

La investigación reportada en esta tesis es parte de los programas de investigación del CICESE (Centro de Investigación Científica y de Educación Superior de Ensenada, Baja California).

La investigación fue financiada por el CONAHCYT (Consejo Nacional de Humanidades, Ciencias y Tecnologías).

Todo el material contenido en esta tesis está protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México). El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo o titular de los Derechos de Autor.

CICESE © 2023, Todos los Derechos Reservados, CICESE

Centro de Investigación Científica y de Educación Superior de Ensenada, Baja California



Maestría en Ciencias en Tecnologías Avanzadas e Integradas

Audiometría automatizada mediante visión computacional

Tesis

para cubrir parcialmente los requisitos necesarios para obtener el grado de
Maestro en Ciencias

Presenta:

Fernando Serrano Monroy

Ensenada, Baja California, México

2023

Tesis defendida por

Fernando Serrano Monroy

y aprobada por el siguiente Comité

Dr. Jesús Emmanuel Gómez Correa

Codirector de tesis

Dra. Ana Laura Padilla Ortiz

Codirectora de tesis

Dra. Rufina Hernández Martínez

Dr. Gabriel Alejandro Galaviz Mosqueda

Dr. Víctor Ruiz Cortés



Dr. Víctor Manuel Coello Cárdenas

Coordinador del Posgrado en Tecnologías Avanzadas e Integradas

Dra. Ana Denise Re Araujo

Directora de Estudios de Posgrado

Resumen de la tesis que presenta Fernando Serrano Monroy como requisito parcial para la obtención del grado de Maestro en Ciencias en Tecnologías Avanzadas e Integradas.

Audiometría automatizada mediante visión computacional

Resumen aprobado por:

Dr. Jesús Emmanuel Gómez Correa

Codirector de tesis

Dra. Ana Laura Padilla Ortiz

Codirectora de tesis

La pérdida de audición se refiere a la disminución de la capacidad auditiva, éste es un proceso progresivo e irreversible, en el que la mayoría de las ocasiones el individuo que presenta pérdida no se da cuenta de su evolución hasta que el proceso se encuentra avanzado. Dado que la audición es la vía principal a través de la cual se desarrolla el lenguaje y el habla, se debe tener presente que cualquier trastorno en la percepción auditiva de un individuo en edades tempranas, afectará su desarrollo lingüístico y comunicativo, sus procesos cognitivos y consecuentemente su posterior integración escolar, social y laboral. Todas estas consideraciones explican la necesidad de detectar este problema en sus comienzos, incluso antes de que el individuo se percate de que existe. La audiometría es una exploración del sistema auditivo, en la cual se mide la capacidad del oído humano de percibir sonidos. El objetivo principal de esta tesis es el diseño y desarrollo de un audiómetro automatizado, que con el reconocimiento de imágenes pueda interpretar la respuesta de la persona al realizar evaluación audiométrica, Para ello, se desarrolló un sistema de visión computacional. Se presentan algunos beneficios del audiómetro digital, tales como la realización de pruebas de manera estandarizada para cada persona, la facilidad de almacenamiento de los resultados en formato digital, y la ampliación del acceso a las pruebas auditivas en lugares remotas.

Palabras clave: audición, audiometría tonal, visión computacional, reconocimiento de patrones, salud auditiva

Abstract of the thesis presented by Fernando Serrano Monroy as a partial requirement to obtain the Master of Science degree in Advanced and Integrated Technologies.

Automated audiometry through computer vision

Abstract approved by:

Dr. Jesús Emmanuel Gómez Correa

Thesis Co-Director

Dra. Ana Laura Padilla Ortiz

Thesis Co-Director

Hearing loss refers to the decrease in auditory capacity, which is a progressive and irreversible process. In most cases, individuals with hearing loss are unaware of its progression until it has advanced. Since hearing is the main pathway through which language and speech development, it is essential to consider that any impairment in auditory perception during the early stages of life will affect linguistic and communicative development, cognitive processes, and subsequently, their educational, social, and occupational integration. All these considerations highlight the need to detect this problem early on, even before the individual becomes aware of its existence. Audiometry examines the auditory system that measures the human ear's ability to perceive sounds. The main objective of this thesis is the design and development of an automated audiometer that, through image recognition, can interpret a person's response during audiological evaluations. To achieve this, a computer vision system was developed. Some benefits of the digital audiometer are highlighted, such as conducting standardized tests for each individual, ease of storing results in digital format, and expanding access to auditory tests in remote locations.

Keywords: hearing, tonal audiometry, computer vision, pattern recognition, hearing health

Dedicatoria

A mis padres Hernán y Alejandra por brindarme todo el apoyo de mi formación académica.

Agradecimientos

A mi familia por apoyarme siempre en la medida de lo posible.

Al Centro de Investigación Científica y de Educación Superior de Ensenada, Baja California por brindarme la oportunidad de realizar mis estudios de posgrado.

Al Consejo Nacional de Humanidades, Ciencias y Tecnologías (CONAHCyT) por brindarme el apoyo económico para realizar mis estudios de maestría (CVU: 1135265).

A mis directores de tesis, el Dr. Jesús Emmanuel Gómez Correa y la Dra. Ana Laura Padilla Ortiz, por su constante y paciente orientación.

A mi comité de tesis, Dra. Rufina Hernández Martínez, Dr. Gabriel Alejandro Galaviz Mosqueda, y al Dr. Víctor Ruiz Cortés por su apoyo y diferentes puntos de vista de mi trabajo de tesis.

A mis amigos Ana, Owen y Fernando por siempre estar dispuestos a escucharme, incluso a la distancia.

A mis compañeros de maestría por haberme acompañado en este camino de aprendizaje, por haber compartido grandes experiencias y buenos momentos.

Tabla de contenido

	Página
Resumen en español	ii
Resumen en inglés	iii
Dedicatoria	iv
Agradecimientos	v
Lista de figuras	viii
Lista de tablas	x
Capítulo 1. Introducción	
1.1. Antecedentes	1
1.2. Justificación	3
1.3. Hipótesis	3
1.4. Objetivos de la investigación	3
1.4.1. Objetivos específicos	3
Capítulo 2. Sistema de visión computacional	
2.1. Etapas del sistema de visión	4
2.2. Adquisición de datos	5
2.3. Técnicas del reconocimiento de patrones	6
2.4. Espacio de color	6
2.4.1. Espacio RGB	7
2.4.2. Modelo HSV	8
2.4.3. Modelos YCbCr y CIELab	9
2.5. Segmentación	10
2.5.1. Filtro HSV	12
2.6. Transformada de Fourier	12
2.7. Clasificación de imágenes	14
2.8. Métricas de desempeño	15
Capítulo 3. Audiómetros digitales	
3.1. Sonido	16
3.1.1. Ondas de sonido	16
3.1.2. Características de una onda	17
3.2. Curvas isofónicas	18
3.3. Audiómetro	19
3.3.1. Características de un audiómetro	20
3.4. Audiometría tonal	21
3.5. Audiograma	22
3.6. Método de Hughson-Westlake	24
3.7. Calibración	26
3.7.1. Intensidad de sonido	26
3.7.2. Método de calibración	27

Capítulo 4. Resultados

4.1.	Sistema de visión computacional	29
4.1.1.	Procesamiento de las imágenes	29
4.1.2.	Clasificación	32
4.1.3.	Evaluación	33
4.2.	Desarrollo del audiómetro automatizado	34
4.2.1.	Implementación	35
4.2.1.1.	Generación de tonos	35
4.3.	Interfaz de Usuario	37
4.4.	Pruebas objetivas	40
4.4.1.	Error absoluto	41
4.5.	Audiometrías	42

Capítulo 5. Discusión**Capítulo 6. Conclusiones**

6.1.	Limitaciones del estudio	48
6.2.	Trabajo a futuro	48

Literatura citada	49
------------------------------------	----

Lista de figuras

Figura	Página
1. Diagrama a bloques del sistema de visión.	5
2. Representación gráfica del modelo de color RGB, los ejes R (rojo), G (verde) y B (azul), tienen valores desde 0 a 255.	7
3. Representación gráfica del modelo de color HSV.	8
4. Imagen original previo al proceso de segmentación.	10
5. Imágenes segmentadas por color. a) rojo, b) verde, c) morado, d) magenta, e) amarillo y f) los demás colores que no han sido segmentados.	11
6. Matriz de confusión.	15
7. Características de una onda de sonido, se muestra una onda con $A = 1$ y $\lambda = 0.05$, tomando la velocidad de una onda de sonido como $c = 343m/s$, da como resultado una onda de frecuencia de 6860 Hz.	18
8. Curvas isofónicas por la norma ISO226, cada curva expresa un mismo nivel de sonoridad medido en fonios.	19
9. Representación de una audiometría (Salud auditiva integral, 2018).	21
10. Ejemplo de un audiograma. Los círculos rojos representan las mediciones para el oído derecho, mientras que las cruces azules representan las mediciones para el oído izquierdo.	23
11. Clasificación de la pérdida auditiva relacionada al umbral auditivo.	23
12. Proceso del método de Hughson-Westlake modificado, implementado para encontrar los umbrales auditivos.	25
13. Proceso para el reconocimiento del patrón de la mano, la imagen A e imagen B son las imágenes de referencia y la imagen objetivo es la que se quiere clasificar.	30
14. Segmentación de tonos color piel utilizando el filtro HSV. En (a) se muestra la imagen sin filtros y en (b) la imagen al pasar por el filtro HSV. Ambas imágenes con 320x240 píxeles.	31
15. Transformada de Fourier de la mano segmentada en la figura 14(b).	31
16. Imágenes tomadas de la base de datos de manos, utilizadas para evaluar el sistema de reconocimiento de patrones, donde a, b y c imágenes de la mano abierta y d imagen de la mano cerrada.	33
17. Matrices de confusión, se evaluó el modelo usando 3 imágenes diferentes de la mano abierta y una imagen de la mano cerrada, a) utiliza la imagen (a) y (d), b) utiliza la imagen (b) y (d), c) utiliza la imagen (c) y (d), imágenes de la figura 15.	35
18. Diagrama de flujo general del proceso para automatizar el audiómetro digital.	36
19. Presentación de bienvenida.	38
20. Datos personales que la persona debe ingresar antes de la prueba.	38
21. Visualización de la etapa de familiarización.	38
22. Instrucciones durante la prueba.	39

Figura	Página
23. Indicación de que dará inicio a la prueba de audición.	39
24. Mensaje simultáneo al estímulo auditivo.	40
25. Interpolación de las mediciones para las frecuencias audiométricas, donde la amplitud tiene unidades arbitrarias (U.A).	41
26. Errores absolutos medido para diferentes frecuencias audiométricas.	42
27. Audiometría del participante AP. Los audiogramas (a) y (b) se realizaron con el audiómetro B&K modelo 1800. Los audiogramas (c) y (d) se realizaron con el audiómetro desarrollado en este proyecto.	43
28. La audiometría 1 fue realizado por el participante EG y la audiometría 2 por el participante RL	44

Lista de tablas

Tabla	Página
1. Propiedades de la transformada de Fourier.	14
2. Amplitudes iniciales para la calibración en unidades arbitrarias.	27
3. Valores del coeficiente de correlación para el caso de una imagen.	32
4. Coeficiente de correlación para el caso de una imagen con la mano abierta y una imagen sin la mano.	33
5. Coeficientes a y b, obtenidos a partir de la ecuación de amplitud.	41

Capítulo 1. Introducción

1.1. Antecedentes

El sentido de la audición es fundamental para percibir los sonidos a través del oído y desempeña un papel fundamental en diversos aspectos de nuestra vida. La audición nos permite comunicarnos, interactuar con otros y percibir el mundo que nos rodea. Nos brinda la capacidad de expresarnos, mantenernos seguros y comprender nuestro entorno. Por lo tanto, es de suma importancia cuidar y proteger nuestra capacidad auditiva para poder disfrutar plenamente de la vida y estar activamente involucrados en el mundo que nos rodea (Castañeda, 1999). Algunos aspectos importantes de la audición son:

- **Comunicación:** la audición es crucial para adquirir el lenguaje, ya que nos permite comunicarnos con los demás y comprender el lenguaje hablado.
- **Seguridad:** la audición nos permite detectar sonidos de alerta, como alarmas de incendio o el ruido de los automóviles, lo que nos ayuda a tomar decisiones para garantizar nuestra seguridad.
- **Orientación espacial:** la audición también nos ayuda a localizar fuentes sonoras, lo que implica una mejor conciencia espacial y percepción de nuestro entorno.

Se estima que en México 2.3 millones de personas padecen discapacidad auditiva. Cerca del 50 % son adultos mayores a 60 años, el 34 % tienen entre 30 y 59 años y el 2 % son niñas y niños (Instituto de Salud para el Bienestar, 2023). A partir de esta premisa, surge el interés de conocer las implicaciones de perder el sentido de la audición.

Algunas consecuencias de la pérdida auditiva son aislamiento social, soledad y depresión. Además, la tasa de desempleo es más alta en esta población. La pérdida auditiva afecta la calidad de vida de la persona que la padece.

Una de las causas de la disminución auditiva en los adultos es la exposición prolongada a sonidos muy fuertes, los cuales se pueden evitar. Otras causas pueden ser enfermedades crónicas, tabaquismo, otosclerosis o degeneración neurosensorial. También existen otras causas tales como un tapón de cerumen, medicamentos ototóxicos, productos químicos ototóxicos, carencia nutricional o infecciones virales, según datos proporcionados por la Organización Mundial de la Salud (2023). La detección oportuna de la pérdida auditiva es una estrategia que ayuda a disminuir o evitar las consecuencias que conlleva.

Para detectar una pérdida auditiva se realiza una audiometría tonal para medir los umbrales auditivos de cada oído. Según la definición de Durrant et al. (1985), el umbral auditivo es el volumen mínimo al cual una persona escucha determinadas frecuencias de un sonido. La prueba se realiza con un audiómetro, el cual genera diferentes tonos a frecuencias que pueden variar entre 250 y 8000 Hz, debido a que en ese intervalo se abarca las frecuencias del habla, que son las más importante para la comunicación. El proceso se repite varias veces para las diferentes frecuencias, con el fin de descartar cualquier error. Este es uno de varios métodos para realizar la exploración auditiva.

Durante la pandemia de COVID-19, se utilizó la telemedicina para ofrecer consultas a los pacientes, evitando así riesgos de contagio. Además, como menciona la Secretaría de Salud en México (2021), es importante brindar una mayor visibilidad a la población que padece problemas de audición.

Los avances tecnológicos recientes han permitido que la teleaudiología sea la primera opción para abordar la falta de servicios de atención médica auditiva y la escasez de profesionales cualificados en áreas remotas. La demanda en el campo de la audiología ha dado lugar a numerosos desarrollos de aplicaciones audiológicas, así como a la creación de audiómetros portátiles y pruebas de audición utilizando teléfonos inteligentes. Esto permite llegar a zonas donde no hay servicios médicos especializados y hacerlos accesibles para los pacientes, sin importar su situación económica (Cardier et al., 2016). Estos avances requieren sistemas automatizados capaces de reconocer e interpretar las respuestas de las personas durante una audiometría.

Por otro lado, la visión computacional es una disciplina que busca emular en las computadoras la capacidad de nuestros ojos para reconocer y analizar imágenes. A través de la interpretación de imágenes capturadas, como aquellas obtenidas mediante una cámara, se busca identificar los objetos presentes en el entorno (Sucar & Gómez, 2011). Esta disciplina ha sido ampliamente estudiada en diversas tareas, como la detección de objetos (Ouyang et al., 2017), el reconocimiento de rostros (Wright et al., 2010), la clasificación de imágenes (Mairal et al., 2008), el reconocimiento de poses humanas (Chen & Yuille, 2014) y la segmentación semántica (Long et al., 2015).

En este trabajo, se usó un sistema de visión computacional para medir la respuesta de una audiometría tonal automatizada, la respuesta será identificada con el gesto de la mano abierta o cerrada. Se presentará el diseño de un audiómetro automatizado que utiliza el reconocimiento de imágenes para identificar la respuesta de la persona que se realice la prueba audiométrica. Este audiómetro proporcionará información de los umbrales auditivos, que son fundamentales para evaluar el estado de salud auditiva de los individuos.

