

# Prototipo de comunicación audible a través del movimiento ocular con procesamiento digital de imágenes

Diego Alejandro Sánchez Alfonso  
Neiry Dayán Martínez Lara  
Ingeniería Biomédica  
Universidad Manuela Beltrán, Colombia  
diego.san97@hotmail.com  
neirydml@gmail.com

Andrés David Pérez Rodríguez  
Jorge Andrés Alvares Triana  
Ingeniería Biomédica  
Universidad Manuela Beltrán, Colombia  
perez.andresdavid@yahoo.es  
Jorge.alvarezt@docentes.umb.edu.co

**Resumen** – En este trabajo se desarrolla un prototipo de comunicación audible basado en el procesamiento digital de imágenes, con el fin de brindar una herramienta económica que mejore la calidad de vida de personas con limitaciones de comunicación, causadas por la parálisis completa de la movilidad del cuerpo como es el caso de la esclerosis lateral amiotrófica donde se conserva únicamente la movilidad ocular. En este estudio, se establece un prototipo con una cámara de 640x480 píxeles de resolución, basado en el rastreo ocular del ojo derecho; clasificando el movimiento a partir de la segmentación del iris y nueve patrones de movimiento preestablecidos. El procesamiento inicia con la identificación de parámetros anatómicos del ojo y continúa con métodos de segmentación donde se incluye: el recorte la imagen, eliminación de componentes como pestañas y brillos especulares, corrección del brillo, transformación al modelo de color HSV, identificación de estructuras circulares y asignación de etiquetas. Para la experimentación se tuvo en cuenta 12 voluntarios en tres condiciones de luz: homogénea alta, homogénea baja y no homogénea, lo cual demostró una sensibilidad mayor al 94% en la clasificación de los nueve patrones en condiciones de luz homogénea alta, es decir, sin sombras y con alto contraste de estructuras en la imagen.

**Palabras clave** — *Comunicación, Clasificador Kernel, Máquina de soporte vectorial, Rastreo Ocular, Trastorno severo de movilidad.*

**Abstract** – This work develops an audible communication prototype based on digital image processing in order to present an economical tool that improves the quality of life of people with communication limitations. This is caused by the body’s complete paralysis as sclerosis lateral amyotrophic, where just eye mobility is preserved. In this study, a prototype is established with a 640x480 pixels resolution camera to make an ocular tracking of the right eye, making a movement classifying based on the iris segmentation and nine preset movement patterns. Processing begins with the identification of eye’s anatomical parameters and continues with segmentation methods where is included: image cropping, removal of components such as eyelashes and specular highlights, brightness correction, transformation to the HSV color model, identification of circular structures and tag assignment. For experimentation, the prototype was tested on 12 volunteers in three light conditions: high homogeneous, low homogeneous and non-homogeneous. It showed a sensitivity greater than 94% in the classification of the nine patterns in conditions of high homogeneous light, it means without shadows and high contrast of structures in the image.

**Keywords** — *Communication skill, Eye tracking, Kernel classificatory, Severe mobility disorder, Vector support machine.*

## I. INTRODUCCIÓN

La esclerosis lateral amiotrófica (ELA) es una enfermedad neurodegenerativa de rápido progreso que provoca el deterioro del sistema nervioso central [1]. En la evolución de la enfermedad, la afectación del tronco encefálico desarrolla progresivamente la alteración de la musculatura facial; por lo cual, los últimos síntomas manifiestos se focalizan en el déficit de la articulación de palabras y debilidad de la musculatura respiratoria, sin embargo, se conserva la movilidad ocular [2]. Hay estudios que incluyen a personas con ELA en etapa tardía que han mostrado una mejora significativa en la calidad de vida de las personas que usan un sistema de Eye-Tracking, además de una carga reducida para los cuidadores [3].

Hoy en día las herramientas tecnológicas son un estímulo para mejorar la calidad de vida de las personas con ELA [4]. Actualmente existen dispositivos de rastreo ocular como: el iRiter, que funciona a partir de una pantalla con interfaz gráfica y permite escribir o dibujar; y el Tobii C15, que permite controlar un ordenador a través de la reconstrucción 2D del ojo. Aunque estos dispositivos operan eficazmente, es importante resaltar que son de alto costo y se deben importar desde otros países. Por esta razón, este trabajo escrito se orienta en el diseño y construcción de un prototipo de comunicación audible a través del movimiento ocular. Su funcionamiento se basa en la clasificación de patrones y generación de palabras, con el fin de presentar una herramienta portátil y económica que permita a las personas con ELA interactuar socialmente y responder conforme a sus emociones, elevando su estado de ánimo y mejorando su calidad de vida.

La estructura metodológica del proyecto se basa en cinco etapas: la primera consiste en identificar los parámetros del diseño incorporando ergonomía y funcionalidad en el rastreo ocular; la segunda etapa abarca los métodos de rastreo ocular para la segmentación del iris en tiempo real; la tercera etapa corresponde a la clasificación de palabras a partir de patrones simples y compuestos; la cuarta etapa corresponde a la reproducción de palabras y detección del parpadeo y por último la quinta etapa valida el funcionamiento del clasificador en diferentes condiciones ambientales. Finalmente, el dispositivo requiere de un hardware compuesto por

una cámara y un procesador, como lo es un ordenador; y un software que permita la ejecución de algoritmos en tiempos cortos. Estos dispositivos generan un alto grado de independencia para personas la enfermedad anteriormente mencionada, y se encaminan a los bajos costos para la asequibilidad de toda la población.

## II. MARCO TEÓRICO

### A. Sistema mecánico del ojo.

Corresponde a los movimientos del ojo, su análisis se lleva a cabo por un sistema fijo de tres coordenadas (X, Y, Z) que atraviesan su centro de rotación. El eje Z siempre coincidirá con la vista del ojo en posición primaria [5]. Respecto al centro de rotación se reconocen los siguientes movimientos:

#### 1) Movimiento horizontal entorno a Y.

Aducción, si la córnea se dirige hacia la línea media del cuerpo; abducción, si la córnea se dirige hacia afuera, de manera distal [5].

#### 2) Movimientos verticales entorno a X:

Elevación, cuando la córnea se dirige hacia arriba; depresión, cuando la córnea se dirige hacia abajo [5].

