

Bases geométricas y de proceso para el seguimiento ocular.

Federico A. López Pardiñas (fede.a.lopez@gmail.com)

Monografía vinculada a la conferencia del Dr. Christer Norström, TOVII A.B. Suecia.-
“Métodos de interacción ocular con terminales informáticos”

Resumen — Eye gaze tracker (EGT) o seguimiento ocular son la base para poder implementar dispositivos interfaz (hombre - máquina) capaz de localizar y calcular hacia donde una persona está dirigiendo su mirada. Existen varios tipos de dispositivos de este tipo basados en diferentes técnicas.

Si bien tiene una infinidad de aplicaciones como pueden ser videojuegos, APP's de celular, cálculo de estadísticas para productos comerciales, entre otras, una de las más interesante es sin duda, la interacción con terminales informáticos utilizados por personas con discapacidad motora y el futuro desarrollo para el diagnóstico médico.

Este documento ilustra cuales son las bases que utilizan estos dispositivos para funcionar y extraer los datos que necesitan. Dada la variedad y combinaciones de métodos y algoritmos existentes, se nombran las distintas variantes existentes y se detalla solo uno de los casos que consiste en el método de pupilas claras y 1er punto de Purkinje.

Palabras claves — Eye tracker, Seguimiento de mirada, Gaze tracker, Gaze point estimation, Purkinje, corneal-glint reflection.

I. INTRODUCCIÓN

EL ojo humano tiene características y particularidades que nos permiten tener una estimación de en qué lugar del espacio estamos enfocando la mirada, para ello es necesario saber a grandes rasgos su funcionamiento. A continuación se detalla la estructura del ojo y se hace una breve descripción de cada punto, en la Figura 1 se muestran las partes del ojo de nuestro interés.

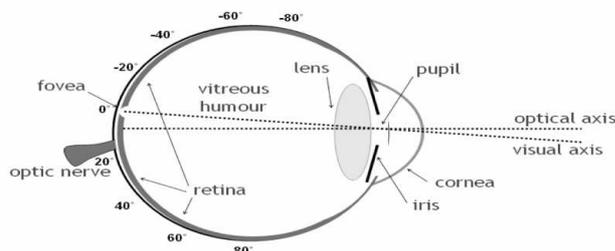


Figura 1. Estructura del globo ocular. Imagen extraída de [4].

A. Estructura del ojo

La retina es un tejido sensible a la luz situado en la superficie interior del ojo. Es similar a una tela donde se proyectan las imágenes. La luz que incide en la retina desencadena una serie de fenómenos químicos y eléctricos que finalmente se traducen en impulsos nerviosos que son enviados hacia el cerebro por el nervio óptico.

La córnea es la parte frontal transparente del ojo que cubre el iris, la pupila y la cámara anterior. La córnea, junto con la cámara anterior y el cristalino, refracta la luz y contribuye en mayor parte del poder de enfoque del ojo.

La fovea es el área de la retina donde se enfocan los rayos luminosos y se encuentra especialmente capacitada para la visión del color. Dirigir la vista hacia un objeto supone colocar su imagen óptica en la fovea.

La pupila es una estructura del ojo que consiste en un orificio situado en la parte central del iris por el cual penetra la luz al interior del globo ocular. Se trata de una abertura dilatada y contráctil, aparentemente de color negro que tiene la función de regular la cantidad de luz que le llega a la retina, en la parte posterior del ojo.

B. Angulo de visión

El ángulo de visión es el ángulo que se forma desde el centro de la fovea y el extremo horizontal, vertical, diagonal de la imagen. Se denominan ángulo de visión horizontal, vertical y diagonal respectivamente e indica el rango visible de la imagen en un plano.

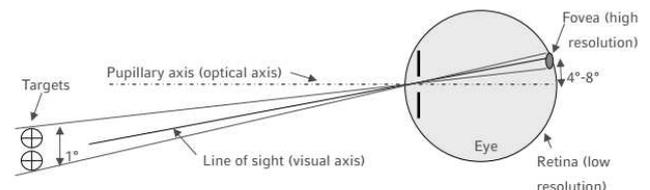


Figura 2. Zona de alta resolución y proyección en la fovea. Imagen extraída de [3].

Un aspecto importante a destacar es que si bien el campo visual del humano es de 220° horizontal y 120° vertical, solo se proyecta sobre el área de la fovea un ángulo de 1° a 2° que llamaremos ángulo de visión de fovea, como se ilustra en la Figura 2. Vale decir que por este motivo sólo somos capaces de enfocar con buena resolución y a una distancia de 1 metro un objeto equivalente al tamaño de la uña de un dedo.