1.2. Justificación

La pérdida auditiva merma significativamente la calidad de vida de quien la padece. Por ello la necesidad de detectar este problema desde sus inicios, incluso antes de que el individuo se percate de que existe. Actualmente, para realizar una evaluación audiométrica es necesario acudir a alguna clínica, o bien, a algún consultorio particular. Implementar un sistema de visión computacional en un audiómetro digital, permitirá que este tipo de pruebas puedan realizarse de manera remota, es decir, sin la necesidad de que el paciente se desplace para sus evaluaciones periódicas.

1.3. Hipótesis

La implementación de un audiómetro digital automatizado que mediante visión computacional registre las respuestas del paciente en la prueba, hará posible realizar audiometrías. El sistema de visión computacional guiará a la persona durante toda la prueba.

1.4. Objetivos de la investigación

El objetivo general de este trabajo de investigación es desarrollar e implementar un audiómetro automatizado que, en complemento de un sistema de visión computacional, realice una evaluación de la salud auditiva sin la necesidad de una persona especializada en el área.

1.4.1. Objetivos específicos

- Identificar y seleccionar técnicas de reconocimiento de patrones.
- Generar un sistema de visión computacional que reconozca imágenes determinadas.
- Diseñar e implementar en MATLAB un audiómetro digital.
- Automatizar mediante visión computacional un audiómetro digital.
- Evaluar el audiómetro desarrollado.

Capítulo 2. Sistema de visión computacional

Un sistema de visión computacional trata de imitar la visión humana, cuyo propósito es dar una interpretación a imágenes digitales. Para ello, se requiere de un dispositivo que capture video o imágenes para ser procesados con un conjunto de algoritmos que permitan obtener de manera automática un resultado. Algunas aplicaciones que ha tenido son: caracterizar lesiones en la piel (Maglogiannis & Doukas, 2009), identificar de manera automática especies de plantas (Kumar et al., 2012), inspeccionar el ensamblaje en naves industriales (Kakani et al., 2020) o determinar el comportamiento de ciertos animales (Nasirahmadi et al., 2017). Estas tareas se pueden realizar utilizando reconocimiento de patrones y procesamiento digital de imágenes en el sistema (Liu et al., 2006).

Los seres humanos poseemos la habilidad natural de identificar objetos, formas y movimientos con facilidad, esto mismo puede lograrse mediante algoritmos de aprendizaje automático. Estos algoritmos se entrenan utilizando grandes conjuntos de datos, lo cual les permite reconocer patrones y obtener resultados precisos (Abburu & Golla, 2015).

El reconocimiento de patrones tiene un papel valioso, ya que con pocos datos es posible lograr resultados acertados. Además, el procesamiento digital de imágenes complementa al reconocimiento de patrones al proporcionar técnicas y algoritmos para mejorar la calidad de las imágenes. Un buen sistema de visión computacional requiere el reconocimiento de patrones, ya que es fundamental para comprender y procesar la información visual.

Los patrones se definen como similitudes compartidas por conjuntos de elementos u objetos, y su categorización se determina según las características que presentan, lo que permite agrupar objetos similares. El reconocimiento de patrones es una disciplina que emplea diversas técnicas para identificar y extraer patrones en los datos. (Sharma & Kaur, 2013).

2.1. Etapas del sistema de visión

El sistema de visión sigue varias etapas importantes. En primer lugar, la medición de una señal es crucial para recopilar los datos necesarios, esto puede incluir señales como la presión del sonido, la intensidad de la luz, la humedad o la temperatura, dependiendo del tipo de reconocimiento de patrones que se esté realizando.

Después de la adquisición de datos, viene la etapa de preprocesamiento. Aquí es donde los datos se procesan y manipulan para mejorar su calidad y facilitar la extracción de características importantes. Por ejemplo, en el procesamiento de imágenes, se puede escalar el tamaño de la imagen para reducir el tiempo de cálculo y mejorar la eficiencia del sistema de reconocimiento (Vera et al., 2015).

Una vez que se han preprocesado los datos, sigue la etapa de extracción de características. Esta etapa es crucial para identificar y resaltar las características clave que se utilizarán en la etapa final de, y se utilizan técnicas como la segmentación de imágenes para separar los objetos del fondo y resaltar las características importantes.

La clasificación es una etapa importante. Implica agrupar las características o patrones extraídos en diferentes categorías, lo que permite al sistema tomar decisiones. Una vez que las características han sido clasificadas, el sistema puede tomar decisiones basadas en esa información.

La Figura 1 muestra en un diagrama a bloques las etapas de un sistema de visión, las cuales se describen a continuación.

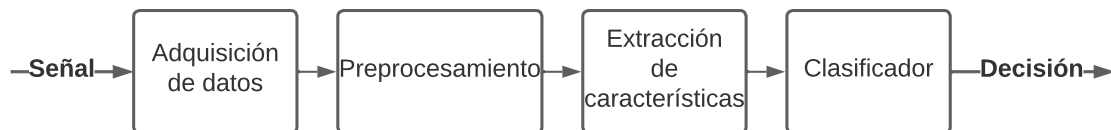


Figura 1. Diagrama a bloques del sistema de visión.

2.2. Adquisición de datos

La adquisición de datos mediante un sistema de visión computacional se refiere al proceso de capturar imágenes o videos utilizando dispositivos ópticos, como cámaras, y utilizar esos datos para realizar el análisis y la extracción de información mediante algoritmos y técnicas de procesamiento de imágenes. La adquisición de datos, se realiza principalmente capturando imágenes utilizando cámaras o sensores ópticos, es decir, se adquieren imágenes o secuencias de imágenes de la escena o el objeto que se desea analizar. Las cámaras pueden estar fijas o montadas en sistemas móviles, como drones o robots.

2.3. Técnicas del reconocimiento de patrones

Algunas de las técnicas comunes utilizadas en el reconocimiento de patrones son:

- **Extracción de características:** esta técnica desempeña un papel importante en el proceso de reconocimiento de patrones. Consiste en identificar y extraer características relevantes de los objetos, lo que permite obtener patrones de interés para su posterior reconocimiento o clasificación.
- **Segmentación de imagen:** consiste en delimitar áreas y objetos en una imagen, lo que facilita el análisis y reconocimiento de patrones en cada segmento (Kuruvilla et al., 2016).
- **Modelos de redes neuronales:** las redes neuronales son una técnica ampliamente utilizada en el campo de la inteligencia artificial. Se trata de sistemas que, mediante el procesamiento de datos de entrenamiento, son capaces de aprender y resolver problemas gracias a su capacidad de adquirir conocimiento previo (Sharma & Kaur, 2013).
- **Clasificación supervisada:** en esta técnica, los algoritmos de clasificación comunes incluyen máquinas de soporte vectorial (SVM por sus siglas en inglés), árboles de decisión y árboles aleatorios (Nasteski, 2017).

Usando una o la combinación de varias técnicas, se logra alcanzar el objetivo del reconocimiento en diversas áreas con un alto porcentaje de identificación.

2.4. Espacio de color

Un espacio o modelo de color es una forma de representar y organizar los colores de manera sistemática. Es un modelo matemático que define un conjunto de colores y proporciona un sistema de coordenadas para especificar cada color dentro de ese espacio.

Existen diferentes espacios de color, cada uno con su propia forma de representar los colores como lo veremos más adelante. Sin embargo, cada espacio de color tiene sus propias ventajas y aplicaciones específicas y la elección del espacio de color depende del contexto y los requisitos del proyecto en particular. Al convertir entre diferentes espacios de color, se pueden realizar ajustes y adaptaciones para garantizar la representación más precisa y consistente de los colores.

2.4.1. Espacio RGB

Uno de los modelos de color más comunes es el modelo RGB (rojo, verde, azul, por sus siglas en inglés), que es un modelo de color aditivo. En este modelo, los colores se crean mediante la combinación de los colores primarios de la luz: rojo, verde y azul. Esta combinación puede producir más de 16 millones de colores diferentes (Fairman et al., 1997). En la representación gráfica de la Figura 2, se muestran todos los posibles colores dentro de este modelo RGB. Cada punto en el gráfico representa un color específico que se puede lograr ajustando los valores de los parámetros rojo, verde y azul. Los colores más oscuros y cercanos al origen del gráfico representan combinaciones con valores de parámetros más bajos, mientras que los colores más claros y cercanos a los bordes del gráfico representan combinaciones con valores de parámetros más altos.

El modelo RGB se deriva de la relación entre las longitudes de onda de la luz y los colores que el ojo humano percibe. Además, se basa en la síntesis aditiva de color de los colores primarios de la luz y se compone únicamente de tres colores que se podrán combinar entre ellos, cada uno puede tener valores de intensidad entre 0 y 255.

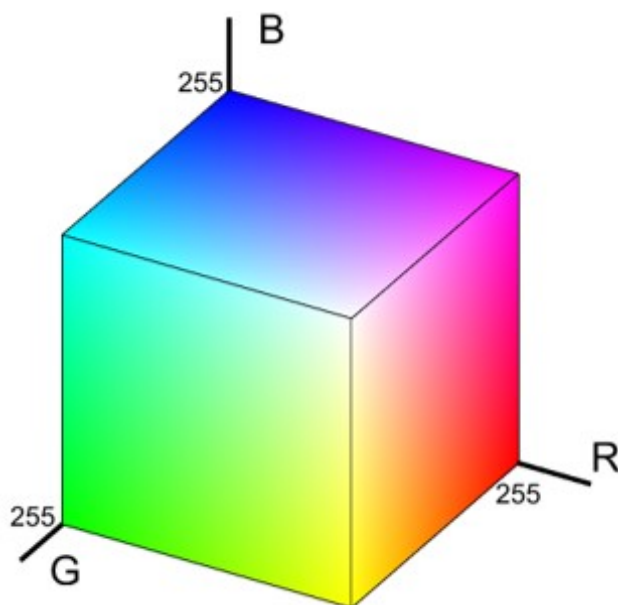


Figura 2. Representación gráfica del modelo de color RGB, los ejes R (rojo), G (verde) y B (azul), tienen valores desde 0 a 255.

2.4.2. Modelo HSV

Una forma de representar los colores es mediante el modelo HSV, por sus siglas en inglés "Hue" (tono), "Saturation" (saturación) y "Value" (luminancia). En este modelo, la variación del tono implica moverse en un círculo de colores, donde cada valor representa un color específico. La saturación se refiere a la opacidad o blancura del color, donde un valor de saturación completo indica un color puro sin mezcla con blanco. La saturación varía desde cero hasta uno, desplazándose desde el centro del círculo hacia afuera. Por último, la luminancia determina el nivel de oscuridad o brillo del color. El valor cero corresponde al color negro en la punta del cono, mientras que el valor uno se asocia con el color blanco, y su apariencia también se ve afectada por la saturación (Liu & Zheng, 2010).

El tono está relacionado principalmente con la longitud de onda de los colores espectrales. Se mide en grados y varía de 0 a 360° . Comenzando desde el rojo y avanzando en sentido de las manecillas del reloj hasta el azul, el tono abarca todo el espectro de colores.

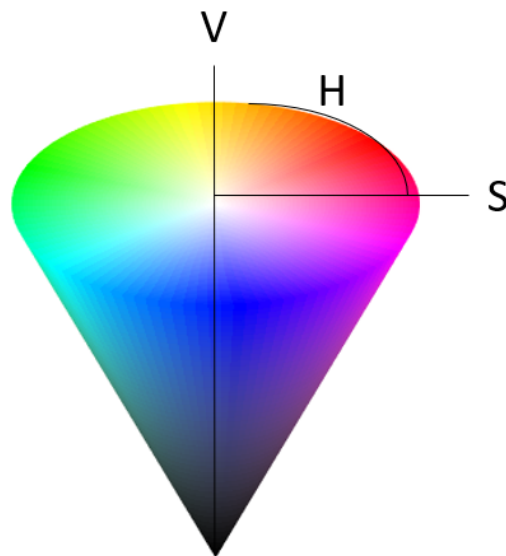


Figura 3. Representación gráfica del modelo de color HSV.

La Figura 3 representa una estructura gráfica en forma de cono para visualizar los colores del modelo HSV. Esta representación se relaciona con coordenadas cilíndricas, donde el tono del color varía el ángulo, la saturación afecta el radio y el valor determina la altura del cono. Además, se impone una restricción de que la saturación (S) siempre sea menor o igual a la luminancia (V).

Dado que cada pixel en la imagen tiene la información del color utilizando los canales RGB, utilizan las

siguientes transformaciones que relacionan el modelo RGB con el modelo HSV:

$$MAX = \text{máximo}\{R, G, B\}, \quad (1)$$

$$MIN = \text{mínimo}\{R, G, B\}, \quad (2)$$

$$H = \begin{cases} \text{indefinido} & \text{si } MAX = MIN \\ 60 \frac{G-B}{MAX-MIN} + 0 & \text{si } MAX = R \text{ y } G \geq B \\ 60 \frac{G-B}{MAX-MIN} + 360 & \text{si } MAX = R \text{ y } G < B \\ 60 \frac{B-R}{MAX-MIN} + 120 & \text{si } MAX = G \\ 60 \frac{R-G}{MAX-MIN} + 240 & \text{si } MAX = B \end{cases}, \quad (3)$$

$$S = \begin{cases} 0, & \text{si } MAX = 0 \\ 1 - \frac{MIN}{MAX}, & \text{otro caso} \end{cases}, \quad (4)$$

$$V = MAX, \quad (5)$$

donde MAX y MIN son los valores máximos y mínimos respectivamente entre los valores R , G y B , H es el tono, S la saturación y V es el brillo.

2.4.3. Modelos YCbCr y CIELab

En el modelo YCbCr, la información de la luminancia se representa mediante un único componente denominado Y , mientras que los parámetros que dan el color se almacena en los componentes Cb y Cr . Cb se refiere a la componente de crominancia azul. Por otro lado, Cr corresponde a la componente de crominancia roja (Phung et al., 2005).

CIELab es un sistema de color utilizado para representar todos los colores que el ojo humano puede percibir. La Commission Internationale d'Éclairage (CIE) es la entidad responsable de establecer los estándares de color, tales como CIEXYZ, CIEluv y CIELab. La componente L representa la luminancia variando entre el negro y el blanco, mientras que a y b son colores cromáticos, a es la coordenada que varía entre rojo y verde, y b la coordenada que varía entre los colores azul y amarillo.

Cada uno de los cuatro modelos de color mencionados anteriormente se clasifica en diferentes grupos. El modelo RGB se encuentra en el grupo de colores básicos, el modelo HSV en el grupo de colores perceptuales, el modelo YCbCr en el grupo de colores ortogonales y el modelo CIELab en el grupo de colores perceptualmente uniformes.

2.5. Segmentación

Una imagen digital es una representación de algo físico, que puede ser modelada a través de una función matemática discreta. La imagen tendrá la estructura de una matriz bidimensional y se compondrá de un pixel en cada espacio de la matriz. Cada pixel contendrá la información del color que representa, donde originalmente serán tres canales los cuales serán igual a los valores de los parámetros RGB. Un filtro es una herramienta utilizada para lograr la segmentación de imágenes, la cual implica la separación de las regiones de interés presentes en una imagen del fondo o de otras regiones. Uno de los métodos de segmentación es el de umbralización binaria, consiste en comparar cada pixel y clasificarlos dependiendo del valor de su intensidad comparada con un umbral (Yousefi, 2011). Si la intensidad del pixel se encuentra dentro del umbral significa que encontramos la región de interés, si es lo contrario, se le clasificará como el fondo.



Figura 4. Imagen original previo al proceso de segmentación.

La detección de color de piel es una técnica ampliamente utilizada en el campo de la visión computacional y el procesamiento de imágenes digitales. Se utiliza para reconocer gestos faciales (Kjeldsen & Kender, 1996) y corporales, así como para aplicaciones de interacción humano-computadora (Hongo et al., 2000).