### B. Parámetros anatómicos del ojo

La córnea se encuentra en la cámara anterior, con un radio de curvatura de 8mm aproximadamente y tiene un diámetro de 12mm con una forma de menisco cóncavo [5]. La zona central de la córnea es la zona óptica, donde se da la formación de la imagen, esta sección mide cerca de 4mm de diámetro. Por otro lado, el iris cumple la función de regular la cantidad de luz que pasa hacia la retina a través de la pupila, esta última tiene un diámetro promedio de 2mm a 3mm cuando se encuentra en presencia de un haz de luz y 8mm en ausencia de luz; este tamaño disminuye cuando aumenta la edad [5].

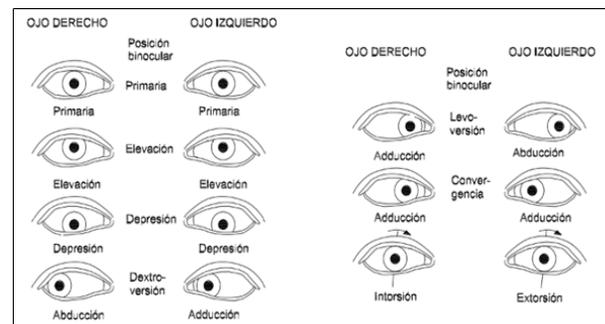


Figura 1. Movimientos monoculares y binoculares del ojo humano [5].

C. *Procesamiento digital de imágenes.*

El procesamiento digital de imágenes (PDI) cuenta con una diversidad de técnicas que se pueden implementar para filtrar componentes, reconocer y contar objetos, resaltar y segmentar regiones, detectar bordes, entre otras. Para esto se basa en la conversión de un campo de imagen continuo a una forma digital equivalente, por lo que se consideran procesos de cuantización (conversión de imagen analógica a imagen digital), muestreo de imágenes, operaciones inversas, análisis de matrices, métodos de representación de imágenes en el espacio vectorial, entre otros [17].

III. MATERIALES.

Para el diseño del prototipo de comunicación audible se utilizaron los materiales bajo costo y fácil acceso para la elaboración física del prototipo. Estas consideraciones garantizaron su bajo costo, sin embargo se resalta la necesidad de un ordenador portátil para ejecutar el software necesario para interactuar con la calibración y puesta en marcha de la adquisición de imagen de la cámara USB. El entorno de programación empleado es Python 3.5, para el procesamiento digital de la imagen se emplean las librerías de *OpenCV* versión 2, esta selección de software se realiza por ser un software libre y gratuito. Por último, el costo final del prototipo es de 109.000 pesos colombianos sin tener en cuenta el ordenador, el costo y especificaciones técnicas de los materiales utilizados en la construcción del prototipo se observan en la Tabla 1.

Tabla 1. Materiales y costos empleados en el prototipo de comunicación audible

Componente	Especificaciones técnicas		
	Resolución de video 640 x 480	No aplica	
Mini Cámara (SmartCam USB Connection)	Iluminación mediante leds blancos	6	\$80000
	Transmisión de datos por USB 2.0	1	
	Distancia focal 4cm – 6cm	No aplica	
	Ángulo de visión 70°	No aplica	
Gorra con broche	Tela y visera	1	\$7000
Acrílico transparente	30cm x 35cm	1	\$13900
Tornillos con tuerca	3.0mm	4	\$1200
Tornillos con tuerca	2.5mm	6	\$1200
Arandela	1.5cm	2	\$400
Arandela	3.0cm	2	\$400
Tomillo sin fin	Plástico	1	\$500
Tuerca	Plástica	2	\$200
Baquela		1	\$1000
Pila CR2032	3 voltios	1	\$1500
Diodo LED	Color blanco	4	\$400
Interruptor	3 pines 2 posiciones	1	\$300
Tubo plástico		1	\$500
Porta pila	Redonda para CR2032	1	\$500
TOTAL			\$109000

1. Link: funcionamiento del prototipo: <https://youtu.be/3i5ZNgQ1Hvw>

IV. METODOLOGÍA

El tipo de estudio es experimental, con un alcance exploratorio ya que muy poco se ha abordado sobre el tema en el país, además es un desarrollo tecnológico orientado para personas con trastornos severos de movilidad con limitaciones comunicativas.

El método empleado para el desarrollo y evaluación de patrones se evidencia en la Figura 2. La metodología se clasifica en cinco fases que se desarrollarán a continuación:

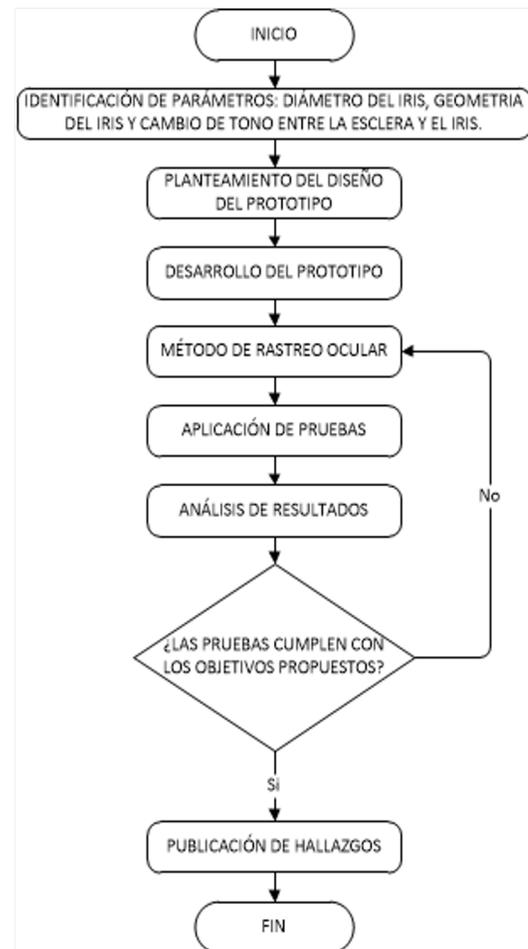


Figura 2. Diagrama de flujo del proyecto.

