

Seguimiento Ocular Para La Interacción Con El Medio

G.A. Nava¹, C. Ramírez², G. Mendizabal-Ruiz³
Universidad de Guadalajara, Guadalajara, Jalisco., México

¹ ga_nh@hotmail.com

² carlogerstein@hotmail.com

³ gerardo.mendizabal@academicos.udg.mx

Resumen—Se desarrolló un programa que permite la interacción del usuario con la computadora a partir del movimiento de los ojos. El sistema cuenta con diversas aplicaciones potenciales dentro del ámbito médico y demás áreas; desde funciones auxiliares en pacientes con poca movilidad hasta aplicaciones que permiten su uso como herramienta dentro del hogar. La presente versión del proyecto expone un método computacional que permite el cambio de las páginas de un libro electrónico, a partir de la posición en la que se encuentra el iris de la persona y auxiliándose de técnicas algorítmicas que facilitan el reconocimiento facial y la detección de círculos. Como resultado se obtuvo un programa que permite a la computadora detectar e interpretar el movimiento de los ojos, mostrando así la página deseada y brindando al usuario la oportunidad realizar su lectura cómodamente.

Palabras clave—**Detección de iris, Seguimiento, Violajones, E-book, Transformada de Hough**

I. INTRODUCCIÓN

En la actualidad existe una problemática que genera un gran impacto en aquellas personas a las que les es imposible realizar tareas que implican el movimiento de sus extremidades, ya sea por la carencia de las mismas o bien por disfuncionalidad. Las personas con discapacidades pueden llegar a sentirse ahogadas en sus aspiraciones, intereses, ilusiones y, en definitiva, en su vida laboral. Su motivación se va adormeciendo por los factores externos (físicos) e internos (emocionales).

Una manera muy viable de sustituir la ejecución de tareas como tener que presionar teclas o manipular el cursor de una computadora, es el seguimiento ocular, mejor conocido como la técnica de *eye tracking* (Fig. 1), una solución tecnológica que tiene por objetivo extraer información del usuario a partir del análisis de sus movimientos oculares.

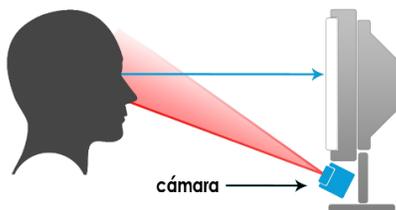


Fig. 1. Proceso de detección ocular (Modificado de goo.gl/3aB5kS).

El trabajo presenta una propuesta para las personas con anomalías congénitas o adquiridas del hombro, brazos, manos (ausencias, defectos longitudinales, transversales de los huesos) o defectos articulares (anquilosis, defectos de función, etc.).

Con la finalidad de encontrar una alternativa a la problemática planteada, se ha desarrollado un sistema que detecta la posición del iris en personas con dificultades motrices, permitiendo ejecutar instrucciones de acuerdo a la intención del usuario. El campo de acción de dicho sistema abarca diversas áreas y representaría una importante herramienta en el sector médico, resultando particularmente útil al momento de buscar mantener la autonomía e independencia de pacientes que se encuentran en condiciones de salud en las que su movilidad se ve afectada o limitada.

Si bien el campo de acción del sistema es amplio, la presente versión del proyecto busca enfocarse principalmente en demostrar el correcto funcionamiento del programa, así como su utilidad. Para dicho fin, se ha desarrollado un sistema que detecta la posición del iris y permite cambiar la página de un libro electrónico a voluntad del usuario.

La finalidad principal del trabajo es hacer posible la interacción entre la computadora y el usuario con las discapacidades aquí abordadas. Una persona con discapacidad motivada puede aprender más, acercarse al conocimiento y potenciar sus capacidades con mayor facilidad. Se espera que el desarrollo de este sistema permita a las personas con discapacidad aumentar su independencia y comodidad al momento de realizar sus actividades.

II. METODOLOGÍA

El método presentado requiere una cámara web para la captación de imágenes del usuario desde la zona del tórax hacia arriba, a manera de fotografía tamaño credencial o pasaporte.

El programa presenta una interfaz que muestra imágenes representativas que se utilizan como indicadores de acciones. Por ejemplo si el cambio de página se realiza a la derecha, izquierda o bien, si se permanece leyendo en la página actual. Antes de comenzar, el programa inicializa los contadores que controlan el desplazamiento de las páginas del e-book.