Sin embargo, la detección de piel puede ser un desafío debido a las variaciones de iluminación y fondo en las imágenes. Por lo tanto, es importante elegir características relevantes para la detección de piel (Contreras & Gallegos, 2016). Los modelos de espacio de color son una de las herramientas más importantes, ya que cada modelo tendrá una sensibilidad diferente a los cambios de color.

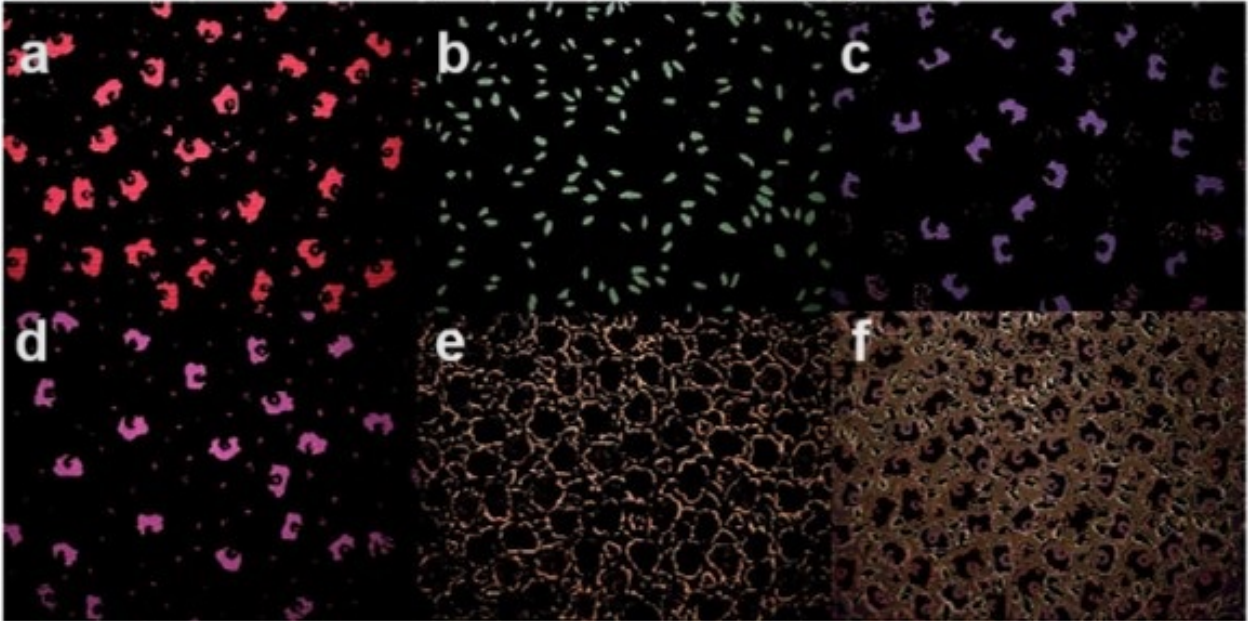


Figura 5. Imágenes segmentadas por color. a) rojo, b) verde, c) morado, d) magenta, e) amarillo y f) los demás colores que no han sido segmentados.

El proceso de la segmentación de la piel permite distinguir en una imagen los colores que se asemejen a los diferentes tonos de piel. La manera de segmentar un color es eligiendo valores para los componentes rojo, verde y azul e ir comparando cuál pixel tiene esta combinación. Dado que no hay un único tono de piel, para segmentar se debe elegir un umbral para cada componente de color. En la Figura 4 se muestra una imagen original que contiene una variedad de colores. A través de un algoritmo de segmentación, se seleccionan parámetros específicos para identificar y separar los diferentes tonos presentes en la imagen. Como resultado, se generan imágenes adicionales que resaltan los colores encontrados. Cada imagen de la Figura 5 contiene un grupo de colores semejantes, lo que permite distinguir de manera clara cada color. Se ha encontrado que segmentar tonos de piel con el modelo RGB no es preciso, ya que los objetos se ven afectados por la intensidad de la iluminación, principalmente. Por ello, se han utilizado diferentes modelos de color para segmentar la piel. Con umbrales del modelo HSV se ha demostrado que permite distinguir mejor los tonos de piel que con propuestas de umbrales del modelo RGB (Ammal et al., 2013).

2.5.1. Filtro HSV

Entre los modelos de espacio de color comúnmente utilizados para la segmentación de piel se encuentran el modelo de espacio de color HSV, YCrCb y CIELab. Estos modelos se utilizan porque separan los componentes de crominancia y luminancia, lo que permite una mejor caracterización de los diferentes tonos de piel (Phung et al., 2005).

En particular, el modelo HSV se utiliza ampliamente debido a su capacidad para separar la información de crominancia y luminancia. Además, este modelo es el que menos se ve afectado por la variación de iluminación en comparación con otros modelos de espacio de color.

Para lograr la detección del color de la piel en una imagen, se aplicó un umbral basado en los parámetros HSV descritos en el artículo de Contreras & Gallegos (2016). Esta técnica de filtrado es utilizada como método de segmentación.

$$0 < H < 0.2, \quad (6)$$

$$0.35 < S < 0.9, \quad (7)$$

$$0 < V < 1, \quad (8)$$

La saturación y el tono desempeñan un papel fundamental en la clasificación de los tonos de piel, mientras que el parámetro del brillo abarca toda su escala.

En una imagen en la que ya hemos segmentado el objeto de interés, no es suficiente para poder clasificarlo, ya que el objeto por sí solo no proporciona patrones fáciles de reconocer. Por lo tanto, es necesario implementar técnicas adicionales para extraer las características necesarias.

2.6. Transformada de Fourier

La Transformada de Fourier es una técnica muy utilizada en el procesamiento de imágenes con sus propias aplicaciones y ventajas. La transformada de Fourier es una operación matemática que transforma una función del dominio del tiempo o del espacio al dominio de la frecuencia, sin modificar la información

que contiene. La transformada de Fourier tiene aplicaciones en campos de la ciencia para análisis de señales (Tao et al., 2006), análisis de imágenes digitales para la implementación de filtros (Beaudoin & Beauchemin, 2002) y reconstrucción de imágenes (Schomberg & Timmer, 1995).

Tanto la segmentación mediante el filtro HSV como la transformada de Fourier son técnicas que se complementan. Después de aplicar el filtro HSV para segmentar la región de piel, se puede utilizar la transformada de Fourier para analizar las frecuencias presentes en dicha región. Esto permite identificar patrones específicos que resultan útiles para el reconocimiento de imágenes.

En la clasificación de imágenes, el objetivo es categorizar los objetos que contiene utilizando la información que la imagen pueda proporcionar para tomar una decisión. El reconocimiento de imágenes puede hacer uso de técnicas de aprendizaje automático ("machine learning") o aprendizaje profundo, así como la implementación de redes neuronales profundas. Una desventaja de las redes neuronales es que se requieren muchas imágenes para obtener buenos resultados (Tajbakhsh et al., 2016). El reconocimiento de patrones es una alternativa a las redes neuronales.

Una imagen digital puede analizarse desde su color, forma, tamaño o características como esquinas. Si un objeto en la imagen puede reconocerse fácilmente a través de alguna de estas características, no será necesario implementar una red neuronal que requiera un gran número de imágenes de entrenamiento. Otra característica que se puede obtener de una imagen es el patrón de la transformada de Fourier para obtener sus componentes en el dominio de la frecuencia.

La transformada de Fourier en dos dimensiones de una función $f(x, y)$ se puede definir como (Brigham, 1988):

$$F(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) e^{-i2\pi(ux+vy)} dx dy, \quad (9)$$

donde u y v son las variables de frecuencia y x y y las variables temporales o espaciales.

Para obtener la transformada inversa de Fourier es con la ecuación:

$$f(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(u, v) e^{i2\pi(xu+yv)} du dv, \quad (10)$$

Para aplicar la transformada de Fourier a una imagen digital se utiliza la transformada de Fourier discreta

para funciones con dos variables independientes.

$$F(u, v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-2i\pi(\frac{ux}{N} + \frac{vy}{M})}, \quad (11)$$

$$f(x, y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) e^{2i\pi(\frac{ux}{N} + \frac{vy}{M})}, \quad (12)$$

donde $F(u, v)$ es la transformada de Fourier discreta de una función $f(x, y)$ en dos dimensiones.

Se presentan algunas propiedades de la transformada de Fourier en la tabla 1.

Tabla 1. Propiedades de la transformada de Fourier.

Linealidad	$\mathbb{F}[af(x, y) + bg(x, y)] = aF(u, v) + bG(u, v)$
Cambio de escala	$\mathbb{F}[f(ax, by)] = \frac{1}{ ab } F(\frac{u}{a}, \frac{v}{b})$
Traslación	$\mathbb{F}[f(x - x_0, y - y_0)] = F(u, v) e^{-j2\pi(\frac{x_0 u}{M} + \frac{y_0 v}{N})}$
Convolución	$f(x, y) * h(x, y) = \frac{1}{MN} \sum_{m=1}^{M-1} \sum_{n=0}^{N-1} f(m, n) h(x - m, y - n)$

2.7. Clasificación de imágenes

En el contexto de la clasificación del sistema de reconocimiento de patrones, se puede aplicar la correlación para comparar las características extraídas de la imagen, con características previamente conocidas. La correlación es una técnica que relaciona a dos señales y determina la similitud que hay entre ellas. A la correlación de dos señales diferentes se le conoce como correlación cruzada y si son iguales es autocorrelación (Godfrey, 1980).

A través de la correlación, se establece un coeficiente de correlación (r) entre dos señales, $A(n, m)$ y $B(n, m)$, que representa la similitud entre ellas. Este coeficiente se calcula mediante la fórmula (Zhang et al., 2010):

$$r = \frac{\sum_m \sum_n (A(n, m) - \bar{A})(B(n, m) - \bar{B})}{\sqrt{(\sum_m \sum_n (A(n, m) - \bar{A})^2)(\sum_m \sum_n (B(n, m) - \bar{B})^2)}}, \quad (13)$$

Con el coeficiente de correlación, categorizamos la imagen según su similitud con los patrones de referencia, asignando la categoría que presente mayor similitud.

2.8. Métricas de desempeño

Una manera de evaluar cuán bueno son los resultados es a través de la matriz de confusión. Los valores obtenidos en la matriz permiten calcular las métricas de desempeño como exactitud, precisión y sensibilidad y están dadas por las siguientes expresiones (Borja-Robalino et al., 2020):

$$Exactitud = \frac{TP + TN}{TP + TN + FP + FN}, \quad (14)$$

$$Precisión = \frac{TP}{TP + FP}, \quad (15)$$

$$Sensibilidad = \frac{TP}{TP + FN}, \quad (16)$$

Las métricas son calculadas con el número de predicciones verdaderas positivas (TP), verdaderas negativas (TN), falsas positivas (FP) y falsas negativas (FN) del modelo (ver figura 6). La exactitud es el porcentaje de muestras que fueron clasificadas correctamente. La precisión es la proporción de predicciones verdaderas positivas entre el total de las predicciones positivas. La sensibilidad es la tasa de verdaderos positivos, es decir, la proporción de muestras positivas clasificadas correctamente.

Matriz de confusión

Predicción	si	TP	FP	Precisión
	no	FN	TN	
		Sensibilidad	Especificidad	Exactitud
		si	no	
		Valor real		

Figura 6. Matriz de confusión.

Capítulo 3. Audiómetros digitales

En este capítulo se revisan algunos conceptos básicos del sonido, como las ondas sonoras y sus características. También se presenta de forma breve la historia y evolución de los audiómetros, así como el funcionamiento de estos dispositivos. Finalmente, se presenta un método automatizado para determinar los umbrales auditivos.

3.1. Sonido

El estudio del sonido ha dado lugar a importantes desarrollos tecnológicos, que han logrado un gran impacto en mejorar la calidad de vida y al progreso en diferentes disciplinas, a continuación se mencionan algunos usos. Un ejemplo destacado es el uso de la navegación con sonido, conocida como SONAR (Neupane & Seok, 2020), que ha permitido la detección de objetos y la medición de distancias en equipos de los submarinos, posibilitando así navegaciones más seguras. Esta tecnología se basa en la investigación de animales que utilizan la ecolocalización para orientarse en el espacio y encontrar alimento en la oscuridad (Bassett & Eastmond, 1964).

En el campo de la medicina, los ultrasonidos han desempeñado un papel fundamental al permitir la obtención de imágenes internas del cuerpo (Chan & Perlas, 2011). Esta técnica no invasiva proporciona información detallada sobre órganos y tejidos. Asimismo, el estudio del sonido en los animales también ha resultado invaluable. Por ejemplo, los apicultores pueden analizar el espectro de sonido del zumbido que producen las abejas para obtener información sobre su estado de salud y bienestar (Ferrari et al., 2008).

3.1.1. Ondas de sonido

Las ondas sonoras son ondas mecánicas longitudinales que pueden desplazarse a través de medios elásticos como líquidos, gases o sólidos, lo que conocemos como sonido. Los sistemas que generan movimiento en el aire, como los instrumentos musicales, producen perturbaciones que se propagan en el aire y llegan a nuestros oídos. Allí, estas vibraciones mecánicas se transforman en señales eléctricas que llegan al cerebro (Resnick et al., 2005).

Si consideramos un tren continuo de compresiones y rarefacciones que se desplaza a lo largo de un tubo con un medio elástico, como el aire, cada elemento del fluido puede desplazarse en dirección hacia la derecha o la izquierda. Por convención, la dirección de propagación de la onda se toma como positiva en el eje x . En este caso, la ecuación que representa el movimiento de un elemento a partir de su posición de equilibrio en x en un tiempo t se denota como $s(x, t)$. Podemos expresar la ecuación del desplazamiento longitudinal como una onda sinusoidal:

$$s(x, t) = A * \sin(kx - wt), \quad (17)$$

donde A es la amplitud de la onda, k es el número de onda y w es la frecuencia angular.

3.1.2. Características de una onda

Una onda de sonido se compone de las siguientes características: longitud de onda, frecuencia y amplitud (Ver Figura 7).

Amplitud [A]: es la distancia entre el punto de equilibrio y la cresta de la onda.

Frecuencia [f]: es el número de ciclos en una unidad de tiempo, donde un ciclo es una oscilación completa de la onda. La frecuencia es una medida que determina el tono de un sonido. La frecuencia natural del sistema tiene la siguiente relación con la frecuencia angular $w = 2\pi f$.

Longitud de onda [λ]: es la distancia entre dos máximos o dos mínimos consecutivos de la onda. Se relaciona con la frecuencia natural a través de la relación $f = c/\lambda$, donde c es la velocidad de propagación.

Los tonos puros son ondas sinusoidales compuestas por una sola frecuencia natural f y un período de oscilación de $1/f$ segundos. Por ello, podemos variar su amplitud y frecuencia. La combinación de dos o más tonos puros resulta en sonidos complejos.

El espectro de un sonido se define como la distribución de la energía sonora en todas las frecuencias del rango audible. Al observar el espectro de un tono puro, se puede notar que la energía está distribuida únicamente en una sola frecuencia.

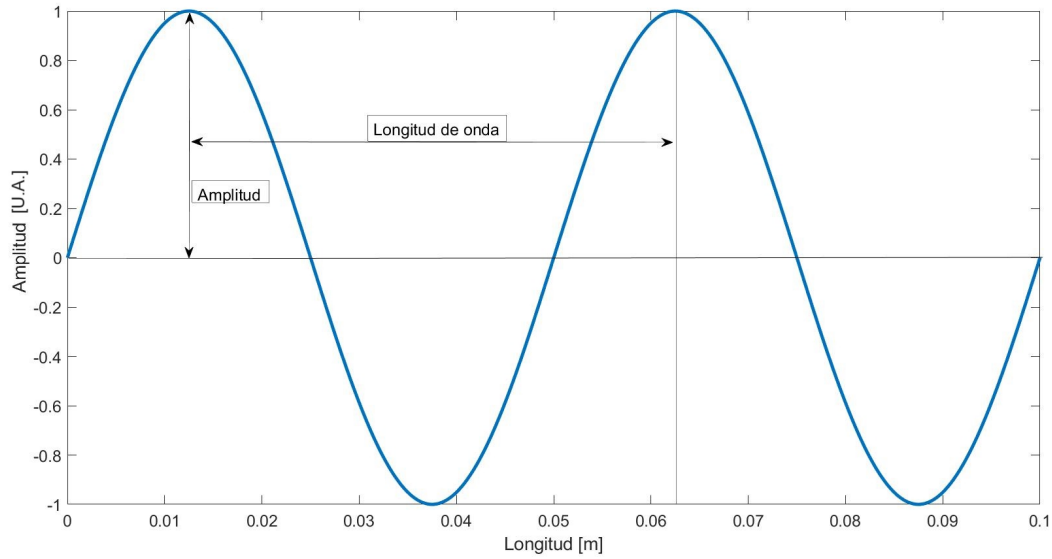


Figura 7. Características de una onda de sonido, se muestra una onda con $A = 1$ y $\lambda = 0.05$, tomando la velocidad de una onda de sonido como $c = 343\text{m/s}$, da como resultado una onda de frecuencia de 6860 Hz.

