

INSTITUTO TECNOLÓGICO DE SAN JUAN DEL RÍO

Amplificación interactiva de contenido por medio de detección de la dirección de la mirada

Autor: J.Carlos ÁVILA RESÉNDIZ

CICATA

25 de enero de 2016

Índice

Introducción

El presente documento esta enfocado únicamente en los resultados del estudio del código existente del proyecto que da nombre a este documento, **Amplificación interactiva de contenido por medio de detección de la dirección de la mirada**.

Dado que en el trabajo previo no se incluyo documentación alguna de sobre la estructura del proyecto y/o de análisis de los requerimientos del sistema, es por ello que antes de seguir con el proyecto se hace necesario el documentar el trabajo existente.

1. Visión artificial

La visión artificial o visión por computadora es la ciencia y la tecnología que permiten a las máquinas ver o extraer información de las imágenes digitales y resolver alguna tarea o entender la escena que están visionando. este conjunto de técnicas nos permiten diseñar interacciones, de modo que el usuario utiliza su movimiento o manipulación de objetos para interactuar con la aplicación.

1.1. Espectros de luz

La

2. Seguimiento ocular

El seguimiento de ojos (traducido del inglés eye tracking) es el proceso de evaluar, bien el punto donde se fija la mirada (donde estamos mirando), o el movimiento del ojo en relación con la cabeza. Este proceso es utilizado en la investigación en los sistemas visuales, en psicología, en lingüística cognitiva y en diseño de productos. ¹

2.1. Tipos de seguimiento

2.1.1. Mediante métodos invasivo

Utilizando algo adjunto al ojo como una lente de contacto especial con un espejo incorporado o un sensor de campo magnético. El movimiento de la unión se mide con

¹[Seguimiento ocular: Wikipedia](#)

el supuesto que no se deslice de manera significativa cuando el ojo gire. Las mediciones realizadas con lentes de contacto han aportado grabaciones extremadamente detalladas de los movimientos oculares. Las bobinas magnéticas es el método que utilizan para realizar estudios sobre la dinámica y la fisiología subyacente al movimiento del ojo.

2.1.2. Mediante métodos no invasivo

El segundo tipo de seguimiento sería sin necesidad que haya contacto. A través de la luz, por lo general luz infrarroja, se refleja en los ojos y se capta mediante una cámara de vídeo o algún otro sensor óptico. La información recogida se analiza para extraer la rotación de los ojos y los cambios en los reflejos.

Los métodos ópticos, especialmente los basados en la grabación de vídeo son ampliamente utilizados para el seguimiento de la mirada y están bien considerados porque no son invasivos y el coste es bajo.

2.1.3. Mediante potenciales eléctricos

Finalmente, el tercer tipo utiliza el potencial eléctrico medido con electrodos colocados alrededor de los ojos para detectar el movimiento. Los ojos son el origen de un constante campo de potencial eléctrico que también se puede detectar en total oscuridad aunque estos estén cerrados. Puede estar modelado para generar un dipolo con el polo positivo en la córnea y el polo negativo en la retina. La señal eléctrica que se puede derivar al uso de dos pares de electrodos de contacto colocados en la piel alrededor del ojo se llama electrooculograma ² (EOG). Si los ojos se mueven de la posición del centro hacia la periferia, la retina se acerca a uno de los electrodos, mientras que la córnea se acerca al opuesto.

Este cambio en la orientación de los dipolos cambia consecuentemente los resultados del campo potencial eléctrico la señal (EOG) medida.

2.1.4. Tecnologías y técnicas

La extensa mayoría de diseños actuales son seguidores de ojos basados en vídeos. Una cámara enfoca uno o los dos ojos y graba sus movimientos mientras el sujeto mira una serie de estímulos. Los seguidores de ojos más modernos usan el contraste para

²[Electrooculograma](#)

localizar el centro de la pupila y crear un reflejo de la córnea a través de luz infrarroja e infrarroja cercana no colimada.

El vector entre estas dos características puede usarse para computar la intersección de la mirada con una superficie después de una simple calibración individual. Se usan dos tipos generales de técnicas de seguimiento de ojos: pupila brillante y pupila oscura. La técnica de la pupila brillante genera un mejor contraste iris/pupila debido a un seguimiento de ojos más correcto en relación a la pigmentación del iris y reduce significativamente las interferencias producidas por las pestañas y otras características ocultas.

