Free Head Motion Eye Gaze Tracking Without Calibration

Carlos H. Morimoto
Department of Computer Science – IME/USP
1010 Rua do Matão, São Paulo, SP 05508 Brazil
hitoshi@ime.usp.br

Arnon Amir and Myron Flickner
IBM Almaden Research Center
650 Harry Road, San Jose, CA 95120 USA
{arnon,flick}@almaden.ibm.com

ABSTRACT

This paper introduces a novel technique for remote eye gaze tracking and detection of point of regard that is specially designed for wide use in HCI. It addresses and eliminates two of the major problems of commercial remote eye gaze tracking, namely the need for user calibration before each session and of accuracy degradation with head movement. The new technique uses a single calibrated camera, several light sources with known positions and a physical model of the eye to estimate the 3D position of the eye and its gaze direction. Simulation results using ray tracing are used to study the accuracy and robustness of the system, and demonstrate its operability.

Keywords

Gaze based interfaces, remote eye gaze tracking.

INTRODUCTION

Gaze enhanced interfaces can significantly improve the usability of computer systems. By determining where the user is looking at, and analyzing the trace of eye movements and fixations, a system can infer the current user's task and infer about its context. It may predict a set of possible actions in order to be more reactive when the user requests such actions, or simply use the task context information to attentively offer/execute a set of actions. Although gaze based interfaces have proven to be useful for physically disabled users, and shown useful in other general applications [3,4], problems with the current eye tracking technology prevents such interfaces from reaching the consumer market as a standard input device.

Eye gaze trackers (EGTs) can be head mounted or remote. Head mounted EGTs are more accurate and allow for free head motion, but requires the user to wear the device. This is inadequate for regular daily use. Remote EGTs do not require any physical contact with the user, but cannot handle large head motion. Both techniques require an initial calibration procedure and typically achieve about 1 degree accuracy. A recent contribution towards solving these problems using multiple cameras is presented in [2]. The next section presents a new remote EGT that allows free head motion without calibration, and uses only one single camera, thus greatly improving the usability of EGTs without an increase in complexity or cost of the system.

CHI 2002, April 20-25, 2002, Minneapolis, Minnesota, USA.
ACM 1-58113-454-1/02/0004.

REMOTE EYE GAZE TRACKING

Current remote eye tracking techniques use a single near infrared light source to generate a reflection (glint) on the surface of the cornea (1st Purkinje image). This reflection is generally used as a reference point. Consider a spherical cornea, the position of the glint does not change with cornea rotations although the pupil position does. The 2D vector defined by the glint and the pupil center can be used to compute the point of regard using a direct mapping from the pupil-glint vector to computer screen coordinates. The mapping is computed during the calibration procedure, where typically the user has to fixate his/her gaze at known scene points in a particular order. Although the calibration procedure might take only a few seconds, it has to be made before each user session. In addition, this simple mapping model does not allow for free head motion. Hence the system might also require recalibration when the user changes his head position, or move her chair (typically, a remote EGT requires recalibration if the head moves only a few inches). These two problems must be solved to increase the usability of eye tracking as input devices.

Free Head Motion and Calibration Free Eye Tracking

We introduce a new technique based on the theory of spherical mirrors and the Gullstrand's eye model. Figure 1 shows the image formation process for a cornea modeled as a convex spherical mirror of center *C*, focal point *F* and vertex *V*. Further details about image formation on convex mirrors can be found in [1].

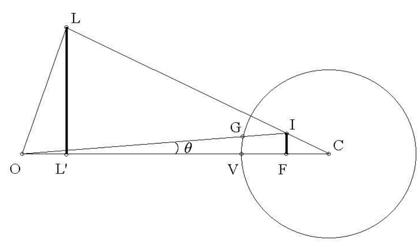

Figure 1: Image formation for an eye modeled as a convex mirror.

Consider a camera located at the origin *O*, and a light source at a known position *L*. Since the distance of the camera to the center of the cornea is much greater than the radius of the cornea, the image of *L* is formed at *F*, i.e., the image *I* of light source *L*, actually reflected from *G*, is seen as if coming from *I* by the camera at *O*. If the camera is calibrated, it is possible to recover the direction of the line *OGI*.

Now consider a second light source positioned on the optical axis of the camera very near *O*. This light generates a second glint on the surface of the cornea, and is reflected at *V*. Therefore, with a second on-axis light it is possible to compute the direction of the line *OVFC*, and the angle θ between the lines *OVFC* and *OGI*. The coordinates of *C* can now be computed if the cornea radius *r* is known. Note that all points in Figure 1 lie on a plane, and since the direction *OVFC* can be computed from the camera, only the module of the vector |*OC*| needs to be computed. Let the coordinates of *O*, *L* and *C* be (0,0), (*Lx*, *Ly*), and (*Cx*, 0) respectively. Then it is easy to show that *Cx* is the smallest positive value of the second order polynomial: $Cx^2 - Cx(Lx+r/2) + (r/2)(Lx - Ly/\tan(\theta)) = 0$.