C. Movimientos oculares

El ojo humano realiza a lo largo del día una serie de movimientos voluntarios e involuntarios, estos movimientos son tenidos en cuenta por los algoritmos de detección.

Movimientos compensatorios

Son movimientos reflejos y por lo tanto involuntarios e implican la participación coordinada de ambos ojos. Surgen como un mecanismo de fijación del campo visual ante movimientos de la cabeza o el tronco.

Movimientos sacádicos

Son movimientos voluntarios que nos permiten visualizar diversas zonas de una escena. Su objetivo es disponer la imagen visual en la fovea.

Movimientos de fijación

Los ojos sólo permanecen relativamente quietos para enfocar una zona concreta de la escena durante períodos de tiempo muy breves, comúnmente de 200 a 350ms.

Resumiendo, percibimos las imágenes moviendo los ojos y fijando los mismos en áreas de interés, partes concretas de la escena. De esta forma el ojo está dirigiendo la atención a una parte concreta de la imagen que está captándose entera en la retina y dirigiéndola para que encaje en la fovea, de este modo se puede ver con mayor grado de precisión y detalle.

II. EYE TRACKER'S (EGT)

Estos sistemas en general se dividen en dos partes, hardware y software. A su vez el software se divide en dos grandes partes, por un lado algoritmos de detección de mirada y por otro convertir la mirada a eventos en el PC, por ejemplo movimientos del ratón o eventos en el programa utilizado por el usuario final.

A. Tipos de Eye Tracker

Existen varios tipos de dispositivos de seguimiento ocular, ya que existen diferentes formas de detección de mirada. En todos ellos es muy significativo el efecto de la iluminación que se utiliza, posición, intensidad, longitud de onda y cantidad de focos luminosos.

En lo que sigue se enumeran diferentes métodos para sistemas de seguimiento ocular.

- 1- Oculografía eléctrica (EOG): Estos sistemas de EGT originales siguen los cambios en campos magnéticos cuando los ojos se mueven, ya que hay diferencias entre las polaridades del ojo de atrás hacia delante. Normalmente, electrodos son colocados encima y debajo y en los lados de los ojos.
- 2- Sistemas de bobina: Estos métodos siguen los movimientos del ojo observando una bobina magnética insertada en el ojo, sea quirúrgicamente o como parte de una lente de contacto. La cabeza del usuario debe estar sujeta, u otra bobina debe ser usada para analizar la posición de la cabeza. Este tipo de EGT es invasivo y peligroso, por lo que no es muy usado
- 3- Sistemas Purkinje duales: Estos sistemas siguen reflejos de luz en las partes delantera y trasera de la córnea. Calculando geoméricamente la orientación de estos reflejos, se puede determinar la posición del ojo. Normalmente este método es muy preciso, pero requiere tener la cabeza inmovilizada.
- 4- Sistemas de pupila claros: Si enviamos infrarrojos

directamente al ojo, se produce un efecto brillante en la córnea, similar al efecto de ojos rojos en una fotografía. Al seguir los movimientos de este reflejo claro, estos sistemas siguen movimientos del ojo orbitales. Usando un algoritmo calibrado, el sistema puede traducir estos movimientos del ojo a posiciones de la de la cabeza, o extrapolar el movimiento del ojo a otro plano.

- 5- Sistemas de pupila oscuros: Al igual que con el método anterior, estos sistemas iluminan el ojo con una luz infrarroja desde una cámara sensitiva a los infrarrojos. El ojo y la cara reflejan esta iluminación, pero la pupila absorbe la mayoría de los infrarrojos y aparecer 'a como una elipse oscura con mucho contraste. Se determina el centro de la pupila con un software de análisis de imagen, y 'esta es puesta en correspondencia con la posición de mirada usando un algoritmo. Estos sistemas son versátiles y más fáciles de configurar que los demás, pero también requieren algún tipo de compensación para los movimientos de cabeza.

Este documento centra su contenido en el estudio de un sistema de Pupila claros ya que al parecer es un método muy eficiente y facilita la detección de la pupila y glint's.

B. Principio de funcionamiento EGT de pupila claros

El dispositivo está basado en un sistema de cámaras y luz infrarroja, la luz infrarroja es reflejada por la retina a la superficie de la córnea creando puntos de reflexión los cuales son proyectados en la imagen percibida por la cámara como puntos brillantes. Estos puntos brillantes son conocidos como puntos de Purkinje y llamados "glints".

