For the first time, we have shown that it is possible to do optical tracking of the pupil, in a cheap, light-weight mechanism, without the use of a full camera. This is accomplished by exploiting the retroreflective features of the retina, the so-called red eye effect, in conjunction with cycling of light across multiple flashes spaced temporally near one another.

Related Work
Major related work has been done on camera based techniques. There have been several mainstream commercial products [1,2]. As well as several low cost DIY or opensource alternatives, namely opengazer and eyewriter [3,4]. The camera based techniques rely on determining the brightest pixel, the purkinje image, which is a corneal reflection of a light in a known fixed position and the ellipse of the iris. The shape of the ellipse and its relative position to the purkinje image is enough to estimate gaze. While this is accurate, it requires a high resolution camera to get an accurate ellipsoid curve. As an alternative technique the before mentioned electrooculogram method is currently under heavy development by NTT Docomo to effectively classify a series of eye gestures, but as stated due to its lack of accuracy with eye fixations it is not an effective methodology for full gaze estimation [5].

Our approach was highly influenced by the Non-photorealistic camera and other computational camera methods [6, 7], where the combination of multiple spatially arranged lights and the analysis of photos created with this custom illumination allowed a more accurate capturing of three dimensional features about the real-world environment.

Limitations and Benefits
The major limitations of the proposed system include an on body device, a calibration step, and the exposure of the participant to high amounts of IR light. We have yet to test and fully understand the upper limit of IR light exposure and intensities and their resulting potential damage to the human eye.

The major benefits are that our method is of significantly lower cost, less intrusive and cumbersome. Additionally it is very fast as it is not constrained by camera frame rates.

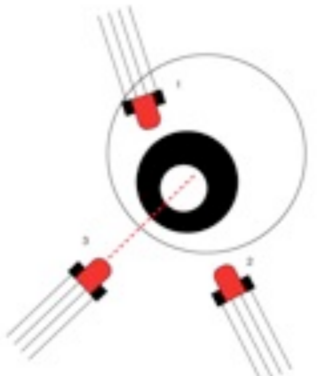

Figure 1. Diagram of the proposed setup: three leds and phototransistors pointed at the pupil will be triggered in sequence

METHOD
The core of the process is in managing the relationship between the four IR leds and the four IR phototransistors. The process is described below for our setup assuming that

leds and phototransistors referred to by the same number are placed spatially along the same angle of incidence to the eye and that n=4

```
i=0
loop {
  flash led #i
  read from all phototransistors and store
  subtract value #i with avg of last n-1 values
  store as size of pupil from position #i
  i = (i+1) % n
}
```

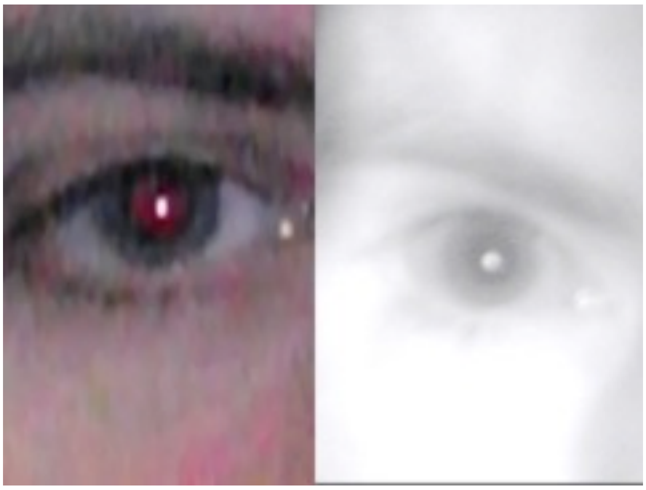

Figure 2. Retroreflection vs regular reflection: What appears as a retroreflection usually is just a large and bright reflection off the cornea (right).

What this expresses is for every led/phototransistor pair, we can determine the amount of retro-reflected light in that direction after a complete flashing cycle around the ring. The amount of retro-reflected light is correlated to the area of the ellipse of the pupil from said perspective. All ambient light effects and purkinje images are removed from the calculation by doing the background subtraction step from the average of all non retro-reflected light generated by flashing they eye at non-incident angles with equal intensity.

IMPLEMENTATION

First step was proving the principle of the device. As such it was necessary to prove that the desired conditions could be achieved in the human eye. And second that given those conditions the tracking system would function.