With several light sources, a set of equations is obtained. This set can be used to obtain a more robust result, using least squares or an iterative method. The set of equations can also be used to estimate the radius of the cornea.

Pupil Center Computation

The vector connecting C to the center of the pupil defines the gaze direction. The pupil is approximately circular, and its diameter varies from 2 to 8mm, depending on ambient lighting conditions. It is located between the cornea and the eye lens, and lies inside the aqueous humor. The cornea and the aqueous humor have different indexes of refraction, but the air-cornea boundary is responsible for most of the refraction.

Due to the symmetry of the pupil, the refraction of the center of the pupil will be contained by the plane defined by the center of the pupil image, and the points *C* and *O*, called gaze plane. If two cameras are available, two gaze planes can be computed, and the direction of gaze is defined by the intersection of these planes [2]. For a single camera, it is required to estimate the true 3D location of the center of the pupil. A simple refraction model was adopted to estimate the 3D location of the pupil center. The pupil is modeled as an object 4.1mm behind a concave spherical surface (the cornea and aqueous humor) with index of refraction 1.376. Under these conditions, the image of the pupil will be formed about 3.5mm behind the cornea, i.e., the pupil image is seen from the camera as coming from a point 3.5mm behind the cornea, defining an inner sphere. The 3D pupil center is computed from the intersection of this inner sphere and the vector connecting the camera center *O* to the image of the pupil.

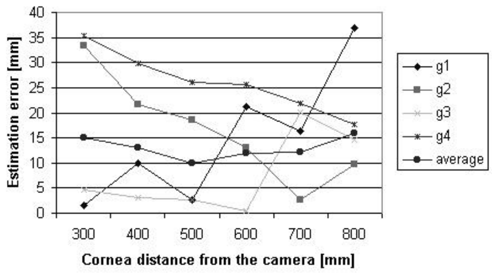

Figure 2: Estimation error for each glint.

SIMULATION RESULTS

Several synthetic images of a physical model of the eye were generated using ray tracing. The position of the glints and the center of the pupil were manually detected on images of resolution 320x240 pixels. The camera is placed at the origin, with 1.7 degrees vertical field of view. To generate the images of the eye model, it was placed at different distances along the camera's principal axis. Five light sources were used. One was positioned at the origin, two at the right (g1 and g2) and two above (g3 and g4) the camera, at 150 and 300mm from the origin in each direction. Figure 2 shows the results for the estimation of the cornea center for distances varying from 300 to 800mm. All images were generated using a single fixed gaze direction towards the point (100,200,0). Figure 3 shows the estimation error in the computation of the direction of gaze. Observe that the average error of 2.5 degrees is maintained over all positions.

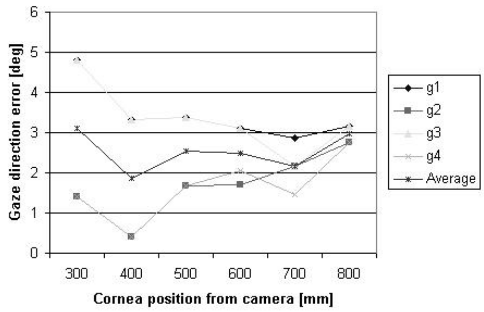

Figure 3: Gaze direction estimation error.

CONCLUSION

This paper introduces a new model for remote eye gaze tracking that eliminates the need for system calibration and allows for free head motion, using just a single camera and at least two light sources. The simulation results indicate that the accuracy of this method still has to be improved. Better results could be achieved by simply increasing the resolution of the images, but we believe that a more accurate refraction model combined with automatic subpixel glint and pupil detection algorithms and robust estimation techniques will significantly improve the robustness and accuracy of the system.

ACKNOWLEDGMENTS

We thank the Fundação ao Amparo à Pesquisa do Estado de São Paulo (FAPESP) and the Conselho Nacional de Pesquisa e Desenvolvimento (CNPq) for their financial support.

REFERENCES

1. Jenkins, F.A., and White, H.E. *Fundamentals of Optics*. 4th ed., McGraw-Hill, 1976.
2. Shih, S.W., Wu, Y.T., and Liu, J. *A calibration free gaze tracking technique*. ICPR 2000, vol.4, pp. 201-204.
3. Sibert, L.E. and Jacob, R.J.K. Evaluation of Eye Gaze Interaction. Proc. ACM CHI'2000, pp. 281-288.
4. Zhai, S., Morimoto, C., and Ihde, S. *Manual and Gaze Input Cascaded (MAGIC) Pointing*. Proc. ACM CHI'99, pp. 246-25