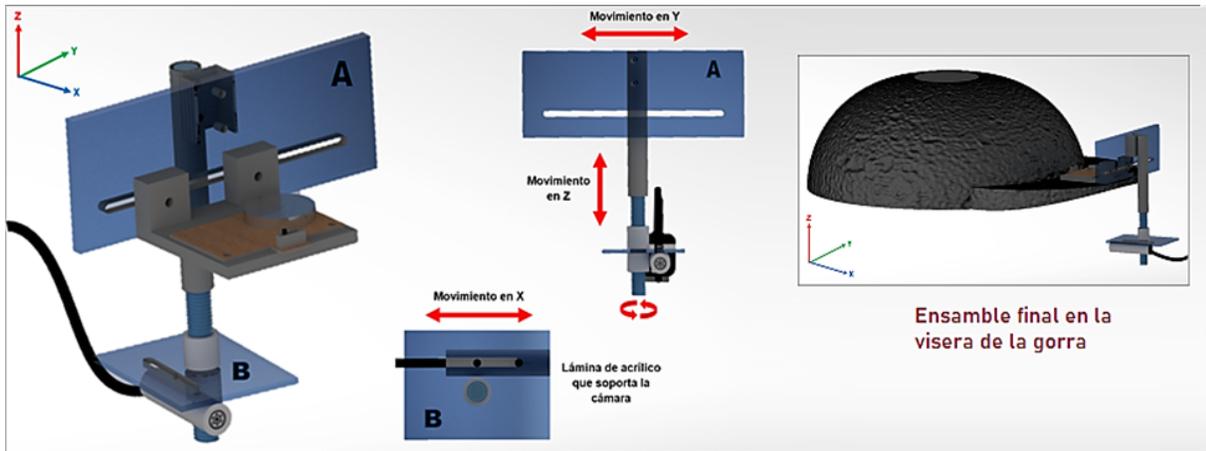


Figura 3. a) Soporte principal, movimientos en Y y Z, b) Soporte de la cámara, movimiento en X.

A) Fase 1: Identificación de parámetros y diseño del prototipo:

1) Identificación de parámetros.

Se tiene presente el diámetro del ojo quien en promedio mide 23mm y 25mm y se considera una estructura casi esférica [14]. Las estructuras visuales de la imagen del ojo permiten identificar la semejanza del iris a una circunferencia lo cual favorece su identificación. Por último se considera únicamente las rotaciones monoculares correspondientes al ojo derecho. Tomando como referencia central y neutra el movimiento primario, los movimientos correspondientes a aducción, elevación y abducción se etiquetarán con un número para seleccionar el patrón.

2) Planteamiento del diseño del prototipo.

Después de plantear diferentes diseños de soporte para la cámara USB, se descarta el uso del marco de lentes y la diadema a causa de la inestabilidad y la causa de sombras en la imagen durante su uso. Por esto se usa una gorra como base de soporte del prototipo con un sistema de rieles que permite mover la cámara en los tres planos (Ver Figura 3). Además, posee anclada una baqueta en la visera con un circuito de alimentación básico que permite la iluminación LED durante 6 a 8 horas continuas aproximadamente. Los materiales utilizados en su construcción se enlistan en la Tabla 1.

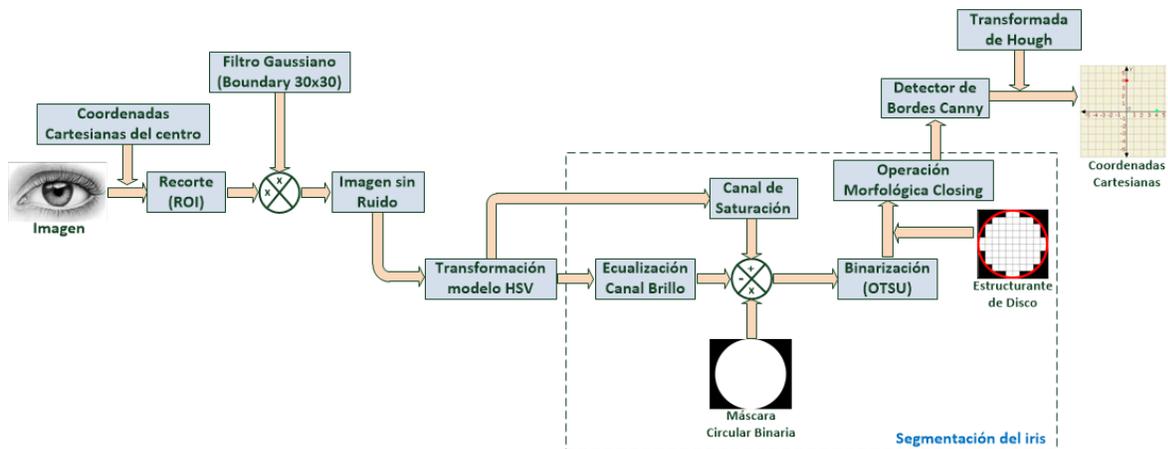


Figura 4. Diagrama de bloques para el proceso de segmentación del iris

1. Link: funcionamiento del prototipo: <https://youtu.be/3i5ZNgQ1Hvw>

Estos materiales favorecieron tres características principales:

- a) El ajuste a la anatomía craneal de cualquier persona gracias al broche de la parte trasera de la gorra.
- b) La ubicación correcta de la cámara para la captura de la imagen gracias a los 3 grados de libertad en el movimiento.
- c) La reducción de peso en el diseño lo que favorece la portabilidad y enfoque de la imagen.

*B) Fase 2: Método de rastreo ocular.*

A continuación se profundiza en el procesamiento digital de la imagen adquirida, esta parte del proceso permite segmentar el iris de la imagen y adquirir su ubicación bidimensional en la imagen con el fin de clasificar patrones posteriormente (Ver Figura 4)

*1) Recorte de la imagen.*

Debido a la distancia focal de la cámara, en la imagen se pueden observar estructuras diferentes al área ocular como cejas, orejas y cabello. El recorte permite eliminar estos componentes considerados ruido para la segmentación igualando las dimensiones de la imagen a las dimensiones de la cavidad orbitaria. A partir de la relación centímetro-píxel identificada en la cámara y la regresión lineal de los datos (Ver Tabla 2)

**Tabla 2. Dimensiones de la imagen.**

Distancia Ojo-Lente (mm)	Dimensiones de la imagen (mm)	
	Alto	Ancho
40	23	31
60	34	45
80	46	62
100	57	77

$$\text{Ancho} = 0.775x - 0.5 \quad (1)$$

$$\text{Alto} = 0.57x + 0.1 \quad (2)$$

*2) Eliminación de componentes.*

Las pestañas y los brillos especulares ocasionan obstrucción en el iris dificultando su segmentación, para eliminarlos se propone utilizar un filtro gaussiano que suavice los bordes y elimine la mayor parte de estos componentes. El estudio elaborado por

Daugman propone el uso de un filtro de Gabor [15], pero éste supone más carga computacional y ralentiza el proceso de identificación. El filtro gaussiano utiliza un estructurante de tamaño 30x30 con el que se logra difuminar los componentes. Adicionalmente, se implementó una máscara circular binaria de radio r que elimina cualquier sombra que se encuentre en los bordes de la imagen y ocasione falsos positivos durante la identificación del iris.