Consideremos la relación que hay entre la frecuencia angular y la frecuencia natural:

$$\omega = 2\pi f, \quad (18)$$

al sustituirlo en la ecuación 17 y considerando que $x = 0$, entonces queda reescrito:

$$S(t) = A \sin(2\pi ft), \quad (19)$$

Es otra forma de representar la ecuación de una onda de sonido y la manera en la que se usará para el presente trabajo.

La Figura 7 muestra una onda sinusoidal, con una frecuencia natural $f = 0.01$ Hz, lo que representa un tono puro, la amplitud provee información de la energía de la onda y se relaciona con el volumen o intensidad de sonido.

3.2. Curvas isofónicas

Fletcher y Munson definieron las curvas isofónicas como la relación entre el nivel de presión sonora de un sonido y la percepción de su magnitud de volumen, en función de la frecuencia del sonido (Fletcher

& Munson, 1933). Estas curvas representan la sensibilidad del oído humano a diferentes frecuencias y son una característica importante del sistema auditivo (Suzuki & Takeshima, 2004).

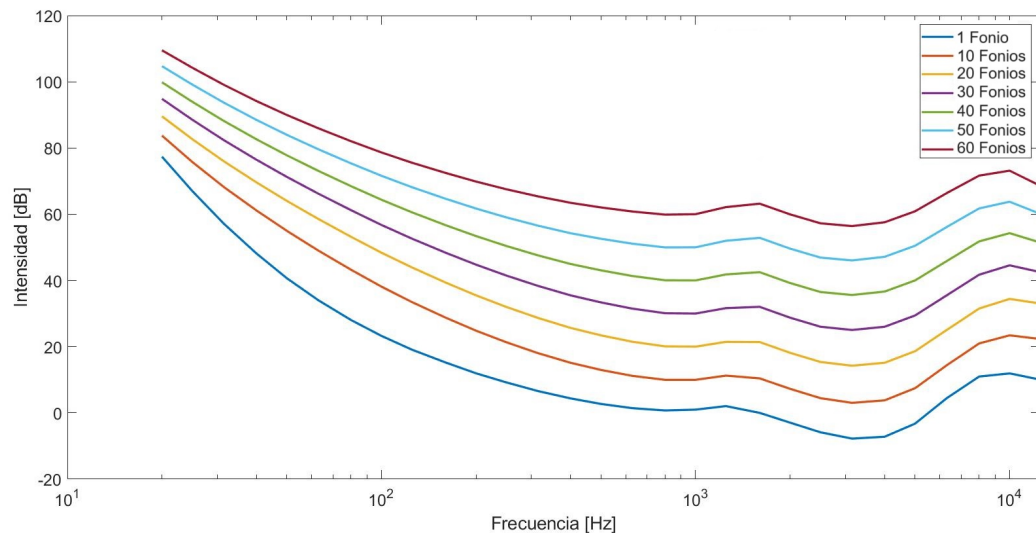


Figura 8. Curvas isofónicas por la norma ISO226, cada curva expresa un mismo nivel de sonoridad medido en fonios.

Las curvas de Fletcher y Munson, ilustradas en la Figura 8, proporcionan información fundamental sobre el sistema auditivo al describir su sensibilidad a distintas frecuencias. Estas curvas pueden servir como punto de referencia para determinar el volumen más bajo que una persona puede percibir. No obstante, la interpretación de los datos puede resultar complicada sin una referencia clara. Una forma de mejorar la legibilidad es realizar una transformación de las curvas isofónicas, representándolas como líneas horizontales. Esta representación facilitaría la comparación de los umbrales auditivos, manteniendo las unidades en decibeles de pérdida auditiva (dB HL).

3.3. Audiómetro

A lo largo del tiempo, se han desarrollado herramientas para el estudio de la capacidad auditiva de las personas. Las primeras pruebas para evaluar la salud auditiva de un individuo se realizaron con un diapasón¹. La prueba de Rinne, presentada en 1845, comparaba la transmisión del sonido a través del aire y del hueso utilizando un diapasón a 512 Hz. La prueba de Weber es otra prueba con diapasón que suele realizarse junto con la prueba de Rinne. Se utiliza frecuentemente para determinar la posible causa

¹Dispositivo de acero doblado en forma de horquilla con pie, que al vibrar emite sonido a una frecuencia específica (Real Academia Española, 2022, definición 1).

y ubicación de la pérdida auditiva.

Harvey Fletcher, graduado en física de la Universidad de Chicago, realizó importantes contribuciones al estudio del habla y la audición en la década de 1920. Junto con su grupo de investigación, adquirió una amplia experiencia en el desarrollo del audiómetro electrónico y en la comprensión de la salud auditiva, como se evidencia en los artículos publicados durante ese período (Fletcher, 1992).

Fue en la década de 1950 cuando surgieron los primeros audiómetros automáticos que permitían registrar los resultados de las evaluaciones auditivas en un gráfico, tal como se describe en el artículo de Epstein (1960). Estos avances tecnológicos proporcionaron una forma más precisa y eficiente de evaluar la audición de las personas.

3.3.1. Características de un audiómetro

Los audiómetros modernos son instrumentos electrónicos diseñados para generar sonidos que varían en frecuencia e intensidad. Los sonidos generados se emiten a través de un transductor que envía el estímulo auditivo al paciente. Las características del audiómetro pueden variar según el modelo. En general, consta de un panel de control que incluye botones, perillas y una pantalla, una sección para configurar las características de la señal de sonido (control de frecuencia e intensidad) y un transductor que se conecta al audiómetro. Otro componente es un micrófono, que tiene funciones tanto para que el especialista pueda comunicarse con el paciente como para detectar el ruido ambiental que pueda afectar los resultados de la prueba (Kapul et al., 2017).

El control de la frecuencia se utiliza para aumentar o disminuir la frecuencia de la señal que el paciente escucha. Las frecuencias permitidas corresponden a las frecuencias audiométricas estándar. Estas frecuencias pueden variar según los diferentes modelos de audiómetros, pero las principales son 250, 500, 1000, 2000, 4000 y 8000 Hz. Estas frecuencias son octavas de un mismo tono o el doble de la frecuencia anterior. Sin embargo, es importante destacar que esta relación de frecuencias no siempre se cumple en todos los audiómetros.

El control de la intensidad permite aumentar o disminuir el nivel de intensidad sonora del estímulo auditivo. Los cambios de intensidad se realizan en incrementos de 5 dB y el rango de intensidad puede ir desde -10 dB hasta 120 dB. Según la Asociación Estadounidense del Habla, Lenguaje y Audición (ASHA, por sus siglas en inglés), los estímulos auditivos deben tener una duración entre 1 y 2 segundos

(American Speech-Language-Hearing Association, 2005).

Los transductores pueden ser audífonos que se insertan en el canal auditivo, audífonos circumaurales o altavoces para las pruebas de audiometría vía aérea, y transductores de vibración ósea para las pruebas de audiometría vía ósea.

3.4. Audiometría tonal

La audiometría tonal es una prueba utilizada para el diagnóstico auditivo, en la cual se mide la capacidad del oído humano para percibir sonidos. El principal objetivo de una evaluación auditiva es identificar problemas auditivos, determinar el nivel y la localización de la pérdida auditiva y, en general, evaluar la salud auditiva. Esto se logra presentando estímulos auditivos de diferentes frecuencias e intensidades que el paciente debe escuchar y determinar la mínima intensidad de sonido.

La audiometría se lleva a cabo mediante el uso de un audiómetro y se realiza en una cabina insonorizada, donde solo ingresa la persona que está siendo evaluada. Antes de comenzar la prueba, al paciente se le proporcionan instrucciones claras y se le explica la acción que debe realizar cuando perciba un sonido. El paciente se coloca unos audífonos y se inicia la prueba con una de las frecuencias audiométricas. Se evalúa cada oído por separado.

Durante la prueba, se presentan diferentes tonos o frecuencias, y el paciente debe indicar levantando la mano cuando escucha cada tono, como es mostrado en la Figura 9. El audiómetro registra la respuesta del paciente, lo que permite al profesional de la salud obtener información sobre los umbrales auditivos en cada frecuencia evaluada (Shearer et al., 1999).

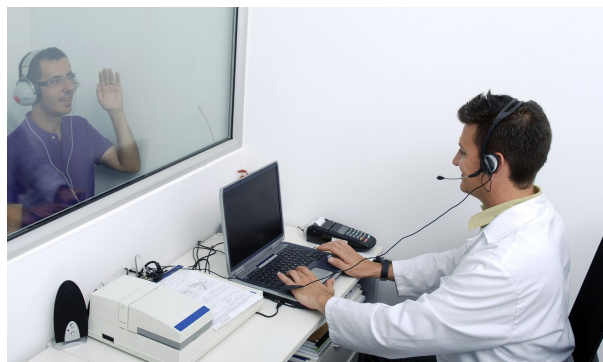


Figura 9. Representación de una audiometría (Salud auditiva integral, 2018).

Al finalizar las mediciones de los umbrales auditivos, los datos se representan gráficamente en un audiograma. En el eje horizontal, se muestran las frecuencias, mientras que en el eje vertical, se muestra la intensidad medida en dB HL (decibeles de pérdida auditiva).

Un audiómetro tradicional está compuesto por botones o perillas que permiten su manipulación. Algunos modelos pueden incluir una pantalla para mostrar los parámetros de los estímulos. En el panel de control del audiómetro se encuentran los controles para ajustar y seleccionar los diferentes parámetros. Además de generar tonos, suele contar con un control para comunicarse con el paciente a través de un micrófono.

Los controles de frecuencia se utilizan para aumentar o disminuir la frecuencia de la señal emitida. De manera similar, los controles de intensidad se utilizan para aumentar o disminuir el nivel de intensidad de la señal. Los controles de señal permiten generar cambios en el tipo de señal, como tonos puros o ruido. Los transductores más comunes utilizados en una audiometría son los auriculares supraaurales, altavoces y vibradores óseos.

3.5. Audiograma

El audiograma es la representación gráfica de una audiometría (ver figura 10). En el eje vertical se representa el volumen o la intensidad de sonido, medido en decibeles de audición (dB HL). Los valores de menor volumen se encuentran en la parte superior del eje, mientras que los de mayor volumen se ubican hacia la parte inferior. En el eje horizontal se representan las frecuencias audiométricas, medidas en hertz (Hz). Las frecuencias están separadas equidistantemente y aumentan a medida que avanzamos hacia la derecha.

Por convención, los símbolos audiométricos utilizados en los audiogramas son diferentes para cada oído. Las mediciones para el oído derecho se representan con el color rojo y se marcan con un círculo, mientras que las mediciones para el oído izquierdo se representan con el color azul y se marcan con una cruz (American Speech-Language-Hearing Association, 1990).

La Figura 11, muestra el grado de pérdida auditiva dependiendo de la gravedad de la misma. Una persona con audición normal es aquella que no tiene dificultades para oír los sonidos débiles, los cuales tienen intensidades entre -10 y 14.9 dB. Con pérdida auditiva ligera, la persona tiene el umbral auditivo entre 15 y 24.5 dB. Se considera pérdida auditiva leve si el umbral está entre 25 y 39.9 dB, moderada entre

40 y 54.9, moderadamente grave si está entre 55 y 69.9 dB, grave si está entre 70 y 90 dB y pérdida profunda si el umbral es mayor a 90 dB (Clark, 1981).

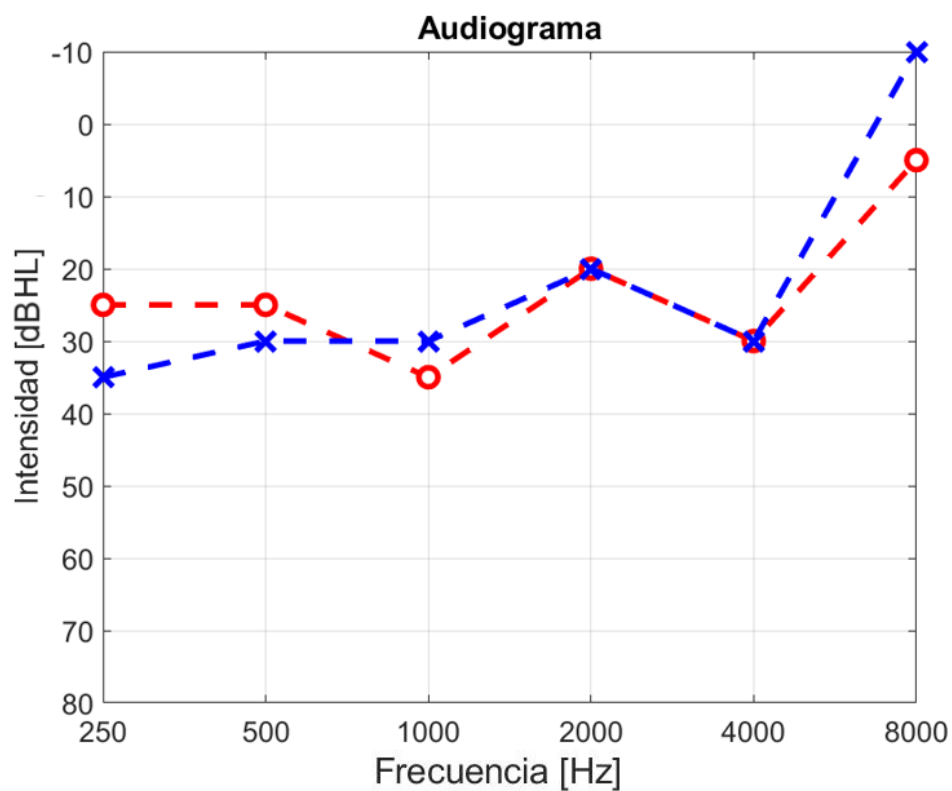


Figura 10. Ejemplo de un audiograma. Los círculos rojos representan las mediciones para el oído derecho, mientras que las cruces azules representan las mediciones para el oído izquierdo.

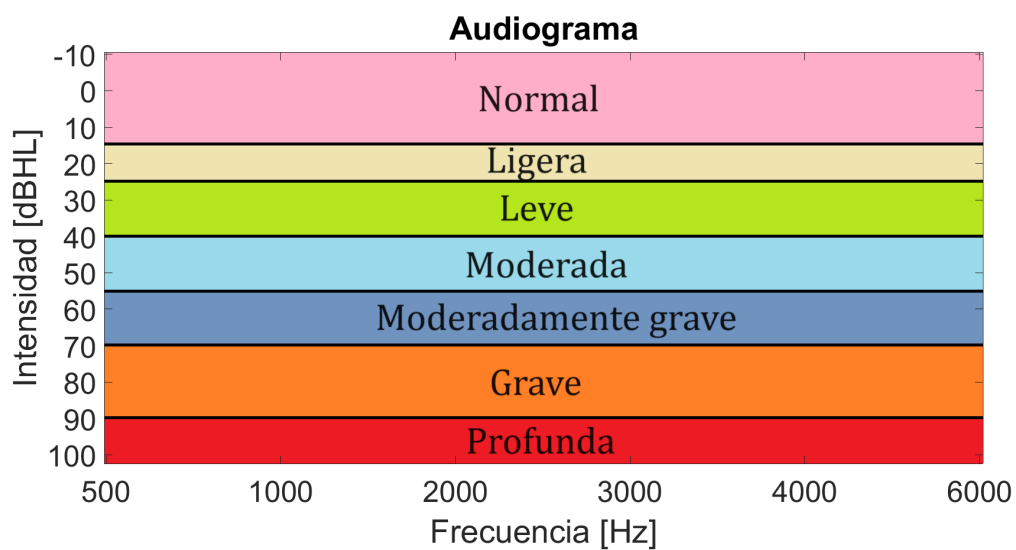


Figura 11. Clasificación de la pérdida auditiva relacionada al umbral auditivo.