Esto además permite un seguimiento en condiciones que van desde la total oscuridad hasta una claridad alta. Estas técnicas no son muy efectivas para hacer seguimientos en exteriores ya que se producen interferencias en su monitorización.

Las configuraciones de los seguidores de ojos varían mucho; algunos se montan en la cabeza, otros requieren la cabeza solo para ser estables (por ejemplo con un apoyaméntón), y algunos siguen los movimientos de la cabeza de forma remota y automática durante el movimiento. La mayoría usan una frecuencia de muestreo de al menos 30 Hz. Aunque 50/60 Hz es lo más común, actualmente muchos seguimientos de ojos basados en vídeo funcionan a 240, 350 o incluso 1000/1250 Hz, frecuencia que se necesita para captar en detalle los rápidos movimientos durante la lectura o durante los estudios de neurología.

El movimiento de los ojos normalmente se divide en fijaciones y salidas, cuando la mirada se detiene en cierta posición y cuando se mueve hacia otra posición respectivamente. Las series resultantes de las fijaciones y las salidas se llama *scanpath*. La mayoría de la información de los ojos se hace disponible durante la fijación, pero no durante la salida.

Los uno o dos grados centrales (*la fovea*) aporta la mayor parte de información; los inputs de las excentricidades más extensas (la periferia) dan menos información. Por lo tanto, la localización de las fijaciones a lo largo del *scanpath* muestran que puntos de información de los estímulos son procesados durante una sesión de seguimiento de ojos. De media, las fijaciones duran alrededor de 200ms durante la lectura de textos lingüísticos y 350ms durante la visión de una escena. Preparar la salida hacia un nuevo objetivo lleva alrededor de 200ms.

3. Análisis de requerimientos

3.1. Funcionales

El sistema debe ser capaz de realizar todas y cada una de las siguientes acciones:

- Detectar o reconocer los ojos.
- Determinar con alto nivel de exactitud la dirección de la mirada.
- Realizar el seguimiento de la dirección de la mirada.
- Efectuar la ampliación del área de la pantalla en la cual se está enfocando la mirada.
- Tener una interfaz de configuración, que contemple lo siguiente:
 1. Calibración inicial.
 2. Entrenamiento del usuario.
 3. Configuraciones personalizadas.
- En su versión inicial debe correr en computadoras con sistema operativo Windows³.

Cumplidos estos requisitos definidos con anterioridad se considera que el prototipo funcional para máquinas con SO Windows está completo.

3.2. No funcionales

- Debe ejecutarse en máquinas con características básicas⁴.
- Debe trabajar en imágenes de baja resolución.
- Debe ejecutarse en tiempo real.
- Debe ser lo suficientemente fluido *optimizado* al realizar el seguimiento.

³Pese a que ese es el requerimiento inicial, se busca que sea multi-plataforma y en especial enfocado a dispositivos móviles

⁴Consideramos características básicas un procesador con al menos dos núcleos y 1Gb de memoria RAM

4. Diseño

Los componentes que se describirán a continuación son única y expresamente los que ya existen en el proyecto existente y sobre el cual se se esta trabajando, y que la finalidad del mismo es tener un documento que describa de la manera mas certera posible el funcionamiento tanto individual como en conjunto.

4.1. Módulos del sistema

Por razones que escapan de mi conocimiento, no hay documentación de ningún tipo ya sea técnica, de usuario o de análisis de factibilidad del proyecto, las siguientes secciones son de cierta forma meras interpretaciones en base al código existente del proyecto existente, por ahora solo con el único fin de tener una perspectiva clara de lo que ya existe y poder ver que partes de ello se pueden reutilizar en un futuro, ya sea parte del código o módulos completos.

4.1.1. Main Window

4.1.2. New Hough

Módulos de la clase:

- NewHough
- circle_hough
- circle_houghpeaks
- circle_points
- getLUT
- getEllipseLUT
- ellipse_hough
- ellipse_houghpeaks
- intersection

4.1.3. Eye-Gaze

Módulos de la clase:

-

4.1.4. Trainer

Su función principal es de entrenar tanto al algoritmo de reconocimiento en cascada de OpenCV como al usuario, en etapa de desarrollo se usa para el entrenamiento del algoritmo con la finalidad de que aprenda a discriminar todo aquello que parece ser ojos de los reales ⁵.

4.1.5. Addaptative Canny

⁵ Considero que lo ideal es separar estas dos características, por un lado un modulo que se encargue de calibrar el sistema y otro que entrene al usuario en el uso de la aplicación.s

Referencias