*3) Corrección del brillo y segmentación.*

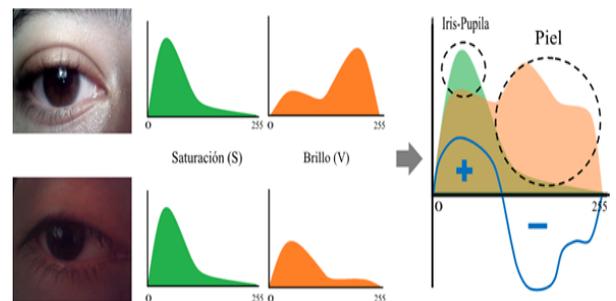
Para el procesamiento del brillo en las imágenes, se transforman los valores del campo de color RGB al campo de color HSV. Al igual que las sombras, el iris presenta una propiedad relacionada con la falta de cromaticidad, lo que significa presencia de estructuras sin color, refiriéndose, para el campo del HSV, a las tonalidades negras [16]. Una vez obtenido el histograma correspondiente para cada canal, se concluye que la piel y el iris poseen características cromáticas distintas entre ellos. Mientras la piel se aleja de las tonalidades puras, el iris y la pupila mantienen un equilibrio de cromaticidad en todos los canales (Ver Figura 6.(E))

*4) Binarización de la imagen.*

El proceso de binarización de la imagen permite convertir los valores de la matriz en dos tipos de datos (0 y 1), permitiendo finalizar con la segmentación del iris usando el método OTSU; este método usa distribuciones estadísticas para umbralizar la imagen, El resultado se puede observar en la Figura 6(F).

*5) Detección de estructuras circulares en la imagen.*

En este proyecto se propone el uso de la transformada circular de Hough la cual, en conjunto al detector de bordes Canny, encuentra círculos en la imagen con radios previamente establecidos.



**Figura 5. Comportamiento de los histogramas y diferencia de los canales S y V.**

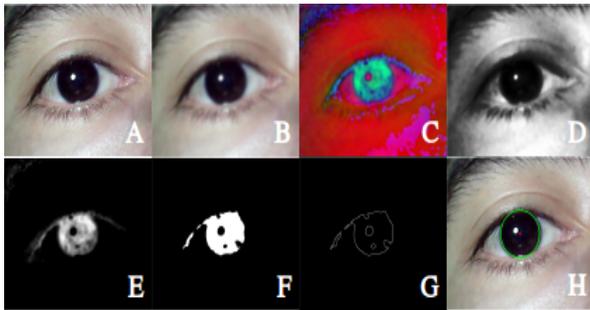


Figura 6. Etapas del procesamiento de la imagen: a) Imagen ROI, b) Filtro gaussiano, c) campo de color HSV, d) Ecuación del histograma del canal V, e) Sustracción de canales S y V, f) Binarización con método Otsu, g) Detección de bordes Canny, h) Transformada de Hough.

La transformada circular de Hough permite detectar círculos con base a la ecuación de la circunferencia donde se establece un radio como parámetro principal. Se requiere establecer un rango de radios que facilite la búsqueda de los círculos. Para determinar estos valores, se tiene en cuenta 3 datos:

Con las ecuaciones 1 y 2 y el tamaño fisiológico del iris que es aproximadamente 12 mm se determinan los radios máximo y mínimo a usar en la transformada de Hough, teniendo en cuenta las distancia máxima y mínima de la cámara al ojo correspondiente a 90mm y 70mm respectivamente.

Tabla 3. Medidas del iris en pixeles.

Distancia Ojo-Lente (mm)	Dimensiones pixel (mm)	Diámetro Ojo (px)	Radio Ojo (px)
70	0.084	142	71
90	0.107	112	56

En la Tabla 3 se pueden observar los valores correspondientes en pixeles para el diámetro y radio del iris; para evitar un posible rango de error en la medición, se establece un radio mayor de 55px y menor de 75px. Se toman coordenadas cartesianas del punto central del círculo identificado y su respectivo radio, estas coordenadas son utilizadas para determinar la ubicación del iris en la imagen.

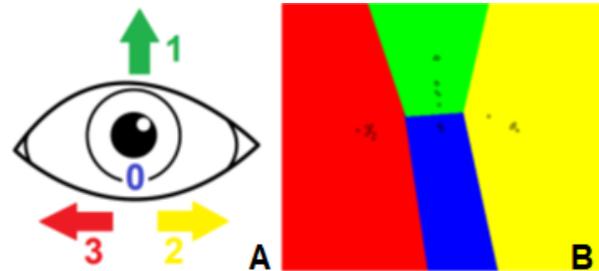


Figura 7. Aplicación de SVM: a) Etiquetado con número y color para cada movimiento, b) resultado del entrenamiento con SVM.

C) Fase 3: Entrenamiento de Super vector machine (SVM)

1) Calibración del SVM.

Se realiza una calibración previa al proceso de clasificación, donde se solicita al usuario por medio de indicaciones audibles observar hacia: arriba, izquierda, derecha y al centro cada 5 segundos<sup>1</sup>. Esto permite al algoritmo guardar 10 coordenadas cartesianas para cada movimiento en una matriz asignándole una etiqueta correspondiente a cada movimiento (Ver Figura 7(A)), la cual se empleará para el entrenamiento de la máquina de soporte vectorial. Esta matriz en conjunto con las etiquetas, suministran la información necesaria para entrenar el SVM, el cual mediante un Kernel lineal multiclase separa los datos en hiperplanos lineales. El resultado del entrenamiento se muestra en la Figura 7(B).