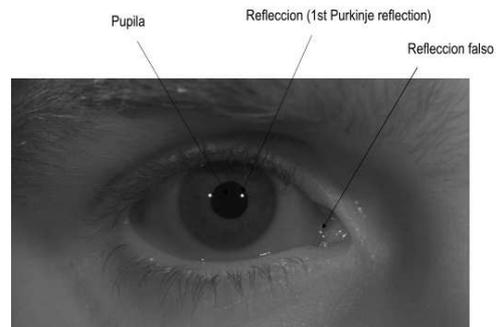


Figura 3. Glint's y falsos reflejos. Imagen extraída de [8].

La posición de los glints respecto a la pupila son usados por los algoritmos de detección, en los cuales es necesario contemplar el hecho de que existen falsos positivos, por ejemplo si el usuario utiliza gafas o incluso reflejos en el borde inferior del párpado, esto se ilustra en la Figura 3.

Como ya se mencionó es clave en estos sistemas hacer un buen uso de la iluminación, usando la intensidad y posición adecuada e incluso en algunos casos se utilizan filtros que permiten pasar hacia la cámara únicamente las longitudes de onda de interés. Además un hecho que no se puede ignorar es que la iluminación se debe calcular para que no ocasione daños al usuario dado que se le esta irradiando directamente al ojo.

Más adelante se enumeran los parámetros que deberán

ser determinados por los algoritmos de detección ya que hay distintos algoritmos algunos basados en geometría y otros con procesamiento matemático o redes neuronales. Estos algoritmos son los que definen el método y la cantidad de puntos necesarios para la calibración.

C. Bases geométricas

La distancia entre el reflejo sobre la superficie de la córnea conocido como el primer reflejo de Purkinje (P1), y la pupila varía con la rotación del ojo, pero no varía significativamente con el movimiento de la cámara ya que la cámara y la luz infrarroja están sujetas al mismo dispositivo y por ende se mueven en conjunto, entonces P1 sirve como punto de referencia con respecto a la posición de la pupila. La línea estimada de visión es calculada midiendo la separación entre el centro de la pupila y el centro de P1.

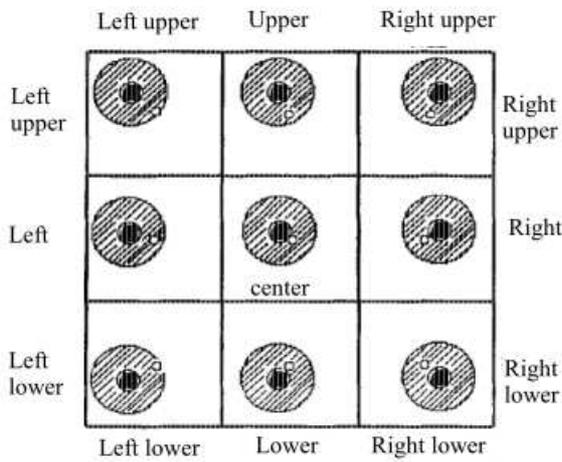


Figura 4. Posición cornea-glinc. Extraído de [1].

El cambio en la línea de visión es proporcional al vector diferencia entre estos puntos, esto se puede observar en la Figura 4. La relación geométrica (en una dimensión) entre la mirada y la separación del reflejo pupila-córnea (PCR) en caso de que la iluminación y la cámara estén en un mismo ángulo, está dado por la ecuación (1) y se ilustra en la Figura 5.

$$PCR = k \sin(\theta) \tag{1}$$

Donde θ es el ángulo de visión de mirada con respecto a la fuente de iluminación y la cámara, k es la distancia entre iris y el centro de la córnea, el cual se asume esférico. Este sistema funciona mientras no se sobrepase giros de 40° , pues el reflejo se pierde.

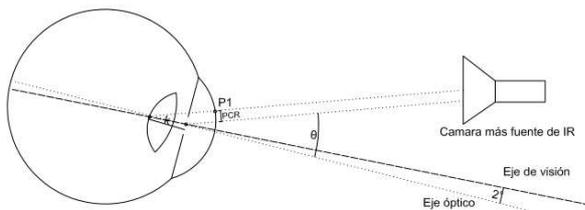


Figura 5. Características geométricas con la fuente de iluminación y cámara en un mismo lugar.

Una característica importante de este método es que permite ciertas variaciones en la posición de la cabeza.