El método propuesto consta de cinco pasos:

1. Detección de elementos faciales.

Para el reconocimiento facial, se aplicó el algoritmo de Viola-Jones (VJ) [1], mediante el cual también se detectó:

1. Rostro: cuadro delimitador de la cara (Fig. 2 a)
2. Par ocular: cuadro delimitador de los ojos (Fig. 2 b)

El algoritmo VJ (Fig. 2) es un método de detección de objetos que se caracteriza por su bajo coste computacional, hecho que le permite ser empleado en tiempo real [1]. Su principal aplicación, misma que se emplea en el presente trabajo, es la detección de caras. Para dicha detección, VJ utiliza elementos básicos de iluminación (*Haar-like features*) característicos de un rostro común, que se basan en la diferencia de intensidades luminosas en las regiones rectangulares adyacentes de la cara.

Para lograr realizar la suma de píxeles de un rectángulo, VJ utiliza la representación de la imagen real en una imagen integral, misma que en un punto (x, y) contiene la suma de todos los píxeles que están arriba y hacia la izquierda de ese punto en la imagen original. El proceso de aprendizaje se basa en un algoritmo llamado AdaBoost (*adaptive boosting*) [4], cuya función es clasificar las características de la imagen integral y así crear una serie de clasificadores que en conjunto se conocen como Clasificador en Cascada. Este clasificador tiene como objetivo principal distinguir entre dos estados: si existe rostro o no. Consiste en una serie de “filtros” que van aumentando su complejidad para llegar a una conclusión efectiva [1].

2. Delimitación de la región del par ocular.

Posterior a la detección, se procedió a la captación de un fotograma (*snapshot*) que fue sometido a un proceso de adecuación, en el cual se convierte la imagen a escala de grises para poder ajustar el contraste y trabajar con ella de una manera más adecuada en cuestión de las funciones utilizadas. Se creó un cuadro delimitador (*bounding box*) para el rostro, esto con la finalidad de aislar dentro del cuadro de menor tamaño posible, la región que se considere rostro según el algoritmo mencionado. Un posible caso a presentarse es la detección de más de un rostro. Para evitarlo, se creó un comparador, el cual evalúa la medida del ancho del cuadro delimitador de los rostros encontrados y se queda con el más grande para trabajar con él. Se llevó a cabo un proceso similar para crear el cuadro delimitador del par ocular, sometiéndolo a la misma comparación y por ende, ambas decisiones del cuadro más grande deben coincidir.

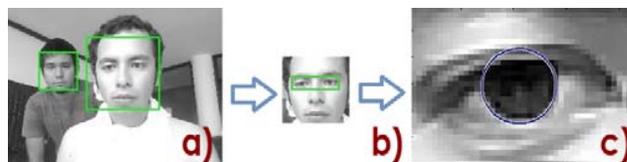


Fig. 2. Procesos de detección de rostro, par ocular y segmentación a un solo ojo.

3. Segmentación de un solo ojo.

Una vez teniendo el cuadro delimitador para el par ocular, se redimensionó, recortando el ancho del cuadro a $1/3$ para aislar un solo ojo y así facilitar el procesamiento de la imagen. La Figura 2 muestra los procesos de detección.

4. Obtención de la posición del iris.

Utilizando la Transformada de Hough (TH) [2], se aplicó una máscara sobre el par ocular recortado a un solo ojo. Esto permitió la identificación del iris mediante una función que proporciona la posición de los centros de los círculos identificados y los valores de los radios de los mismos.

La TH es una herramienta muy comúnmente utilizada para la detección de figuras en una imagen digital [2]. Se basa prácticamente en el análisis de una imagen para encontrar puntos que, alineados, satisfagan la ecuación de cierta figura, una línea recta o curva por ejemplo [5]. En este caso se detectó la presencia de círculos.

Una vez detectada la posición y el tamaño del círculo que representa al iris, se trazó una circunferencia de color en los círculos encontrados para la mejor apreciación de su localización y movimientos. A partir de la posición en coordenadas de los centros de los círculos encontrados (iris), se definieron las distancias de los bordes laterales hacia el centro del círculo del iris de la siguiente manera: La distancia del borde izquierdo al centro del círculo es equivalente a la suma de 0 más la coordenada en x del centro del círculo. Por consiguiente la distancia del borde derecho al centro del círculo equivale al ancho del cuadro delimitador para el par ocular dividido entre 3, menos la distancia del borde izquierdo al centro del círculo, previamente obtenida.

5. Ejecución de la instrucción deseada.

Una vez teniendo el conocimiento de las medidas de los bordes laterales hacia el centro del círculo localizado, para identificar la posición del iris, ya sea a la derecha, a la izquierda, mirada al frente o ausencia de ojo, se utilizó la lógica descrita en la Figura 3.