2) Configuración de patrones simples y compuestos.

Una vez adquirido el rastreo ocular se propone un sistema de patrones simples y compuestos a partir de las etiquetas adquiridas por el clasificador. Caracterizando al movimiento primario del ojo derecho como el indicador final de un patrón. Un patrón simple se compone de un movimiento simple más el retorno del movimiento ocular a la posición primaria (ver Figura 9); Así mismo, un patrón compuesto se compone de un movimiento simple más otro movimiento simple (diferente al inicial) más un tercer movimiento de retorno a la posición primaria del ojo (Ver Figura 9).

1. Link: funcionamiento del prototipo: <https://youtu.be/3i5ZNgQ1Hvw>

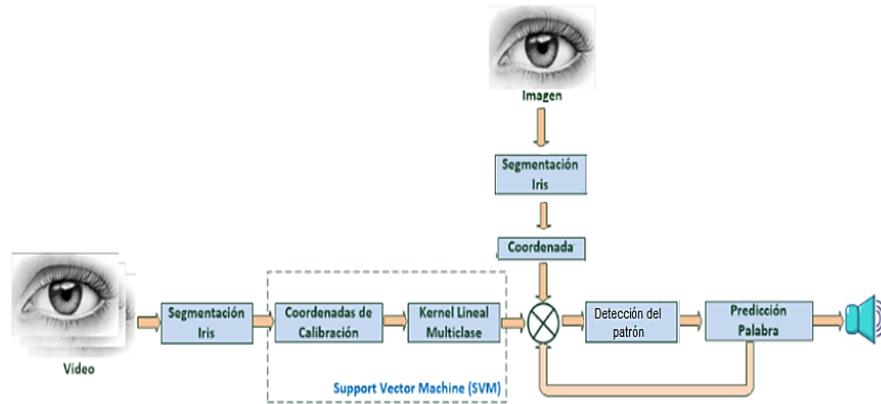


Figura 8. Etapas de los movimientos simple y compuesto.

Este sistema de patrones garantiza al usuario la capacidad de elegir palabras previamente definidas en el vector nombrado “Words” (ver Tabla 4). Para realizar la detección de los patrones se almacena el valor de las etiquetas clasificadas en la etapa 1 y 2 (Ver Figura 9) en las variables  $a$  y  $b$  respectivamente, estas variables determinan la posición  $m$  del vector Words en el que se almacenan las palabras, tal como se profundiza a continuación:

$$m = a + 3b \quad (3)$$

La variable  $m$  indica la posición del vector “Words”,  $a$  corresponde a la etiqueta correspondiente al movimiento asignado en la primera etapa y  $b$  etiqueta correspondiente al movimiento capturado en la segunda etapa. Determinando así la elección de una palabra a partir de un movimiento simple o compuesto.

Tabla 4. Vector de palabras para la comunicación

Posición del vector: WORDS	Palabra
Cero (0)	----
Uno (1)	Baño
Dos (2)	Si
Tres (3)	No
Cuatro (4)	----
Cinco (5)	Hambre
Seis (6)	Sed
Siete (7)	Dolor
Ocho (8)	----
Nueve (9)	Hola
Diez (10)	Gracias
Once (11)	Jajajaja

D) Fase 4: Reproducción de palabras y detección del parpadeo.

1) Reproducción audible de palabras.

Una vez asignada la etiqueta y establecido el patrón, el algoritmo reproduce la palabra seleccionada del vector “Words” a partir del sintetizador de voz digital *playsound* de las librerías de Python (Ver Figura 9).

2) Identificación del parpadeo.

El procedimiento utilizado para la detección del parpadeo tiene como fundamento la comparación entre dos capturas continuas. El cálculo del parpadeo se obtiene a partir de la ecuación 4, donde  $I[i]$  corresponde a la captura de la imagen actual e  $I[i-1]$  a la captura de la imagen anterior. La imagen resultante  $I_p$  es binarizada teniendo como umbral la máxima intensidad de la primera imagen. Una imagen resultante totalmente negra indica la ausencia de movimiento, mientras que la aparición de píxeles blancos indica el parpadeo.

$$I_p = I[i] - I[i-1] \quad (4)$$

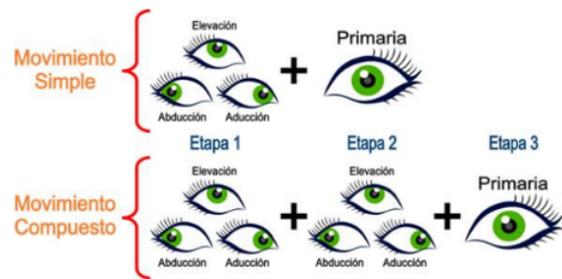


Figura 9. Etapas de los movimientos simple o compuesto

1. Link: funcionamiento del prototipo: <https://youtu.be/3i5ZNgQ1Hvw>

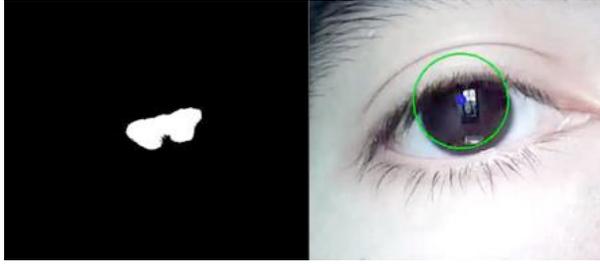


Figura 10. Detección del parpadeo a partir de la comparación de capturas en video.

La velocidad de cierre palpebral cambia según el tipo de parpadeo, el parpadeo reflejo demora 30ms aproximadamente y un parpadeo espontáneo alrededor de 75ms [18]. El prototipo realiza capturas continuas de 100ms, esto suprime cualquier parpadeo que sea involuntario permitiendo la activación del prototipo cuando el usuario lo desee (Ver Figura 10).

Se identificó el parpadeo para activar y/o desactivar la segmentación del iris, con la finalidad de suspender el sintetizador de no cuando el usuario no desee “hablar”. Este proceso se ejecuta cuando el usuario realiza 3 a 4 parpadeos voluntarios consecutivos.