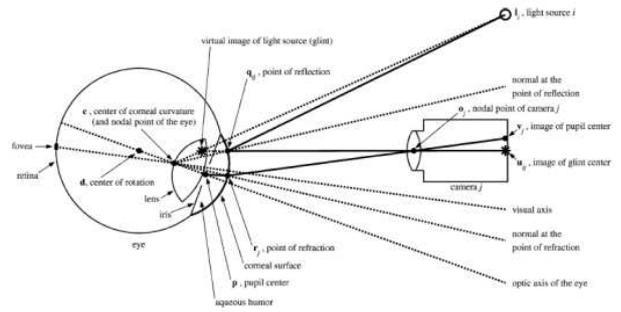


Figura 6. Características geométricas con la fuente de iluminación y cámara separadas. Extraído de [8]

En la Figura 6 se muestran las distintas configuraciones geométricas en caso de que la fuente de IR y la cámara estén en distintos ángulos, resulta más complejo deducir igualdades geométricas en este caso pero igualmente permite ilustrar que cumple con las bases vistas anteriormente y que la variación de la distancia cornea-glinc es proporcional a cambios en la posición del ojo.

Otra magnitud de interés para el sistema y para poder ubicar al objetivo, es la distancia a la que se encuentra el plano en que se fija la mirada del observador, esta se puede calcular usando las relaciones trigonométricas clásicas como se muestra en la ecuación 2.

$$d = a \cdot \tan^{-1} \left(\frac{\alpha}{2} \right) \tag{2}$$

Donde d es la distancia horizontal entre el observador y el plano observado, a es el tamaño del objetivo o punto de observación y α el ángulo de visión de fovea.

D. Bases de proceso

El procesamiento necesario para el EGT puede separar en partes bien definidas. Por un lado es necesario un algoritmo de detección de bordes que permita determinar el centro de la pupila y los glint's. Por otro se necesita mapear las variaciones captadas al mover los ojos en el plano de interés, por ejemplo el monitor de un PC.

Para la detección de la pupila y los glint's a partir de una imagen digital, existen varias técnicas por ejemplo las que se estudian en [9] y [10].

Para definir como se mapean estos parámetros en el plano en el que se fija la mirada, se puede usar directamente geometría 3D como la que se estudia en [7] y [2] haciendo una especie de extrapolación de la posición del centro de la pupila respecto a P1 (origen) hacia el plano de visión donde será necesario también fijar un origen. Otro camino es utilizar algoritmos que usan funciones basadas en parámetros de la geometría en 2D, redes neuronales CNN, aproximación por polinomios de orden alto como la que se plantea en la ecuación 3, u otros de mayor complejidad.

Por ejemplo una aproximación por polinomios de segundo orden sería:

$$\begin{cases} x_p = a_0 + a_1 x_e + a_2 y_e + a_3 x_e^2 + a_4 y_e^2 \\ y_p = b_0 + b_1 x_e + b_2 y_e + b_3 x_e^2 + b_4 y_e^2 \end{cases} \tag{3}$$

Donde (x_p, y_p) son las coordenadas del punto en la pantalla al que se está mirando, (x_e, y_e) es el vector pupilarglint, los parámetros a y b son los coeficientes de los polinomios, que se pueden determinar experimentalmente mediante un procedimiento de calibración. Utilizar un polinomio de mayor orden implica más puntos de calibración.

Una característica que no hemos tratado aún es como determinar si el ojo está posicionado en un lugar fijo o se está moviendo en busca de otro objetivo. Aquí entran en juego los movimientos compensatorios, sacádicos y de fijación. Se debe tener en cuenta que aunque estemos fijando la mirada en un objeto el ojo realiza movimientos que el algoritmo de detección debe tener en cuenta, ya que esto podría derivar en movimientos involuntarios del cursor (jitter), o detecciones erróneas en la dirección de mirada.

Para tener en cuenta estas características, los algoritmos existentes son de varios tipos:

- 1- Basados en velocidad. Utilizan la velocidad del ojo para detectar cuando hay una fijación, basándose en que el ojo tiene una velocidad menor en una fijación que en un movimiento sacádico.
- 2- Basados en dispersión. Estos algoritmos utilizan la posición del cursor para hallar la fijación, y se basan en que la dispersión es baja cuando hay fijación, para lo cual existe un umbral de dispersión, y cuando se baja de ese umbral, hay fijación.
- 3- Basados en área. Estos algoritmos miran solo si hay interacción en determinadas áreas de interés, lo que hace que si detecta que está allí la fijación, la sitúa directamente en el centro de esta área.
- 4- Basados en velocidad + dispersión. Este método tiene dos etapas: una de detección que combina velocidad y dispersión, y otra de smoothing, que calcula el promedio de las N fijaciones anteriores.