3.6. Método de Hughson-Westlake

En la década de 1930, los estudios clínicos de audiometrías mostraron que la detección del umbral parecía ser más acertada cuando se emiten sonidos en breves ráfagas y con breves pausas de silencio a medida que el nivel de sonido cambiaba. Se observó que pasos de 5 dB en el cambio del nivel de sonido, mostraba una menor diferencia en los umbrales con cambios ascendentes y descendentes. En el año de 1943, Cordia C. Bunch publicó el libro *Clinical Audiometry*, donde describe una técnica para encontrar los umbrales auditivos. Basado en el método ascendente Bunch emitía sonido continuamente en un nivel muy por debajo al que el paciente puede escuchar, el sonido sube de volumen hasta que el paciente le indica que ha escuchado, repitiéndolo entre tres o cuatro veces, definiendo el umbral como el nivel a mitad de camino entre el promedio de las corridas descendentes contra las ascendentes (Jerger, 2017).

Fue hasta el año 1944 que Walter Hughson y Harold Westlake publicaron un artículo en el que describían como medir los umbrales auditivos, basados en la técnica de Bunch, con la diferencia que los sonidos deben ser emitidos en breves ráfagas de sonido en lugar de que el estímulo sean tonos continuos. Agregando que el umbral sea definido únicamente por una serie de tonos ascendentes. El método de Hughson-Westlake es conocido como el método ascendente. Definieron el umbral como el nivel más bajo en el que hubo respuesta en 2 de 3 respuestas ascendentes. Posteriormente el Instituto Americano de Estándares Nacionales (2004) adaptó la definición que debían ser 2 de 3 respuestas positivas. Fue hasta que Carhart y Jerger (1959) publicaron el artículo con el método de Hughson-Westlake y discutieron el tamaño de paso de 5 dB en el nivel de sonido. Este método funciona de la siguiente manera: cuando la persona escucha el tono, disminuye 10 dB, cuando la persona no escuche el tono, aumenta la intensidad 5 dB, se habrá encontrado el umbral para un tono cuando el paciente escuche al menos el 50% de la veces.

Los enfoques actuales para automatizar la audiometría de tonos puros se dividen en tres categorías: Automatización del método de límite (técnica de Hughson-Westlake), automatización del método de ajuste (audiograma de seguimiento tipo Békesy), y automatización de método adaptativo.

El algoritmo utilizado para medir el umbral auditivo en el audiómetro de este proyecto se basa en el método de Hughson-Westlake. En la Figura 12 se muestra el diagrama de flujo que determina el umbral auditivo para una sola frecuencia, pero este proceso debe repetirse para todas las frecuencias audiométricas.

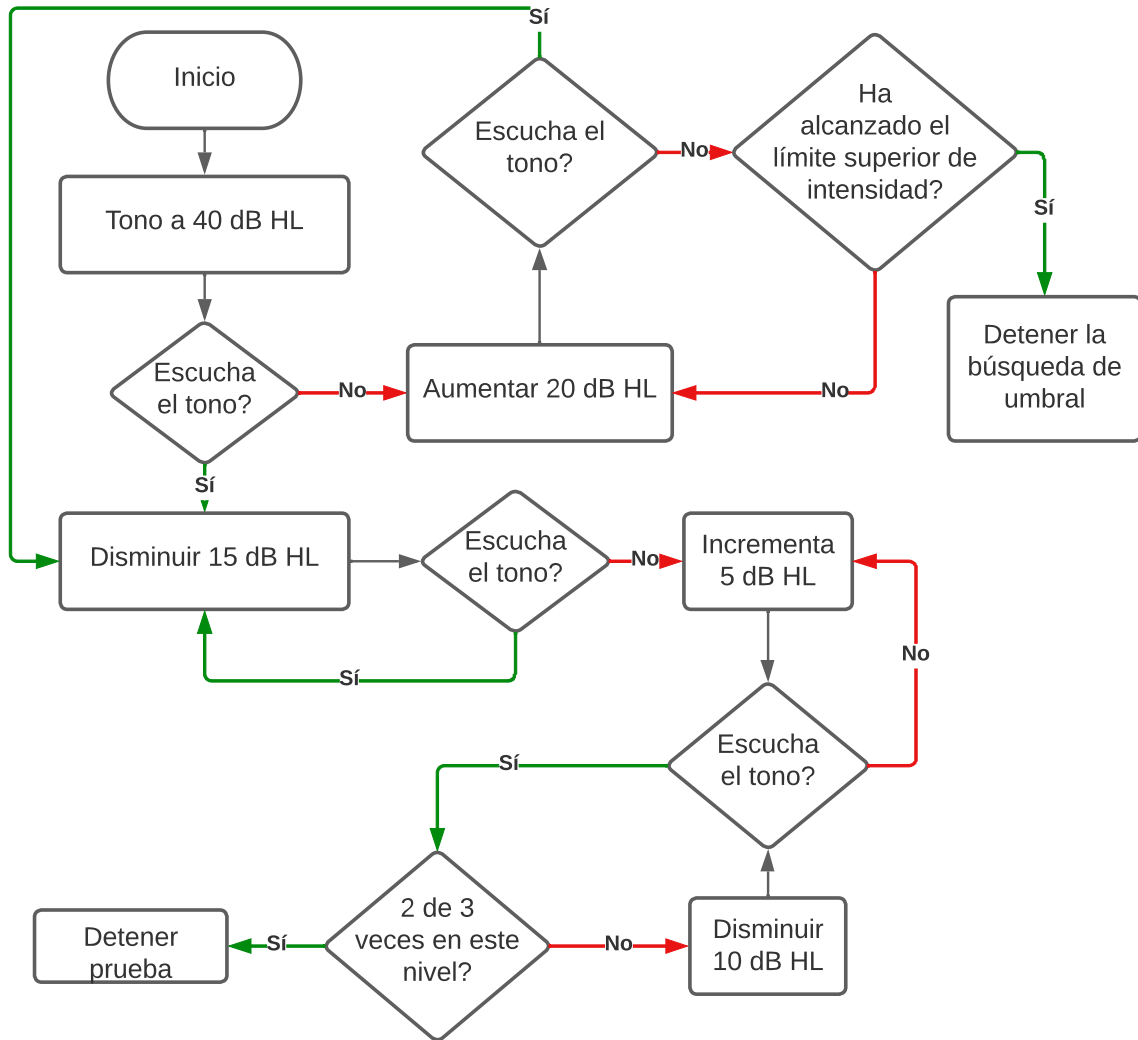


Figura 12. Proceso del método de Hughson-Westlake modificado, implementado para encontrar los umbrales auditivos.

En la primera etapa del algoritmo, se realizan cambios en la intensidad del estímulo auditivo. Dependiendo de la primera respuesta del paciente, los cambios pueden ser incrementos de 20 dB o disminuciones de 15 dB. Esta etapa inicial permite aproximarse rápidamente al rango de intensidad en el cual el paciente puede detectar el estímulo.

Por último, los cambios en la intensidad del estímulo se realizan en incrementos de 5 dB para aumentar la precisión y disminuciones de 10 dB para obtener una mayor resolución en la determinación del umbral auditivo. Esto se basa en el método tradicional utilizado en las pruebas de audiometría.

3.7. Calibración

La calibración es un proceso de ajuste o corrección de un instrumento de medición para que proporcione resultados precisos y consistentes. Este proceso es esencial en una amplia variedad de aplicaciones donde la precisión de las mediciones es crítica. La calibración se realiza comparando las lecturas del instrumento que se va a calibrar con las lecturas de un estándar de referencia o un instrumento de referencia de alta precisión.

La calibración de un audiómetro permite tener mayor confianza en los resultados de una audiometría y es una de las etapas más importantes para el diseño del audiómetro digital. Específicamente, la calibración para un audiómetro se ajusta el nivel de presión sonora que se emite a través de los audífonos del audiómetro.

3.7.1. Intensidad de sonido

Recordemos que la ecuación de una onda de sonido está dada por la ecuación 19, prestando atención únicamente a la frecuencia y a la amplitud de la onda, serán las variables que modificarán la intensidad de la onda. Como menciona Tippens (2007), la variación de la intensidad de una onda sonora tendrá un efecto sensorial en el nivel de fuerza (volumen) de quien lo escuche. La ecuación de intensidad es la siguiente:

$$I = 2\pi^2 f^2 A^2 \rho c, \quad (20)$$

donde f es la frecuencia, A la amplitud de la onda, ρ es la densidad del medio de propagación y c es la velocidad del sonido en el medio.

Considerando que la densidad del medio y la velocidad del sonido son constantes en el medio de propagación, dejando fija la frecuencia se obtiene que la intensidad para una frecuencia determinada depende únicamente de la amplitud de la onda.

3.7.2. Método de calibración

Con el proceso de calibración, buscamos determinar las amplitudes necesarias para emitir intensidades desde -10 dB hasta 90 dB para cada una de las frecuencias audiométricas.

Se utilizará un sonómetro para medir la intensidad del sonido en el proceso de calibración. Inicialmente, desconocemos la magnitud de las amplitudes necesarias para emitir las intensidades mencionadas a través de los audífonos. Por lo tanto, el primer paso consiste en definir un conjunto de amplitudes que variarán mientras medimos la intensidad emitida por los audífonos. La Tabla 2 muestra las amplitudes utilizadas para la calibración, estas no representan el valor real de la amplitud de la onda física, son amplitudes utilizadas en MATLAB para crear una señal. Dependiendo de la frecuencia de la onda de sonido, estas amplitudes serán multiplicadas por un coeficiente específico, resultando en:

$$AmplitudFinal = C * (AmplitudInicial), \quad (21)$$

donde C es el coeficiente específico.

Tabla 2. Amplitudes iniciales para la calibración en unidades arbitrarias.

Amplitud				
				1
0.8	0.6	0.4	0.02	0.1
0.08	0.06	0.04	0.02	0.01
0.008	0.006	0.004	0.002	0.001

Para la frecuencia de 500 Hz se utiliza un coeficiente de 2.5, mientras que para las frecuencias de 1000, 2000, 3000, 4000 y 6000 Hz utilizan el coeficiente 1.

El segundo paso implica medir la intensidad a través de los audífonos para cada frecuencia, lo que proporciona una relación entre la intensidad y la amplitud como resultado. El tercer paso consiste en realizar una interpolación de los datos para encontrar una ecuación que se aproxime a los valores medidos, obteniendo así una ecuación de intensidad que depende de la amplitud.

$$I = a + b * \ln(A), \quad (22)$$

$$A = e^{\frac{I-a}{b}}, \quad (23)$$

donde a y b son los coeficientes. Estos coeficientes se ingresan en la ecuación 23 para calcular las 21 amplitudes que corresponden a las intensidades entre 5 dB y 90 dB, en incrementos de 5 dB.

Hasta este punto, se ha calculado la ecuación para obtener las amplitudes a partir de una intensidad en dB. Sin embargo, es importante recordar que el audiograma mide la intensidad en dB HL, por lo que es necesario realizar una conversión. De acuerdo con la Norma ANSI S3.6-1996, la transformación de decibeles a decibeles de pérdida auditiva se realiza mediante una relación lineal.

Para llevar a cabo esta conversión, se requiere una audiometría obtenida con un audiómetro calibrado. Se realiza una comparación entre los umbrales resultantes de la audiometría obtenida con el audiómetro digital y los resultados de la otra audiometría. Para cada frecuencia, se realiza un ajuste para igualar los resultados. Esta etapa de calibración concluye una vez que obtenemos las amplitudes correctas que permiten al audiómetro generar las intensidades utilizadas en la audiometría.

Capítulo 4. Resultados

4.1. Sistema de visión computacional

En una audiometría tradicional, cuando un paciente escucha un sonido, suele levantar la mano para comunicárselo al especialista. Por este motivo, el objetivo de nuestro sistema de visión computacional es reconocer la mano. La mano abierta indica que el paciente ha escuchado el sonido, mientras que la mano cerrada o la ausencia de ella indica lo contrario.

Para lograr el reconocimiento, es necesario que la mano abierta se muestre con la palma de la mano hacia la cámara de video, de manera similar para el estado de la mano cerrada.

El sistema consta de varias etapas. En la primera etapa, se obtiene la información de los dos estados de la mano que se van a evaluar, es decir mano abierta y cerrada. La segunda etapa consiste en la captura de la imagen objetivo, que es la mano cuyo estado se desea reconocer. En la tercera etapa, se aplica el filtro HSV, descrito en el Capítulo 2, a la imagen objetivo y a las imágenes obtenidas en la primera etapa, este filtro nos permite segmentar la mano. Posteriormente, para obtener los patrones de la imagen se calcula la transformada de Fourier de cada imagen filtrada, considerando las propiedades mencionadas anteriormente (tabla 1), la de traslación es útil para esta aplicación, debido a que se pueden extraer patrones independientemente donde se encuentre el objeto de interés. Con los patrones de la transformada de Fourier de las tres imágenes, se calcula el coeficiente de correlación entre la imagen de la mano abierta y la imagen objetivo, así como entre la imagen de la mano cerrada y la imagen objetivo (ver Figura 13). Para determinar a qué estado corresponde la imagen objetivo, se comparan las dos correlaciones y se selecciona la de mayor valor.

4.1.1. Procesamiento de las imágenes

La primera etapa del proceso consistió en la captura de imágenes que tiene una resolución predeterminada de 720x1280 pixeles, significa que son 720 columnas y 1280 filas. El tiempo de procesamiento para una sola imagen fue de 0.1389 segundos. Durante este proceso, las imágenes fueron modificadas para extraer características relevantes, se aplicó un filtro HSV, una transformada de Fourier, una normalización y un filtro paso bajo.

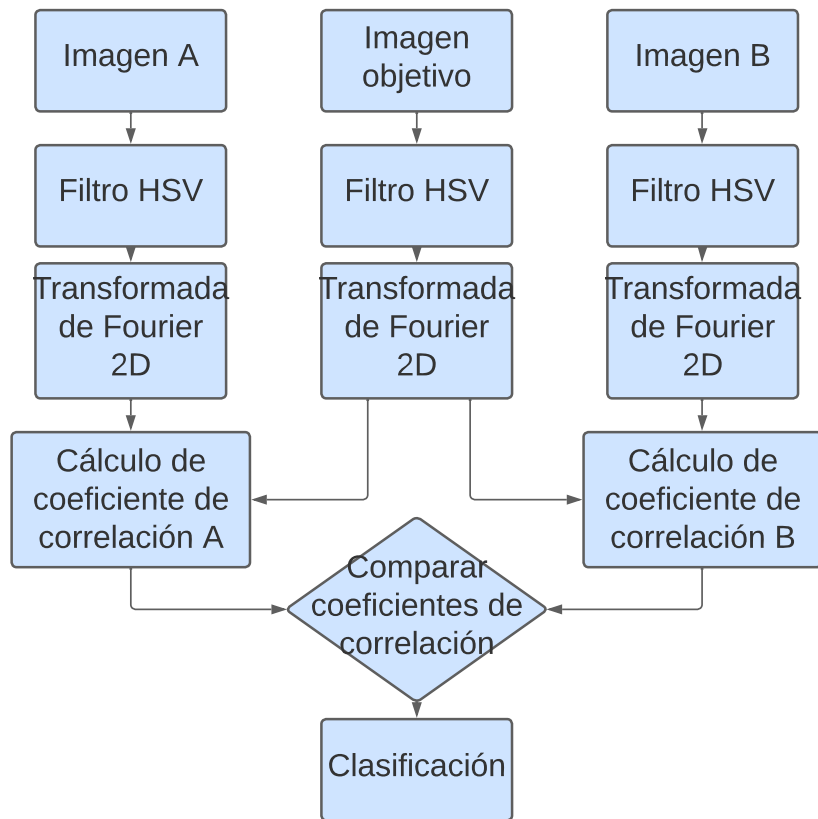


Figura 13. Proceso para el reconocimiento del patrón de la mano, la imagen A e imagen B son las imágenes de referencia y la imagen objetivo es la que se quiere clasificar.