## V. ANALISIS DE RESULTADOS

Como ultima etapa o fase del proceso de investigación, se realizó un experimento para validar el correcto funcionamiento del algoritmo y el clasificador en diferentes condiciones de luz: homogénea alta, no homogénea y homogénea baja (Ver Figura 11). Se realizó la captura en tiempo real

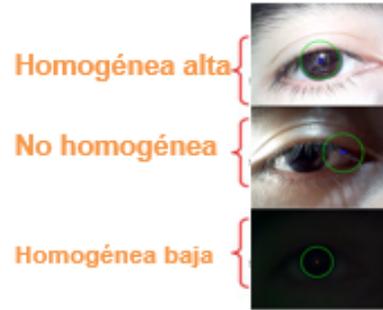
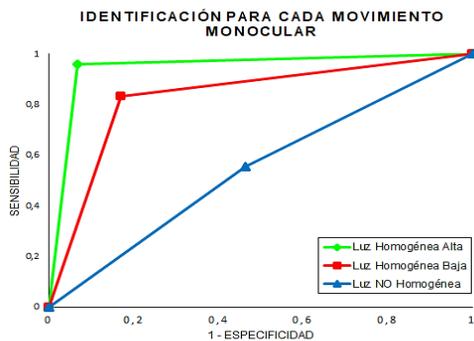


Figura 11. Condiciones de luz del experimento

del movimiento ocular de 12 voluntarios, durante la toma de datos se tuvo en cuenta el siguiente protocolo:

- Ubicar al voluntario en sedestación alternando las condiciones de luz.
- Encender el dispositivo y ubicar la cámara según las condiciones anatómicas del participante.
- Calibrar los movimientos oculares del participante, siguiendo las instrucciones dadas por el programa.
- Indicar al participante que realice 3 parpadeos consecutivos para encender el dispositivo.
- Indicar al voluntario la realización de los patrones y registrar los datos obtenidos.
- Brindar un tiempo de 5 minutos de descanso entre la toma de datos de cada movimiento para evitar la fatiga ocular o si el voluntario expresa necesitarlo.



	Luz H.A.		Luz H.B.		Luz N.H.	
↑	115	16	100	41	51	122
←	224	5	199	20	98	59
→						
	<b>True</b>	<b>False</b>	<b>True</b>	<b>False</b>	<b>True</b>	<b>False</b>
	SE =	95.8%	SE =	83.3%	SE =	46.4%
	ES =	93.3%	ES =	82.9%	ES =	44.5%

Figura 12. Curva ROC para la detección de los movimientos monoculares según valores

1. Link: funcionamiento del prototipo: <https://youtu.be/3i5ZNgQ1Hvw>

Dentro de los factores de inclusión para los sujetos de prueba se tuvo presente:

\* Personas entre 18 y 55 años de edad, con facultades y autonomía de participar en la investigación sin necesidad de un asentimiento.

\* Personas con buen estado de salud físico y mental, sin antecedentes de patologías oculares o de patologías que involucren la pérdida del control autónomo del movimiento ocular.

Para verificar la efectividad del algoritmo de rastreo se realizó la captura de 10 datos por cada movimiento monocular en tres condiciones de luz (Ver Figura 11): homogénea alta (HA), homogénea baja (HB) y no homogénea (NH), donde se obtuvo la curva de ROC de la Figura 12, evidenciando que las capturas en buenas condiciones de luz tienen un mejor resultado que con bajas condiciones de luz. A partir de esta premisa se identificó que en las imágenes con poca luz se detecta el 80% de los movimientos debido a que la escasez de luz genera sectores más oscuros en los costados laterales del ojo, por esta razón en algunas ocasiones la detección en estas direcciones varía hacia el costado contrario. En la Figura 12 se muestran los porcentajes de sensibilidad acorde a las condiciones de luz según el experimento mencionado anteriormente.

De acuerdo con el comportamiento de cada etapa que compone los movimientos simples y compuestos, se procede a determinar la eficacia del clasificador por medio de una curva ROC y la sensibilidad. Se puede identificar que durante las condiciones de luz adecuadas como la homogénea alta, la identificación de cada uno de los patrones se realiza correctamente; algo similar ocurre en condiciones de luz homogénea baja, donde los resultados para cada movimiento no son inferiores a un 70% de efectividad. Contrariamente ocurre en luz no homogénea, donde, la presencia de sombras genera alteraciones en la segmentación del iris.

En los resultados obtenidos en la luz no homogénea se puede observar que para los movimientos que involucran la aducción, el área bajo la curva disminuye considerablemente, mientras que para la abducción aumenta. Se infiere que este comportamiento se debe a que la sombra en la imagen durante la recolección de los datos se encontraba al costado derecho; es decir que las sombras favorecen la detección del movimiento del

Tabla 5. Asertividad del algoritmo para cada palabra

Palabra	LUZ		
	Homogénea Alta	Homogénea Baja	No Homogénea
Baño	96%	83%	46%
Si	92%	82%	33%
No	94%	84%	56%
Hambre	86%	79%	50%
Sed	91%	75%	43%
Dolor	85%	63%	37%
Hola	92%	74%	42%
Gracias	94%	74%	48%
Jajaja	85%	74%	48%

iris cuando se encuentra en la misma dirección de la sombra, mientras que se altera el mismo proceso para el lado contrario. La clasificación del movimiento se determina por el contraste del iris con las demás estructuras en la imagen, sin embargo, las sombras causan falsos positivos durante la clasificación. Al detectar las palabras “jajaja” y “hola”, a pesar de que existe una luz homogénea baja, la detección es inferior a las demás palabras en general, debido a que estos movimientos pasan por el centro y pueden confundirse con un movimiento simple.

En algunos movimientos compuestos se evidencia un menor porcentaje de asertividad en comparación con los movimientos simples (Ver Tabla 5). La causa de estos errores se debe a que el usuario realiza los cambios de movimiento rápidamente; mientras el dispositivo toma 10 capturas por segundo y además requiere identificar 3 veces la misma posición para definir la ubicación del iris, de tal manera que el algoritmo no alcanza a percibir la totalidad de las posiciones por las que el ojo se desplazó. Aproximadamente el ojo debe permanecer de 1 a 2 segundos en cada posición para una óptima detección.