E. Procedimiento de calibración

En estos sistemas es necesario hacer una calibración por usuario, dado que existen diferencias entre un usuario y otro por ejemplo en las características físicas del ojo como puede ser el radio de la córnea, el uso de gafas, distancia usuario-objetivo. Además se corrigen ciertas no idealidades que se asumen durante los cálculos teóricos como por ejemplo la córnea no es esférica.

Durante este procedimiento se estiman los parámetros necesarios según el método empleado, por ejemplo los coeficientes del polinomio de mapeo o el radio de la córnea o la distancia entre el centro de la córnea y el iris.

El procedimiento consiste en identificar varios puntos en la pantalla o seguir un objetivo que va recorriendo la pantalla. A continuación se presenta una tabla comparativa del error que se obtiene con diferentes métodos.

Calibration Method	Error (degrees)
Linear	0.77
2nd-order polynomial	0.57
3rd-order polynomial	0.64
Homographic	0.58

Tabla 1. Comparación de distintos métodos de calibración. Extraído de [10].

III. CONCLUSIONES

Existen muchos métodos para el seguimiento ocular, cada método se diferencia de otro debido a la complejidad de la teoría utilizada, efectividad de acuerdo al caso de aplicación, costos de acuerdo al hardware necesario para implementarlo, comodidad del usuario, efectos secundarios.

Cualquiera de ellos implementa una solución válida y son de gran utilidad tanto para la ayuda de personas con discapacidad como su utilización para propósitos comerciales e investigaciones.

El método basado en pupilas claras, parece ser una muy buena opción ya que facilita la detección de la pupila y el primer reflejo de Purkinje, y utilizando un algoritmo de mapeo basado en polinomios de 2° orden se obtiene una buena precisión.

IV. REFERENCIAS

Publicaciones periódicas y conferencias:

- [1] CHI Jian-nan, ZHANG Peng-yi, ZHENG Si-yi, ZHANG Chuang y HUANG Ying, "Key Techniques of Eye Gaze Tracking Based on Pupil Corneal Reflection," WRI *Global Congress on Intelligent Systems*, MAY 2009.
- [2] Carol Y. Scovil, Emily C. King, y Brian E. Maki, "Determining the Visual Angle of Objects in the Visual Field: An Extended Application of Eye Trackers," *IEEE Trans. On Biomedical Engineering*, VOL. 56, NO. 3, MARCH 2009.
- [3] Heiko Drewes (2010, Aug.). "Eye Gaze Tracking for Human Computer Interaction," Dissertation an der LFE Medien-Informatik der Ludwig-Maximilians-Universität München 2010. Disponible en: http://edoc.ub.uni-muenchen.de/11591/1/Drewes_Heiko.pdf
- [4] J. C. Moya, "Implementación De Un Sistema Guiado Visual," PFC Universidad Carlos III, de Madrid. Abril 2012.
- [5] Martin Böhme, André Meyer, Thomas Martinetz, and Erhardt Barth, "Remote Eye Tracking: State of the Art and Directions for Future Development," The 2nd Conference on Communication by Gaze Interaction – COGAIN 2006.
- [6] S. C. Kim, M. Sked, y Q. Ji, "Non-intrusive Eye Gaze Tracking Under Natural Head Movements," International Conference of the IEEE EMBS San Francisco, CA, USA • Septiembre, 2004
- [7] Jixu Chen Qiang Ji, "3D Gaze Estimation with a Single Camera without IR Illumination," 978-1-4244-2175-6/08/\$25.00 ©2008 IEEE.
- [8] Christer Norström, "Tobii Assistive Technology," Facultad de ingeniería de la Universidad de la republia – Uruguay. Abril 2014.
- [9] Shashank Mathur, Anil Ahlawat, "Application of Fuzzy Logic on Image Edge Detection," Intelligent Technologies and Applications.
- [10] Dongheng Li, Derrick J. Parkhurst, "Starburst: A robust algorithm for video-based eye tracking," Human Computer Interaction Program Iowa State University, Ames, Iowa, 50011.

Libros:

- [11] T. Ohno, K. Hara, H. Inagaki, "Passive Eye Monitoring Algorithms, Applications and experiments," Ed. Springer, 2008.