Dado que las pruebas se realizaron en tiempo real, lo que implicaba una interacción entre el sistema de visión y la respuesta de la persona, era necesario reducir el tiempo de procesamiento para una interacción más rápida. Por lo tanto, se decidió reducir la resolución de la imagen a 320x240 píxeles. Con esta resolución más baja, el tiempo de procesamiento se redujo a 0.0016 segundos. Se observó que la característica extraída de la imagen con baja resolución no presentaba cambios relevantes para la etapa de clasificación.

El primer cambio realizado en las imágenes fue la modificación de su tamaño, lo cual resultó en una mejora significativa en el tiempo total de procesamiento. En la segunda etapa del procesamiento, se empleó un filtro diseñado para segmentar los colores correspondientes a tonos de piel. Al analizar la Figura 14(a), se puede apreciar una imagen que aún no ha sido sometida al filtro, pero ha sido reducida en resolución. Por otro lado, en la Figura 14(b), se observa el resultado de aplicar la binarización con el filtro HSV, donde las regiones en color blanco representan los tonos de piel detectados por el filtro. Se puede notar que el filtro ha segmentado correctamente la mano y el rostro.

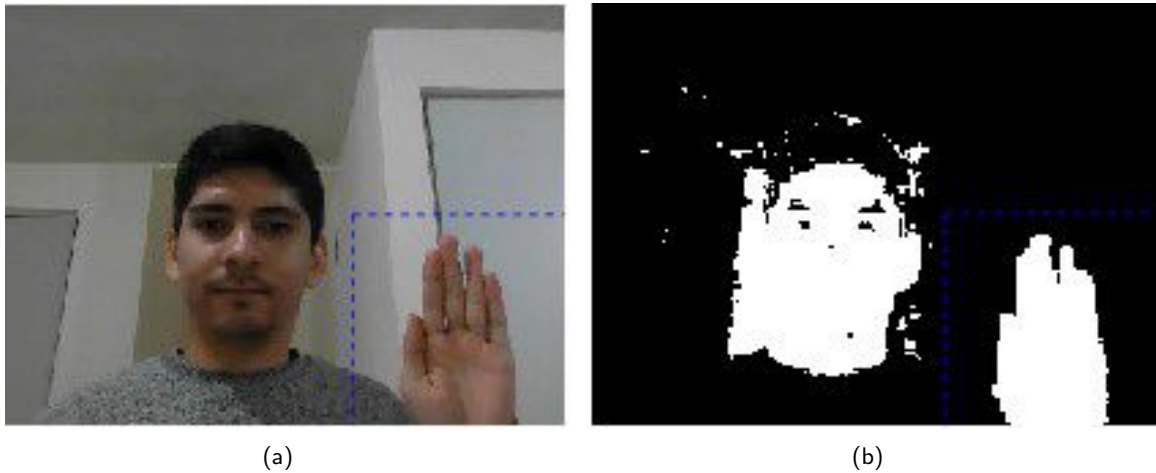


Figura 14. Segmentación de tonos color piel utilizando el filtro HSV. En (a) se muestra la imagen sin filtros y en (b) la imagen al pasar por el filtro HSV. Ambas imágenes con 320x240 píxeles.

En la tercera etapa del procesamiento, se aplicó la transformada de Fourier a las imágenes con el fin de visualizar el plano de las frecuencias. Se observó que los valores más representativos en la imagen resultante de la Figura 15 se localizaban en el centro, lo que indicaba el área era de interés para la extracción de características. Estas imágenes contenían los patrones a reconocer, mostrando las características distintivas de cada categoría que el sistema debía clasificar.

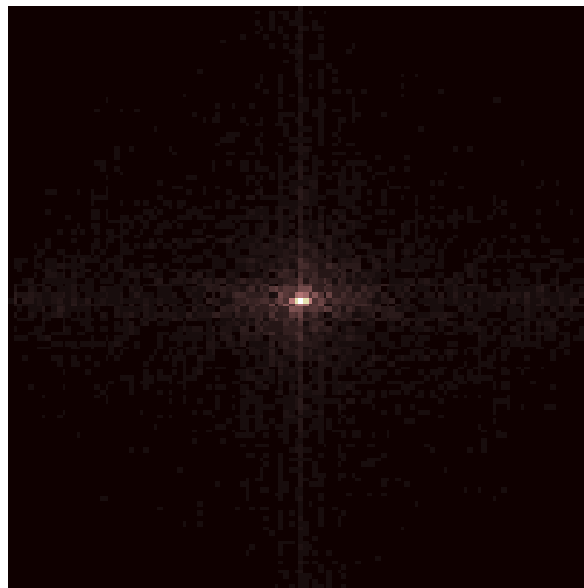


Figura 15. Transformada de Fourier de la mano segmentada en la figura 14(b).

Es importante destacar que en aquellas muestras donde no se detectaba la mano, los píxeles presentaban valores en el plano de los números complejos. Esta información se consideró como una característica útil

en la etapa de clasificación.

Sin embargo, es fundamental tener en cuenta la distancia entre la mano y la cámara, ya que esto afecta la imagen resultante de la transformada de Fourier debido a la propiedad de cambio de escala, por otro lado, si la mano no se encuentra en posición vertical, debido a la propiedad de rotación el patrón en la imagen se verá afectado con una rotación. Por lo tanto, se recomienda que la posición de la mano se mantenga a una distancia y posición constante, y en posición vertical durante la interacción con el sistema. De esta manera, se asegurará una mayor consistencia en las características extraídas y, por ende, en la etapa de clasificación.

4.1.2. Clasificación

Para clasificar una única imagen, se utiliza el coeficiente de correlación para medir la similitud entre las características de la mano abierta o cerrada. Los valores para un único caso se presentan en la Tabla 3.

Además, existe un escenario en el cual el área de interés en la imagen no logra detectar la mano, lo que resulta en la ausencia de una segmentación precisa. Para facilitar la tarea del sistema de visión en estos casos, se consideró que esta situación correspondía a la respuesta 'No escuchó'. Sin embargo, se identificaron características más efectivas para distinguir entre dos respuestas: la ausencia de detección de la mano y la detección de la mano abierta.

En la Tabla 3 se muestran los valores de correlación obtenidos para este escenario en particular. Estos valores de correlación representan una medida de la similitud entre las características extraídas de las imágenes y las respuestas esperadas.

Tabla 3. Valores del coeficiente de correlación para el caso de una imagen.

Mano	Apertura real de la mano	
	Abierta	Cerrada
Abierta	0.9849	0.9192
Cerrada	0.8978	0.9888

Al analizar los resultados de la Tabla 3, se observa que las características utilizadas lograron una diferenciación entre las dos respuestas mencionadas. Se obtuvieron valores de correlación significativamente más altos para la detección de la mano abierta en comparación con la ausencia de detección de la mano, ver Tabla 4. Esto sugiere que las características utilizadas fueron efectivas para distinguir entre estas

respuestas y que el sistema de visión logró un desempeño satisfactorio en esta tarea.

Un espacio físico que presenta un cambio en la iluminación muy grande, como la entrada de luz del sol a través de una ventana, llega a causar ruido en las capturas que se toman desde la cámara. En ese sentido, se recomienda que el fondo donde se encuentre la persona tenga un color liso y se esté alejado de los cambios de luz provocados por la luz del sol.

Tabla 4. Coeficiente de correlación para el caso de una imagen con la mano abierta y una imagen sin la mano.

	Apertura real de la mano	
Mano	Abierta	Sin mano
Abierta	0.9712	0.4866

4.1.3. Evaluación

Se creó una base de datos de 500 imágenes de manos abiertas y 500 imágenes manos cerradas, las imágenes capturadas presentan diferentes características como diferente nivel de iluminación, así como diferentes escenarios que modifican el fondo. En la base de datos están las imágenes de la figura 16. Una manera de evaluar el modelo de reconocimiento es a través de la matriz de confusión, obteniendo la sensibilidad, especificidad y exactitud. Donde la sensibilidad es el porcentaje de casos de mano abierta que clasificó correctamente, y la exactitud el porcentaje total de casos correctos. Estos valores son mostrados en las matrices de confusión, como se muestran en la figura 17.

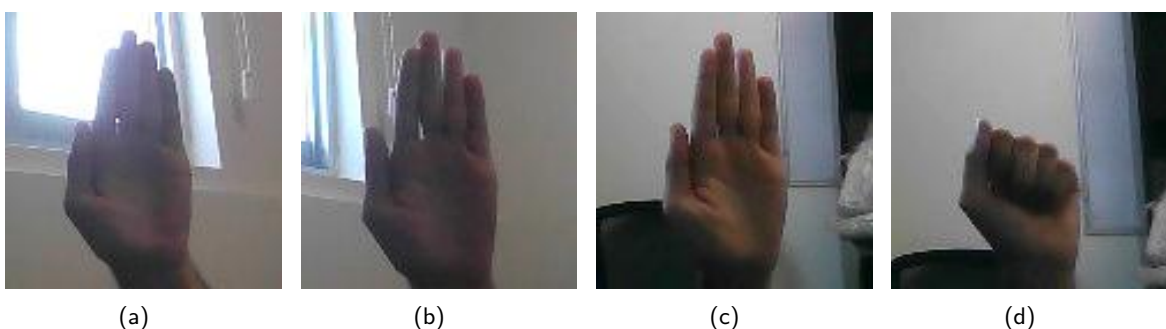


Figura 16. Imágenes tomadas de la base de datos de manos, utilizadas para evaluar el sistema de reconocimiento de patrones, donde a, b y c imágenes de la mano abierta y d imagen de la mano cerrada.

Al evaluar el rendimiento del modelo de reconocimiento utilizando las figuras (a) y (d) de la Figura 16, se obtuvo una exactitud del 49%. Es importante destacar que, si bien el modelo logra un buen

reconocimiento de las imágenes de manos cerradas, su desempeño es deficiente en la identificación de imágenes con manos abiertas. La figura 16(a) presenta un fondo con una ventana por donde ingresa luz exterior, lo que introduce una considerable cantidad de ruido en la imagen y dificulta el proceso de segmentación. Al usar las imágenes (b) y (d) de la Figura 16 en el modelo, se observa una notable mejora en la exactitud, alcanzando un valor de 86.2%. De las imágenes de manos abiertas, un 73.4% fueron clasificadas correctamente, lo que equivale a 367 de las 500 imágenes. A este porcentaje es a lo que llamamos sensibilidad. En cuanto a las 500 imágenes de manos cerradas, 495 fueron clasificadas de manera precisa. Finalmente, al usar las imágenes (c) y (d) de la Figura 16 en el modelo, la matriz de confusión resultante exhibe una exactitud del 99%. La sensibilidad se incrementa al 98.8%, lo que representa una mejora sustancial en comparación con los dos modelos previos. Además, se logra una tasa de aciertos del 99.2% en las imágenes de manos cerradas.

4.2. Desarrollo del audiómetro automatizado

El audiómetro automatizado y el sistema de visión interactúan de manera simultánea durante la prueba. La combinación de ambos da como resultado un audiómetro digital que realiza de manera automática la obtención de los umbrales auditivos, obteniendo la respuesta del paciente mediante visión computacional. En esta sección se describirán los resultados obtenidos con el audiómetro.

El audiómetro digital desarrollado en esta tesis se implementó en MATLAB y tiene la capacidad de reproducir sonidos en un rango de frecuencia que abarca desde 500 hasta 6000 Hz. Las frecuencias seleccionadas son: 500, 1000, 2000, 3000, 4000 y 6000 Hz. Además, el audiómetro es capaz de reproducir sonidos en un rango de intensidad que va desde -10 hasta 90 dB, en pasos de 5 dB.

La intensidad y la frecuencia de la señal emitida por el audiómetro son controladas de manera automática a través del software desarrollado en MATLAB. El sistema permite seleccionar el oído a evaluar y cambiar automáticamente entre las frecuencias audiométricas seleccionadas. El ajuste del volumen se realiza en función de la respuesta del individuo, ya sea aumentándolo o disminuyéndolo, con el objetivo de determinar el umbral auditivo.

La duración de los estímulos emitidos por el audiómetro es de 2 segundos, lo cual está dentro del intervalo recomendado por la Asociación Americana de Audiología (ASHA). Gracias a la automatización implementada, el sistema es capaz de realizar las tareas de manera automática.

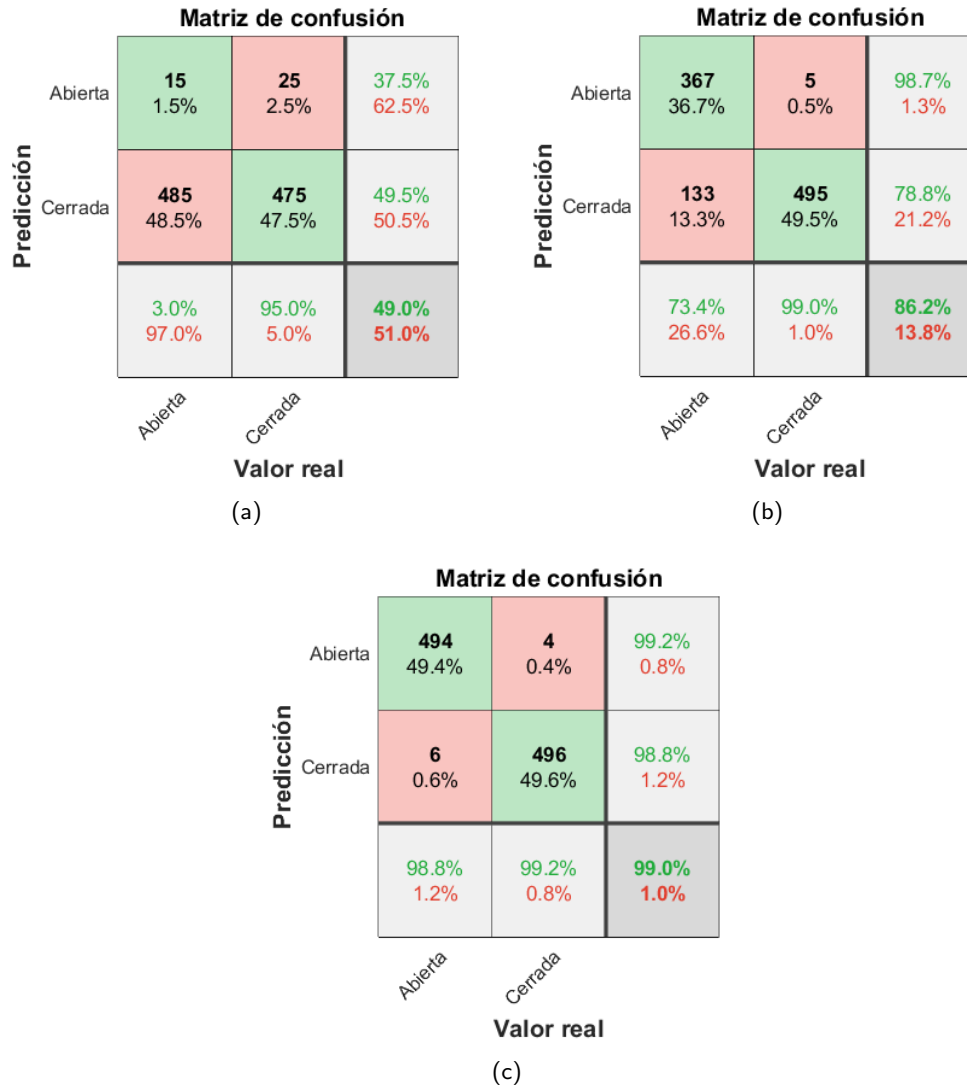


Figura 17. Matrices de confusión, se evaluó el modelo usando 3 imágenes diferentes de la mano abierta y una imagen de la mano cerrada, a) utiliza la imagen (a) y (d), b) utiliza la imagen (b) y (d), c) utiliza la imagen (c) y (d), imágenes de la figura 15.