#### A. Desviación angular de cada etapa

Para determinar la correcta clasificación de cada movimiento monocular ante diferentes condiciones de luz: homogénea alta, homogénea baja y no homogénea, se tomaron 10 muestras de cada movimiento ocular y se calculó la desviación angular el error absoluto con base en la ecuación (8).

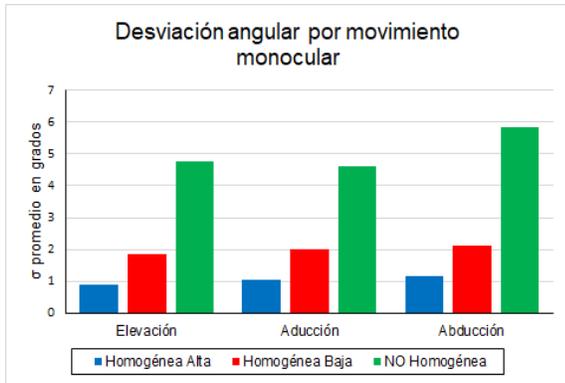


Figura 13. Desviación angular según condiciones de luz.

$$E = |V_{Real} - V_{Aproximado}| \quad (8)$$

Como se puede observar en la Tabla 6, se obtienen desviaciones angulares promedio menores a 1.5 grados para cada movimiento ocular y los datos correspondientes al promedio de grados en cada movimiento durante la calibración y el error promedio de las muestras. Una vez garantizada la precisión en la captura de los datos se evaluó el comportamiento del clasificador.

Teniendo en cuenta el promedio angular de las desviaciones de cada movimiento como referencia se determina la diferencia entre el valor medido y el valor de referencia obteniendo un error absoluto. Como se evidencia en la Tabla 5, la precisión del clasificador se mantuvo más estable en condiciones de luz homogénea comparado a las muestras con sombra en la imagen; es decir, que el prototipo pierde precisión en la detección del iris cuando aparecen sombras de la imagen.

Sin embargo, cabe resaltar de los resultados obtenidos, que a pesar de la pérdida de precisión evidenciada en el error absoluto, aún se garantiza la correcta clasificación de patrones en condiciones de luz homogénea alta (Ver Figura 13), esto se debe al margen máximo establecido por el clasificador mediante el hiperplano lineal del SVM.

Tabla 6. Posición Angula de los movimientos

Parámetro	Movimiento simple			
	Elevación	Abducción	Aducción	
Ángulo promedio	90.5893	3.1384	176.2384	
Desviación	0.8841	1.0401	1.1775	
Error absoluto	H.A	3.3747	2.8146	3.3936
	H.B	7.4181	6.8784	7.0965
	N.H	8.5593	7.9474	8.4889

### B. Detección del parpadeo

La detección del parpadeo tiene en cuenta el contraste de las estructuras oscuras en la imagen y la velocidad de captura. Por esta razón, se realizó el experimento tomando 10 muestras de la ejecución del parpadeo en tres condiciones de luz: homogénea alta, homogénea baja y no homogénea.

En condiciones de luz homogénea, se detecta el 60% de aciertos en el parpadeo, sin embargo, aunque existe un contraste entre la estructura del iris con el resto de la imagen, la velocidad de la captura afectó la detección acertada del parpadeo. En condiciones de luz no homogénea se detecta únicamente el 10% de aciertos, dado a que las variaciones de luz en la imagen impiden la correcta detección del iris, por lo cual no se puede evidenciar una comparación segura de la imagen presente con la anterior. Por último, en condiciones de luz baja y homogénea, se determina un 40% de aciertos en la detección, esto implica que la posibilidad de detectar el parpadeo es incierta (Ver Tabla 7). Se establece que las condiciones homogéneas de luz alta garantizan mayor probabilidad en la detección del ojo abierto en comparación con el ojo cerrado; sin embargo, depende de la velocidad de captura y velocidad de parpadeo del paciente. Por esta razón se da un intervalo de 5 segundos, para que la persona pueda realizar el parpadeo asegurando la detección para el encendido y apagado del sintetizador de voz.

**Tabla 7. Resumen de resultados en la detección del parpadeo.**

Resultado	Condición de luz		
	H.A.	H.B.	N.H
<b>Acierto</b>	61.67%	40%	40%
<b>Error</b>	38.33%	60%	60%

## VI. CONCLUSIONES

En el estudio realizado se resalta la importancia de proveer un dispositivo ergonómico y ajustable que garantice su correcto funcionamiento en diferentes entornos. El prototipo brinda una herramienta portátil, económica y asequible para cualquier persona. Su costo aproximado ronda los \$110.000 pesos colombianos.

En relación a su construcción, la movilidad de la cámara afecta la calibración inicial del prototipo, por ende, se dispuso de un soporte fijo para mejorar la precisión; además, el peso del prototipo puede causar fatiga al usuario durante el uso prolongado, razón que llevó a usar materiales de baja densidad y alta dureza. Igualmente, el diseño con gorra es ajustable a cualquier anatomía craneal, el sistema de soporte permite ajustar en el eje X, Y y Z la cámara, la visera mantiene fija la cámara sin importar el peso de la misma y la fuente de iluminación permite mantener condiciones de luz controladas. Se resalta un porcentaje de acierto entre el 85% y 96% para un tipo de luz homogénea alta, es decir, con alta incidencia de luz y sin sombras en el contorno del ojo.

El usuario tiene la libertad de suspender o activar el prototipo cuando necesite descansar o realizar otro movimiento que no está dentro de los patrones establecidos, de este modo, evitará la reproducción de palabras que no desea expresar. Este proceso se realiza a partir de la detección consecutiva de tres parpadeos en un lapso de 5 segundos. Sin embargo, el algoritmo implementado para la detección del parpadeo presenta opciones de mejora porque tiene un porcentaje de asertividad inferior al 70%.

Finalmente, se concluye que las condiciones de luz necesarias para la detección del movimiento ocular en este prototipo son las fuentes de luz uniformes, las cuales evitan la formación de sombras. Además, la luz homogénea en la imagen, captura una

distribución más amplia de tonos en el campo de color HSV, mejorando la calidad de la segmentación.