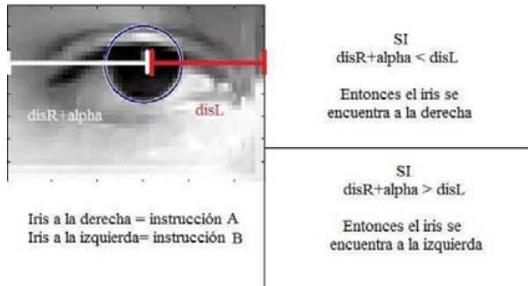


Fig. 3. Lógica para la medición de distancias en la ejecución de instrucciones.

Se puede observar en la Figura 3 que dependiendo de la posición del iris, se ejecuta una instrucción. La instrucción A indica el cambio a la siguiente página, mientras que la instrucción B se refiere al regreso a la página previa. Debido a que al redimensionar el cuadro delimitador para el par ocular, el iris se encuentra un poco más hacia la derecha, entonces fue necesario utilizar un valor de *offset* (α) para emparejar distancias del centro hacia los bordes. Si el iris permanece en la posición en la que se detectó en el *snapshot* anterior, sin desplazarse hacia la derecha o izquierda, las distancias permanecerán iguales, por lo que no habrá cambio de página y se mostrará una imagen que indique que el usuario está leyendo la página actual.

Debido a que en la lectura y en condiciones normales los ojos permanecen en constante movimiento [3], es necesario contar con un sistema de desbloqueo para poder cambiar de página; cuando el programa no detecta el iris, se habilita dicho cambio. A fines prácticos, esto se logra cerrando los ojos por unos segundos y al momento de abrirlos, mover la pupila al sentido en que se quiera mover la página.

III. RESULTADOS

El programa se presenta a través de una interfaz en la que se observa lo siguiente: el rostro delimitado del usuario (Fig. 5 a), una imagen que indica si se está leyendo o se está moviendo la página (Fig. 5 b), su ojo con el iris detectado (Fig. 5 c), una imagen que indica si el sistema de cambio de páginas está activado o desactivado (Fig. 5 d) y, por último, una imagen representativa de la página que se está leyendo (Fig. 5 e).

Antes de continuar, es importante mencionar que la totalidad del programa fue llevada a cabo utilizando el software Matlab, así como funciones propias de dicho entorno. Además, vale la pena aclarar que el valor " α " utilizado durante la programación del código y el cual permitió obtener los resultados más exactos, fue de 16.

En un inicio, el usuario se encuentra leyendo el libro con el sistema de cambio de páginas desactivado, por lo que podrá mover su ojo en cualquier dirección y no generará ninguna modificación. Cuando el usuario se encuentra en proceso de

lectura, se muestra la imagen de un libro con un fondo verde, lo que permite identificar que la persona simplemente está leyendo. Al cerrar los ojos, se puede comprobar cómo se activa el sistema de cambio de páginas. La imagen de cambio de página pasa de estar en "Off" a estar en "On". Una vez activado el cambio de páginas, si el usuario mira a la derecha o a la izquierda, la página cambia hacia adelante o hacia atrás de acuerdo a la posición del iris (Figs. 4 y 5).

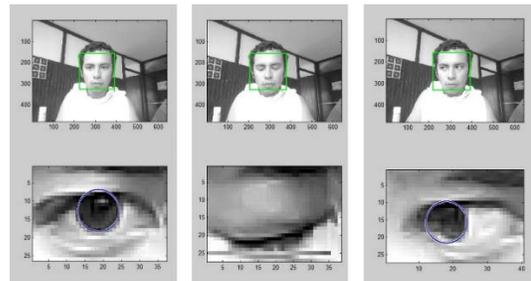


Fig. 4. Detección ocular para la ejecución de tareas.

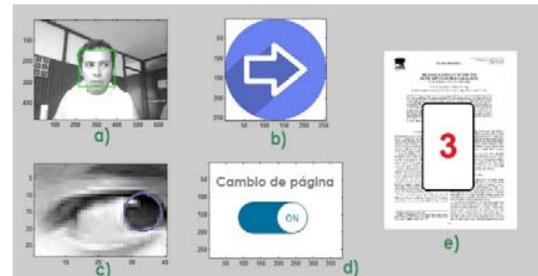


Fig. 5. Usuario provocando el cambio a la siguiente página.



Fig. 6. Sensibilidades de la TH en el software utilizado: A) Sensibilidad de 0.37 B) Sensibilidad de 0.72 C) Sensibilidad de 0.99.



Fig. 7. Sensibilidad de 0.93 (Adecuada).

Se realizaron pruebas para elegir el grado de sensibilidad adecuada para la detección de círculos con la función "imfindcircles" (Transformada de Hough), las cuales tuvieron los resultados mostrados en la Figura 6. Se puede observar que en el rango de sensibilidad de 0.30 a 0.72 no se alcanza a detectar el círculo del iris, sin embargo, con una sensibilidad de 0.99 se detectan círculos que no son de interés. Entonces, tras unas cuantas pruebas más, se concluyó que la sensibilidad adecuada era de 0.93 (Fig. 7).