4.2.1. Implementación

4.2.1.1. Generación de tonos

Es fundamental tener en cuenta que la evaluación de cada oído se realiza de manera individual, lo que implica la generación de dos señales distintas, una para cada audífono. En MATLAB, una forma de representar esto es mediante una matriz donde la primera fila contiene la señal correspondiente al oído izquierdo, mientras que la segunda fila contiene la señal destinada al oído derecho. Podemos utilizar la

expresión $signalFinal = [zeros(size(signal)); signal]$ para definir esta matriz, donde $signal$ representa los valores del tono generado y $signalFinal$ es la señal resultante que se envía a los audífonos.

La función $signal$ describe las características del estímulo auditivo y se define como $signal = A * \sin(2\pi * Freq * tiempo)$. En esta expresión, A representa el término de amplitud, el cual es una matriz que contiene las amplitudes necesarias para generar diferentes niveles de volumen. $Freq$ se refiere a la frecuencia de la señal y $tiempo$ determina la duración de la señal. Para su implementación, se define la variable $tiempo$ como $(1 : Fs * D)/Fs$, donde Fs es la frecuencia de muestreo y D es la duración del sonido.

Es importante destacar que los cambios en la señal no son controlados manualmente por una persona, ya que todo el proceso está automatizado, tal como se muestra en el diagrama de flujo de la Figura 18. El software se encarga de generar y manipular las señales de manera automática, cambiando las frecuencias y ajustando las amplitudes según la respuesta del individuo. Esto permite una evaluación precisa y eficiente de los umbrales auditivos sin la necesidad de intervención manual durante el proceso de prueba.

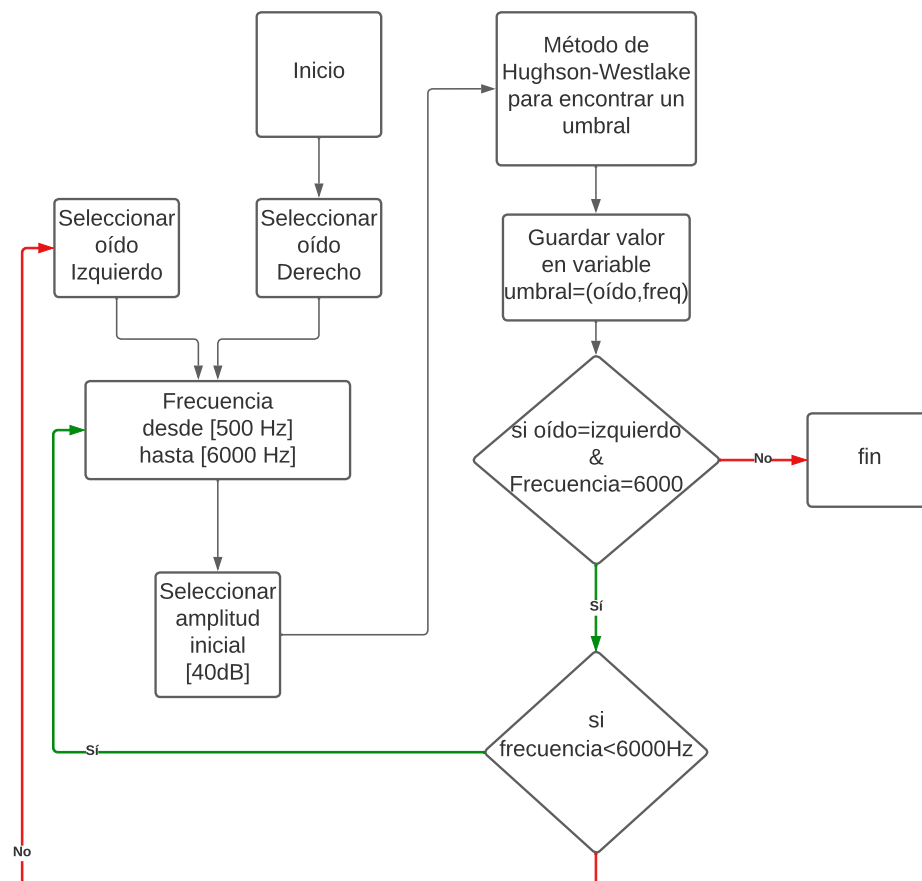


Figura 18. Diagrama de flujo general del proceso para automatizar el audiómetro digital.

Se realiza primero la evaluación del oído derecho, comenzando con la frecuencia de 500 Hz y una intensidad inicial de 40 dB. Una vez determinado el umbral auditivo para esa frecuencia, se procede a seleccionar la siguiente frecuencia y repetir el proceso hasta llegar a la frecuencia de 6000 Hz. Una vez obtenidos los umbrales auditivos para el oído derecho, se inicia la evaluación del oído izquierdo, repitiendo el mismo procedimiento para determinar los umbrales en las frecuencias audiométricas.

4.3. Interfaz de Usuario

Todo lo descrito en este capítulo fue implementado en una interfaz gráfica para complementar al audiómetro digital automatizado. A continuación se describirán las etapas correspondientes a la interfaz de usuario y posteriormente las características del audiómetro automatizado.

La interfaz permite a una persona realizar una audiometría tonal sin la necesidad de contar con la presencia de un experto en el área de audiología. Es importante resaltar que este dispositivo, no suplente al especialista, sino que ofrece una herramienta para la realización de las pruebas audiométricas. El especialista siempre será indispensable para la interpretación y el diagnóstico, en caso de existir alguna deficiencia de la salud auditiva.

El individuo que realice la audiometría debe prestar atención a las instrucciones que se le muestren en la pantalla y llevar a cabo las acciones indicadas durante toda la prueba. La interfaz fue desarrollada utilizando la Interfaz Gráfica de Usuario (GUI, por sus siglas en inglés) de MATLAB.

En la primera etapa, comienza con un mensaje de bienvenida (Figura 19) y pasado unos segundos se muestra la pantalla mostrada en la Figura 20, el paciente debe proporcionar sus datos personales, incluyendo nombre, edad, género, correo electrónico y número de teléfono. Estos datos pueden ser almacenados en una base de datos para su identificación. El paciente debe ingresar sus datos a través del teclado y, una vez que haya completado el proceso correctamente, debe hacer clic en el botón continuar.

A continuación se encontrará con la etapa de familiarización del sonido de prueba. Durante esta etapa, se reproducirá un sonido para que el paciente se familiarice con los tonos utilizados en la audiometría. El paciente tendrá la opción de repetir el sonido tantas veces como sea necesario. Como se muestra en la Figura 21, la interfaz cuenta con un botón para repetir el sonido y otro botón para avanzar a la siguiente etapa de la prueba.

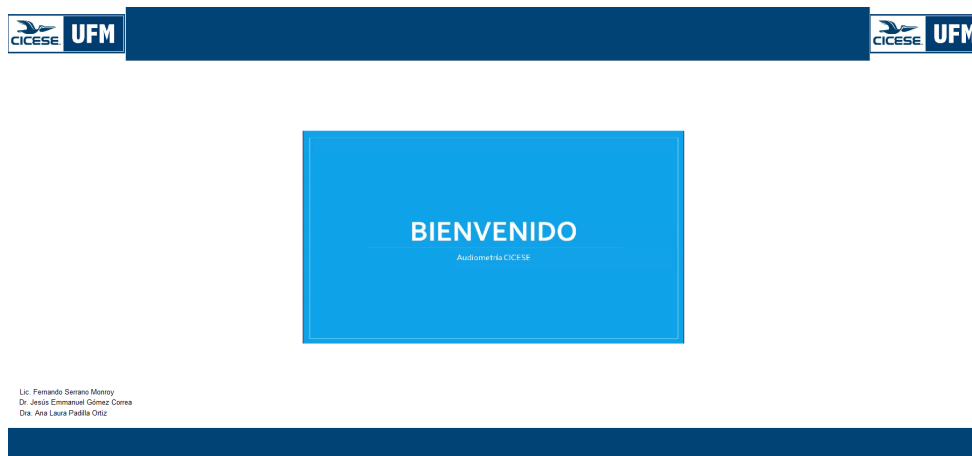


Figura 19. Presentación de bienvenida.

The image shows a form titled "Audiometría" on a blue background. The form has a dark blue header with the CICESE and UFM logos on the left and right. The form itself is white and contains several input fields: "Nombre completo", "Teléfono", "Edad", "Género", and "Correo electrónico". Below the fields, there is a button that says "Haga clic en continuar" and a "Continuar" button.

Figura 20. Datos personales que la persona debe ingresar antes de la prueba.

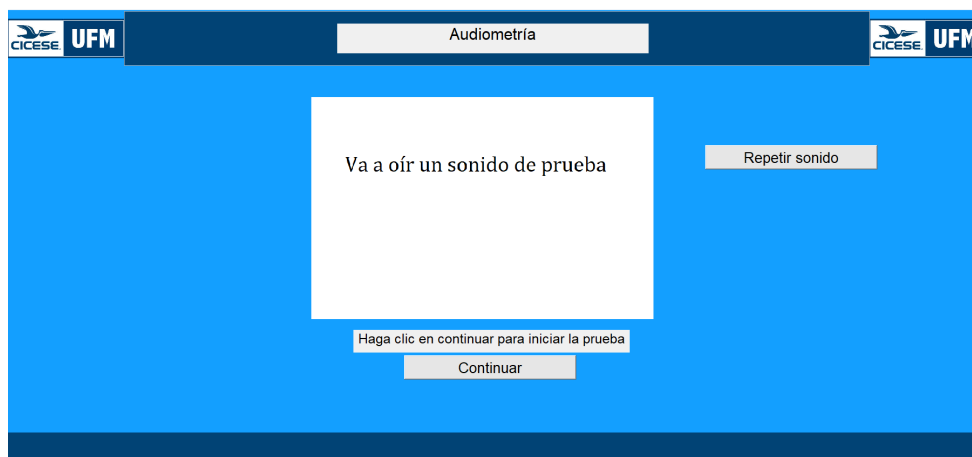


Figura 21. Visualización de la etapa de familiarización.

La siguiente etapa de la prueba consiste en leer las instrucciones. En esta sección, se le indicará al paciente qué hacer en caso de escuchar los sonidos durante la prueba, así como las instrucciones en caso de no escucharlos. Además, se le explicará cómo tomar las fotos de su mano abierta y cerrada, siguiendo las instrucciones que se mostrarán en pantalla (Figura 22) También se proporcionarán videos para que el paciente pueda visualmente comprender lo que debe hacer. Se le indica al paciente cuando va a comenzar la prueba, como lo muestra la Figura 23. La última sección consiste en obtener los umbrales auditivos.

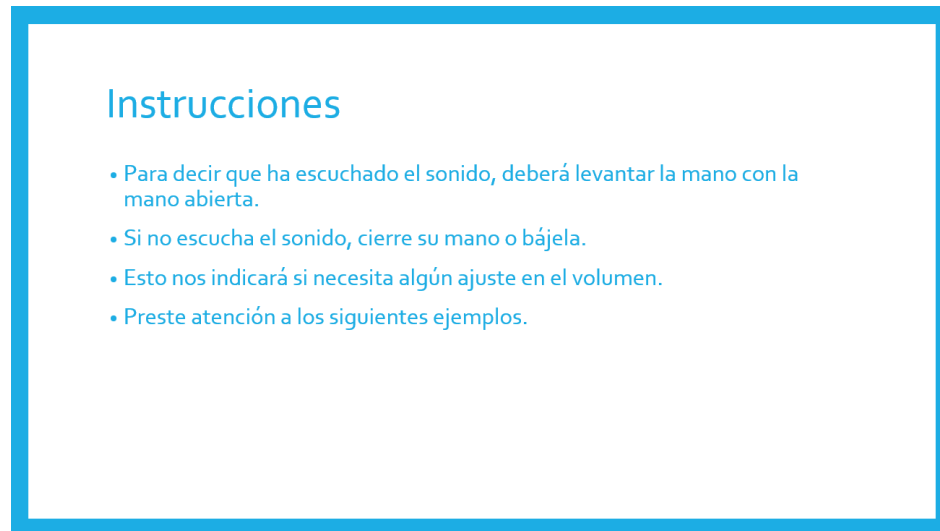


Figura 22. Instrucciones durante la prueba.

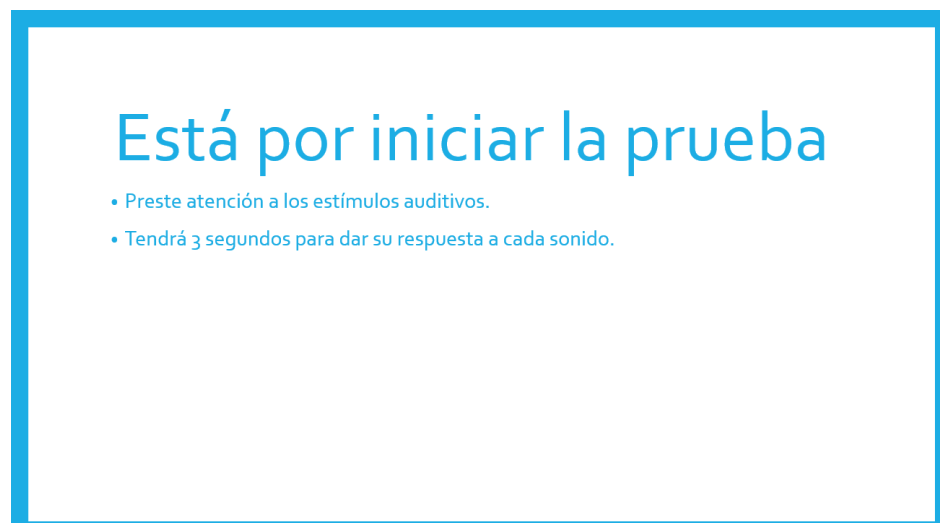


Figura 23. Indicación de que dará inicio a la prueba de audición.

El paciente escuchará durante dos segundos un sonido, simultáneamente aparecerá en pantalla el mensaje ¿Escucha? (Figura 24), luego en pantalla se estará procesando en un tiempo de 3 segundos aproxima-

damente la respuesta de la persona. En los tres segundos se analizarán 15 imágenes para clasificarla en alguna de las dos respuestas, la respuesta final se obtendrá de la respuesta que más se repitió.

Cuando se muestra el audiograma en la pantalla significa que la prueba ha terminado. Además, se brinda al usuario la opción de guardar sus resultados.

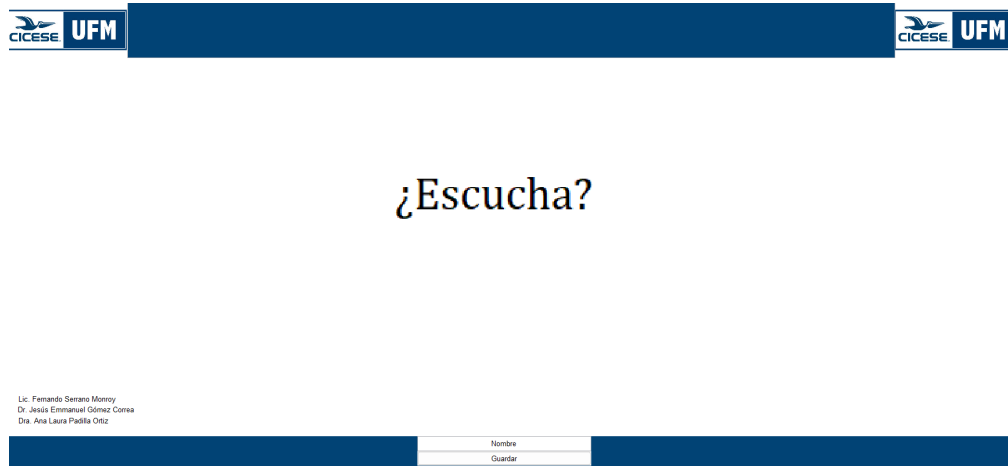


Figura 24. Mensaje simultáneo al estímulo auditivo.