To achieve the first goal, it became important to prove that the eye was: one, retroreflective in IR and under what conditions, two that the variation could be measured and provides enough accuracy for triangulation. We conducted multiple experiments: First we took photos to create the red-eye effect within the visible spectrum. The light source was placed with differing distances to the imaging sensor to estimate the properties of retro reflection of the human eye. We found that it is only possible to achieve the red-eye

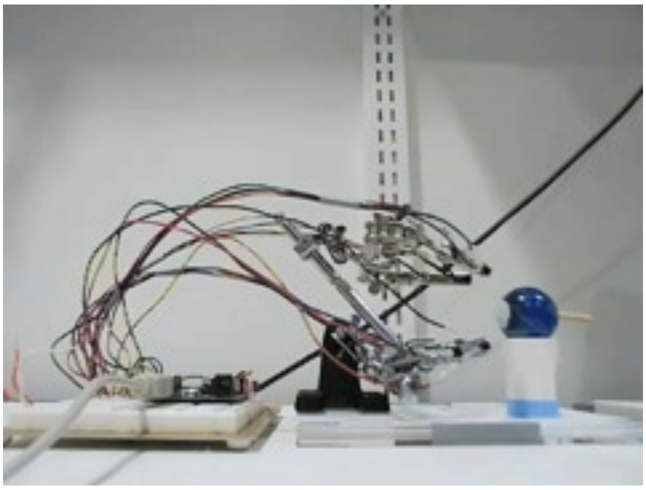

Figure 3. Diagram of the proposed setup: three leds and phototransistors pointed at the pupil will be triggered in sequence

effect when light source is powerful enough (camera flash light) and when light and sensor are very close together, almost within light of sight. The next step was to recreate the retroreflection only using infrared light, which was achieved placing an IR pass filter in front of a camera xenon flash. Using a camera that is sensitive to IR light we were able to create the red-eye effect using IR light only.

To achieve the second goal, first multiple simple prototypes with the proposed setup were built using regular off-the shelf infrared leds. As this approach has so many parameters that can affect the results (voluntary and involuntary motion of eye and head, angle of light source and sensor, light power) and due to safety concerns isolation of the problem was necessary: A model of the eye was built so that the assumed conditions could be studied under a more controlled environment for the creation of the device. The model included a moveable glass eyeball with a retroreflective pupil, before which four sensors and leds were placed. The leds and sensors were arranged in a rectangle and aimed at the pupil.

The rectangular arrangement of sensors proved to be a more viable approach over the original triangular arrangement for two key reasons. One, it increased our range of motion that the eye can travel in and fits more cleanly into a rectangular lcd which could be presented in front of the eye and embedded in the glasses. Second and most importantly, the area of the ellipse that becomes visible to a given sensor while moving orthogonally in the triangular formation follows an elliptical curve rather than a linear one so it is more difficult to interpret. These sort of nonlinear area integrations can be minimized by adding more led/phototransistor pairs in a ring around the rim of the glasses, but for each additional sensor there is an added time cost of cycling through one more ring in the series. Since one complete cycling needs to take place to measure the position of the eye, any movement during the time it takes to cycle through the sensors will cause inaccuracies in the interpreted signal. Hence, for our initial tests four points of data proved to be the optimal choice.

The values from the phototransistors were read and interpreted through an arduino module. Smoothing of the signal was done on board, but triangulation of position was done in an external java application. Eventually all code could be moved onboard. A calibration component was necessary, but post calibration the signal was solid with very little drift after extended use.

RESULTS

Performance Evaluation

The four sensors create a two inch by two inch plane roughly one-half inch from the model eye. We measured the range of motion after calibration (1.2) and the stability of signal (+- 0.002). Using these measurements we can estimate a gaze accuracy equivalent to being able to look at every pixel in a 600 x 600 pixel screen at 300dpi at said one-half inch distance from the eye.

FUTURE DIRECTIONS

Future directions, include building a fully working prototype, human testing, and experiments with form modifications. One experimental test is to coat the glasses surface with a hot mirror coating (a dichroic filter) to inhibit outside IR interference. By coating the lenses themselves in such a solution you both inhibit outside IR interference from sources like the sun as well as more evenly distribute the flashed IR light within the field in a predictable pattern.

CONCLUSION

We proposed a novel method for gaze tracking using low-cost, light-weight, camera-less gaze estimation in a mobile wearable device. The main innovation consists of using only four leds and phototransistors, a novel way to sequence and process the flashing in software. In order to simulate the main idea we built a functional prototype model of the system replacing the human eye with a physical prop.

ACKNOWLEDGMENTS

We thank Ramesh Raskar and Quinn Smithwick for their help.

REFERENCES

1. Tobii Technology AB http://www.tobii.com/corporate/eye_tracking/our_technology.aspx
2. Smart Eye AB http://www.smarteye.se/products.aspx
3. OpenGazer: open-source gaze tracker for ordinary webcams http://www.inference.phy.cam.ac.uk/opengazer/
4. Quan, T., Roth, E., Sugrue, C., Lieberman, Z. , Watson, T., Powderly, J.: The Eyewriter Initiative http://www.eyewriter.org/
5. Hiroyuki Manabe, Masaaki Fukumoto, Full-time wearable headphone-type gaze detector, *CHI '06 extended abstracts on Human factors in computing systems, April 22-27, 2006, Montréal, Québec, Canada* [doi>10.1145/1125451.1125655]
6. R Raskar, K Tan, R Feris, J Yu, M Turk Non-photorealistic Camera: Depth Edge Detection and Stylized Rendering using Multi-Flash Imaging. *Proceedings of ACM SIGGRAPH 2004* 23, 3 (2004) 679 - 688.
7. Retrographic sensing for the measurement of surface texture and shape M.K. Johnson and E.H. Adelson *Computer Vision and Pattern Recognition, 2009*