Validación

Para validar el trabajo, se realizó una serie de pruebas con 10 sujetos. Estas consistieron en comprobar la precisión del sistema mediante la repetición de las ejecuciones de las tres acciones más funcionales. En la Tabla 1 se muestran los resultados del proceso de validación. Cada sujeto realizó 5 ejecuciones por cada acción presentada en la tabla.

A partir de los resultados del estudio de pruebas para validación presentado la Tabla 1, y promediando el error relativo porcentual individual, se puede establecer que el sistema tiene un porcentaje de error del 7.3% en condiciones de iluminación óptima, dato que indica que el sistema puede ser considerado eficiente.

IV. DISCUSIÓN

Observando los resultados gráficos del sistema en funcionamiento, se puede apreciar que se cumplió el objetivo de crear una interfaz usuario-ordenador que controla de manera eficiente el cambio de página de un libro electrónico en el momento en que el usuario lo demande, siempre y cuando se tenga una iluminación adecuada para la correcta detección de los círculos del iris.

Abordando el tema de la iluminación, se notó que el programa no funciona de la manera ideal cuando se ejecuta en un ambiente con una luminosidad deficiente o con la presencia de muchas sombras.

En cambio, cuando se ejecuta en un lugar donde la luz ilumina el rostro desde varios ángulos, se aprecia el funcionamiento que se esperaba como objetivo primordial.

TABLA 1. VALIDACIÓN DEL SISTEMA.

Acción	No. De repeticiones	Ejecuciones correctas	Ejecuciones incorrectas	Porcentaje de error
Activación del cambio de página	50	45	5	10%
Cambio a la siguiente página	50	48	2	4%
Cambio a la página previa	50	46	4	8%

Una dificultad importante fue la de encontrar la sensibilidad idónea para la detección de los círculos que representan al iris. Después de indagar los parámetros más utilizados, se experimentó con distintas cifras, comprobando que debajo del 0.80 no se detectaba el iris como círculo y arriba del 0.95 se detectaban círculos fuera de contexto, por lo que se concluyó que el parámetro de sensibilidad efectivo era de 0.93. Para la realización de este proyecto, se decidió

que cambiar hoja por hoja sería lo más adecuado. Es decir, al activar el sistema de cambio de páginas, el usuario tiene la posibilidad de cambiar hacia la izquierda o derecha una sola vez por cada activación. Posteriormente, cuando termina de leer, si quiere cambiar de página de nuevo, vuelve a cerrar los ojos y el sistema se activa una vez más. En un proceso de lectura normal, esta funcionalidad resulta bastante conveniente, sin embargo, resultaría aún más útil contar con una variación en la que se pudiera avanzar o regresar varias páginas rápidamente y no sólo una por una.

De manera general, se considera que los principales inconvenientes del algoritmo son la incapacidad del código para trabajar en un entorno con una iluminación deficiente, la simplicidad de las actividades que permite realizar y la velocidad de ejecución de la interfaz. El equipo continuará trabajado para solventar dichas carencias y mejorar la calidad y funcionalidad del programa en versiones futuras.

V. CONCLUSIÓN

El sistema funciona de la manera esperada y con un pequeño margen de error, cumpliendo con los objetivos propuestos al momento de plantear el proyecto. Parece ser que el programa será de gran utilidad y tiene potencial. Aún pueden realizarse varias mejoras, por lo que se espera poder continuar con el trabajo durante los próximos meses con la finalidad de pulir distintos detalles.

El equipo se encuentra a favor de una sociedad incluyente en donde se tome en cuenta a toda persona y sin que se discrimine por las capacidades físicas. Este proyecto busca dicha inclusión y se espera que pueda fomentar una actitud similar en los demás miembros de la sociedad.

BIBLIOGRAFÍA

- [1] Parra, E.: Aceleración del algoritmo de Viola-Jones mediante rejillas de procesamiento masivamente paralelo en el plano focal. *idUS* (2015)
- [2] Rhody, H.: Segmentación. Transformada de Hough. Chester F. Carlson Center for Imaging Science (2005)
- [3] Rayner, K.: Eye Movements in Reading and Information Processing: 20 Years of Research. *Psychol Bull.* 124, 372–422 (1998). DOI: 10.1037//0033-2909.124.3.372
- [4] Viola, P., Jones, M.: Robust Real-Time Face Detection. *International Journal of Computer Vision* 57, 137–154 (2004). DOI: 10.1109/ICCV.2001.937709
- [5] Ito, Y., Ohyama, W., Wakabayashi, T., Kimura, F.: Detection of Eyes by Circular Hough Transform and Histogram of Gradient. *Pattern Recognition (ICPR)* (2012)