Un dispositivo de estas características podría contribuir positivamente en la comunicación verbal entre un paciente con enfermedades como esclerosis lateral amiotrófica o síndrome de enclaustramiento y su cuidador porque el paciente sería capaz de reproducir voluntariamente las palabras que desea expresar. Además, puede ser utilizado en otras herramientas de rastreo ocular como el control de prótesis biónicas u otros artefactos que realicen movimientos simples.

## AGRADECIMIENTO

Agradecemos a nuestras familias por su apoyo permanente ya nuestro tutor de tesis Jorge Alvarez Triana quien nos guio en este proceso de aprendizaje, atendió nuestras inquietudes y compartió sus conocimientos. Finalmente agradecemos a Dios por brindarnos la sabiduría para llevar a cabo este proyecto y por darnos salud y bienestar durante este tiempo.

## REFERENCIAS

- [1] A. Sánchez & E. Planells, “Necesidad de una evaluación e intervención en esclerosis lateral amiotrófica,” *Revista de Logopedia, Foniatría y audiología*. vol.38, no. 4, pp. 174-183, Sep., 2018.
- [2] F. Orient, R. Terré, D. Guevara y M. Bernabeu, “Tratamiento neurorrehabilitador de la esclerosis lateral amiotrófica”. *Rev Neurol.*, vol. 43, no. 9, pp. 549-555, 2006.
- [3] R. Spataro, M. Ciriaco, C. Manno, y V. La Bella, “The eye-tracking computer device for communication in amyotrophic lateral sclerosis,” *Acta Neurologica Scandinavica*, vol. 130, no. 1, pp. 40-45, 2014.
- [4] R.J. O’Connor, S. Cano, L. Ramió, A. Thompson y E. Playford, “Factors influencing retention for people with multiple sclerosis: Cross-sectional studies using qualitative and quantitative methods,” *Journal of Neurology*, no. 252, pp. 892-896, Ago., 2005.
- [5] MC. Puell Marín, *Óptica fisiológica: el sistema óptico del ojo y visión binocular*, Madrid, España: Unv. Complutense de Madrid, 2006, pp. 188-193.
- [6] WK. Pratt, *Digital Image Processing: PIKS Inside*. 3er ed. Nueva York, Estados Unidos: John Wiley and Sons Inc., 2001, pp. 89-158.
- [7] S. Muniyappan, A. Allirani and S. Saraswathi, "A novel approach for image enhancement by using contrast limited adaptive histogram equalization method," 2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT), Tiruchengode, India, pp. 1-6, Jul., 2013, doi: 10.1109/ICCCNT.2013.6726470.
- [8] N. Jamil, S. Tengku Mohd y ZA. Bakar, "Noise removal and enhancement of binary images using morphological

- operations," 2008 International Symposium on Information Technology, Kuala Lumpur, Malaysia, pp. 1-6, 2018, doi: 10.1109/ITSIM.2008.4631954.
- [9] HA. Hassan, NM. Tahir, I. Yassin, A. Zabidi, CHC. Yahaya y S. M. Shafie, "Automated Optic Disc removal in fundus images using iterative heuristics and morphological operations," 2013 IEEE Conference on Systems, Process & Control (ICSPC), Kuala Lumpur, pp. 230-233, 2013, doi: 10.1109/SPC.2013.6735137.
- [10] K. Lochan, P. Sah y KK. Sarma, "Innovative feature set for retinopathic analysis of diabetes and its detection," 2012 3rd National Conference on Emerging Trends and Applications in Computer Science, Shillong, pp. 240-245, 2012, doi: 10.1109/NCETACS.2012.6203267.
- [11] RK. Sidhu, "Improved canny edge detector in various color spaces," Proceedings of 3rd International Conference on Reliability, Infocom Technologies and Optimization, Noida, pp. 1-6, 2014, doi: 10.1109/ICRITO.2014.7014744.
- [12] Xiaochuan Zhao, Peizhi Liu, Min Zhang y Xinxin Zhao, "A novel line detection algorithm in images based on improved Hough Transform and wavelet lifting transform," 2010 IEEE International Conference on Information Theory and Information Security, Beijing, pp. 767-771, 2010.
- [13] O. Chapelle, P. Haffner y VN. Vapnik, "Support vector machines for histogram-based image classification," in IEEE Transactions on Neural Networks, vol. 10, no. 5, pp. 1055-1064, Sep. 1999, doi: 10.1109/72.788646.
- [14] MA. Mora, JD. Bernal y JE. Paneso. "Anatomía quirúrgica del ojo: Revisión anatómica del ojo humano y comparación con el ojo porcino," *Morfología*, vol. 8, no. 3, pp. 21-44, 2016.
- [14] A. López, E. Valveny y M. Vanrel (2019) "Detección de objetos". [Internet]. Disponible en <https://www.coursera.org/learn/deteccion-objetos/home/welcome>. [Accedido: 20-ene-2020]
- [15] J. Daugman, "How iris recognition works," IEEE Transactions on circuits and systems for video technology, vol. 14, no. 1, pp. 21-30, Ene., 2004.
- [16] P. Gil, F. Torres y FG. Ortiz, "Detección de objetos por segmentación multinivel combinada de espacios de color," *XXV Jornadas de Automática*, Ciudad Real: Universidad de Alicante, Sep., 2004.
- [17] WK. Pratt, *Digital Image Processing: PIKS Inside*. 3er ed. Nueva York, Estados Unidos: John Wiley and Sons Inc., 2001, pp. 89-158.
- [18] AT. Bahill, MR. Clark y L. Stark, "Computer simulation of overshoot in saccadic eye movements," *Comput Programs Biomed*, vol. 4, no. 4, pp. 230-236, Ago., 1975, doi: [https://doi.org/10.1016/0010-468X\(75\)90036-7](https://doi.org/10.1016/0010-468X(75)90036-7).

## **Autores**

### **Diego Alejandro Sánchez Alfonso**

Ingeniero biomédico Universidad Manuela Beltrán, Colombia.

### **Neiry Dayán Martínez Lara**

Ingeniera biomédica Universidad Manuela Beltrán, Colombia.

### **Andrés David Pérez Rodríguez**

Ingeniero biomédico Universidad Manuela Beltrán, Colombia.

### **Jorge Andrés Álvarez Triana**

Master en ingeniería Biomédica, Universidad Nacional de Colombia e Ingeniero Biomédico, Universidad Antonio Nariño, Colombia.