4.4. Pruebas objetivas

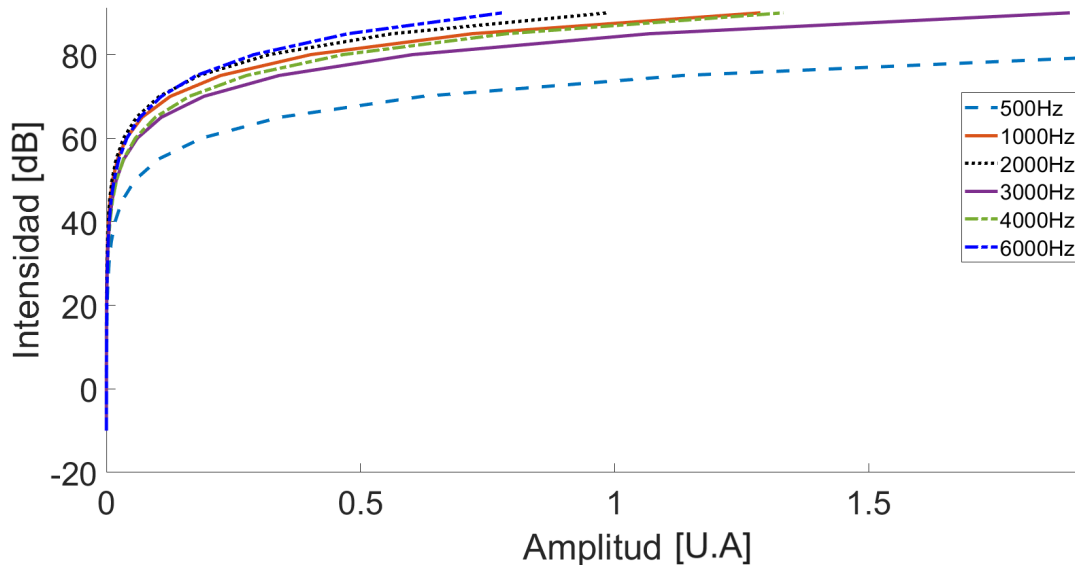
En esta sección se presentan los resultados de las pruebas objetivas, las cuales consistieron en obtener mediciones de intensidad de sonido mediante el uso de un sonómetro. Estas mediciones se compararon con los valores reales de intensidad que debería emitir el audiómetro automatizado.

Durante el proceso de calibración, se utilizó un sonómetro de clase 2, específicamente el modelo SM-10 de la marca Amprobe. Este sonómetro es adecuado para mediciones de campo y cuenta con una resolución de 0.1 dB, una ponderación de tiempo rápido de 125 ms, un rango de frecuencia desde 31.5 Hz hasta 8 kHz y una ponderación A de 30 a 130 dB. Además, presenta una precisión de ± 1.5 dB.

Para llevar a cabo las pruebas audiométricas, se utilizó la interfaz de audio Behringer U-Phoria UMC202HD, la cual posee características técnicas de Audiophile 2x2 y una resolución de 24 bits/192 kHz en audio USB. Dado que durante la audiometría no habría una persona controlando el audiómetro, se mantuvo la perilla del volumen de la interfaz de audio fijo en dirección horizontal hacia la izquierda.

Tabla 5. Coeficientes a y b, obtenidos a partir de la ecuación de amplitud.

Frecuencia [Hz]	500	1000	2000	3000	4000	6000
Coeficiente a	76.9576	87.8330	90.1249	84.4211	87.2755	92.5297
Coeficiente b	8.3284	8.5967	8.8877	8.7249	9.5170	10.1248

**Figura 25.** Interpolación de las mediciones para las frecuencias audiométricas, donde la amplitud tiene unidades arbitrarias (U.A).

De la Figura 25 las curvas corresponden al ajuste de la función que mejor se aproxima a los datos de las mediciones realizadas con el sonómetro. Cada curva corresponde a una frecuencia diferente. Recordemos que la función utilizada corresponde a la ecuación 23 del Capítulo 3. Los coeficientes 'a' y 'b' utilizados para modelar cada frecuencia se muestran en la Tabla 5.

4.4.1. Error absoluto

En la evaluación objetiva, el objetivo es medir el valor real de intensidad emitido por el audiómetro y compararlo con el valor fijo de intensidad que debería emitir. El error absoluto nos dice cuan preciso es el audiómetro digital. En la Figura 26, se muestra el error de cada frecuencia para cada nivel de intensidad, desde 45 hasta 90 dB con pasos de 5 dB, únicamente por limitaciones del sonómetro la frecuencia de 6000 Hz tiene muestras entre 55 y 90 dB. Un valor de error pequeño indica que las mediciones del

audiómetro son consistentes con el valor real de intensidad emitido.

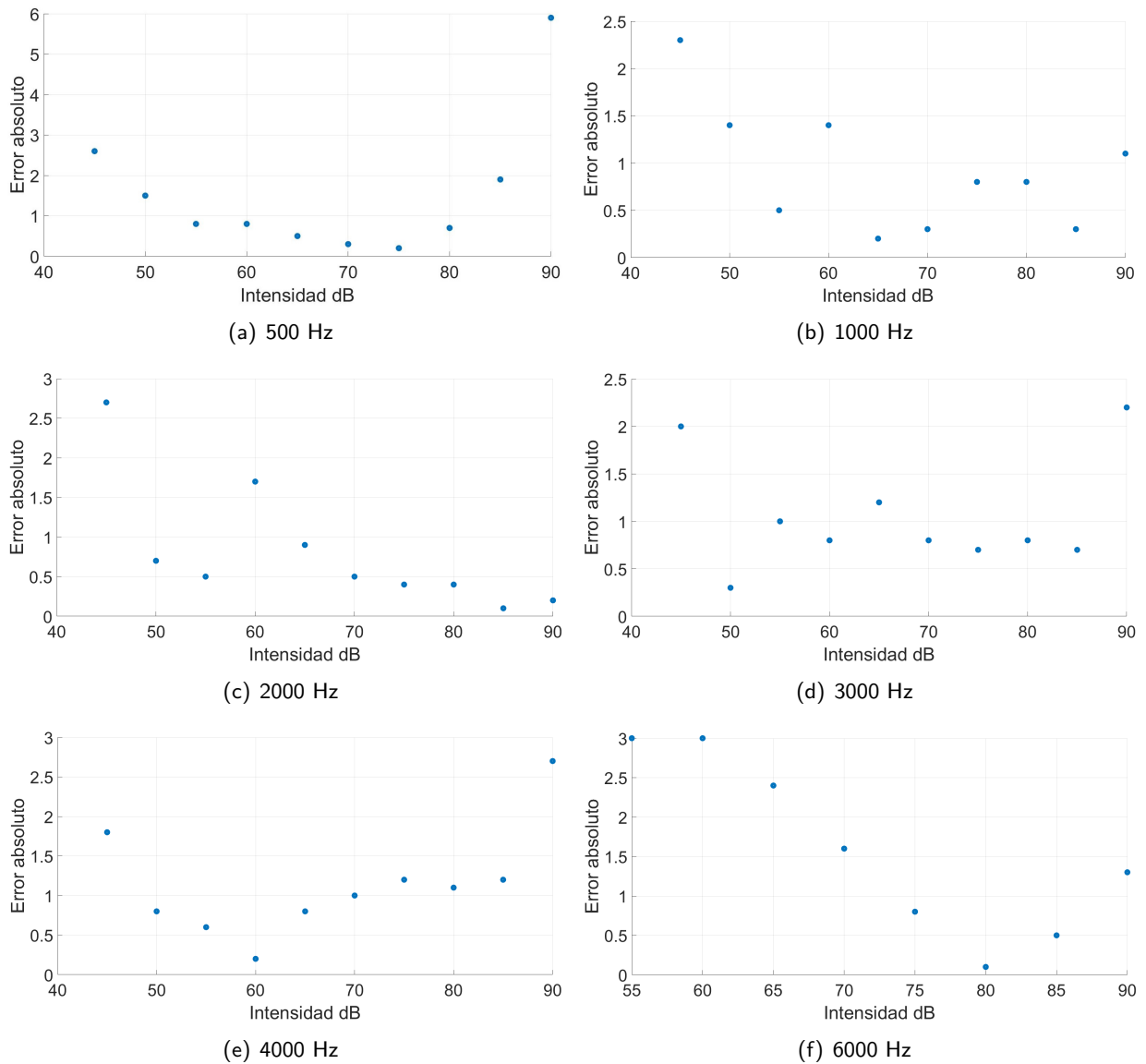


Figura 26. Errores absolutos medido para diferentes frecuencias audiométricas.

4.5. Audiometrías

En esta sección se muestran los resultados de algunas audiometrías realizadas con el audiómetro automatizado. También se presentan dos audiogramas obtenidos de pruebas audiométricas realizadas con un audiómetro comercial con el fin de hacer una comparación con los resultados obtenidos.

En la Figura 27(a) y 27(b) se muestran las audiometrías que fueron obtenidas mediante el uso de un

audiómetro de la marca Brüel & Kjaer modelo 1800 de registro automático, con una atenuación de -20 dB. Este audiómetro tiene la capacidad de reproducir 7 frecuencias, con un intervalo de intensidad que va desde -10 hasta 90 dB, y un atenuador que puede variar desde -30 hasta 70 dB.

El funcionamiento del audiómetro consiste en que el paciente, a través de los audífonos, debe escuchar estímulos auditivos y mantener presionado un botón mientras los escucha. Debe soltar el botón cuando ya no pueda escuchar el sonido. Simultáneamente, el audiómetro registra las mediciones mediante el movimiento de una pluma sobre el audiograma.

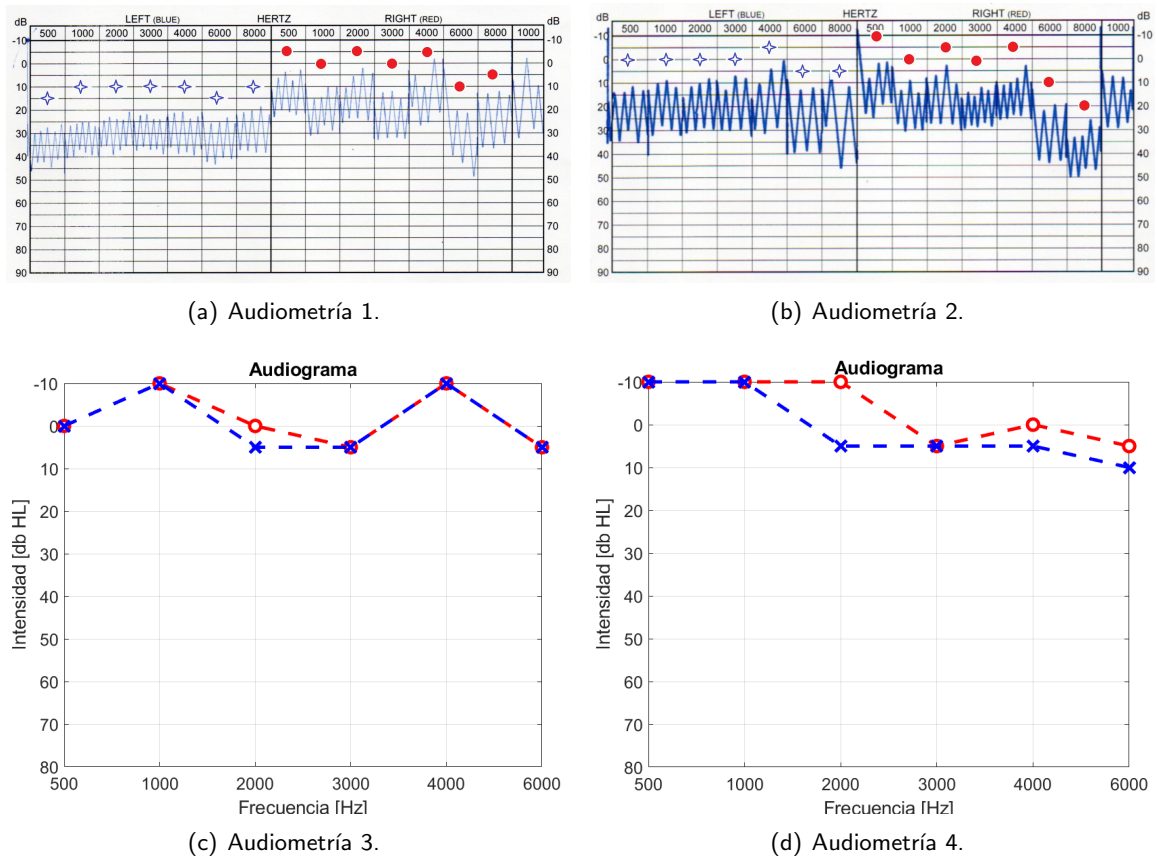


Figura 27. Audiometría del participante AP. Los audiogramas (a) y (b) se realizaron con el audiómetro B&K modelo 1800. Los audiogramas (c) y (d) se realizaron con el audiómetro desarrollado en este proyecto.

Una característica de este audiómetro es que la intensidad del volumen varía de forma continua, a diferencia del audiómetro desarrollado en esta tesis, que tiene una variación discreta. En las Figuras 27(c) y 27(d) se presentan los resultados obtenidos con el audiómetro automatizado. Todas las audiometrías fueron realizadas por el participante AP. Las audiometrías del participante AP fueron utilizadas como datos de referencia para la calibración.

Finalmente, en la Figura 28 se presentan los resultados de dos audiometrías realizadas por dos participantes. El primer participante ha manifestado no tener problemas para escuchar o comunicarse con otras personas, mientras que el segundo participante ha mencionado experimentar dificultades para escuchar sonidos y entender algunas palabras durante conversaciones uno a uno. Sin embargo, en estos casos no tenemos alguna audiometría de referencia realizada con algún audiómetro comercial con la que se puedan comparar los resultados.

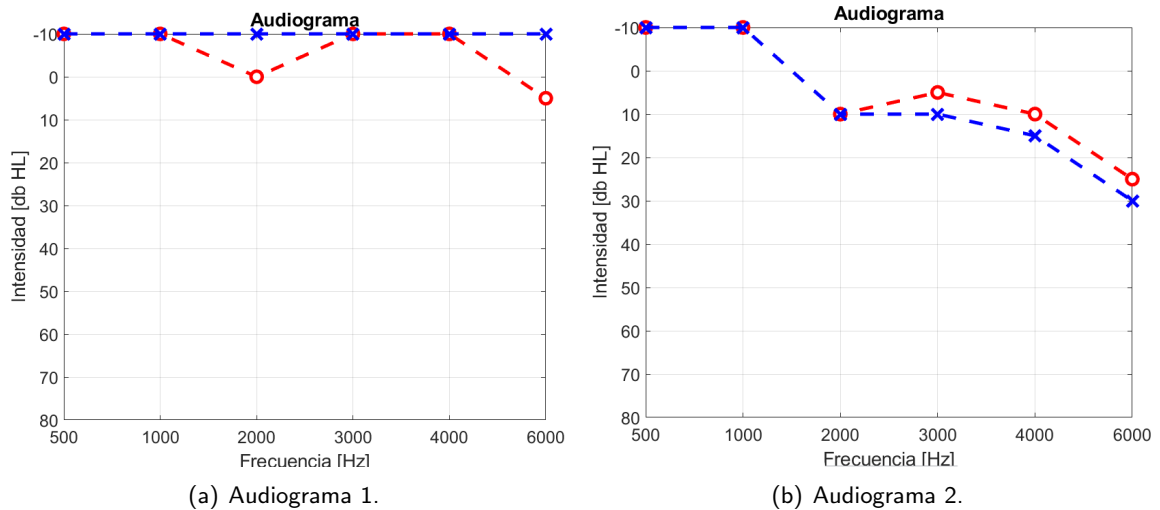


Figura 28. La audiometría 1 fue realizado por el participante EG y la audiometría 2 por el participante RL

Capítulo 5. Discusión

Dentro de los objetivos específicos de esta tesis, se incluye la identificación y selección de técnicas de reconocimiento de patrones. La metodología desarrollada implica la creación de un modelo de reconocimiento que prescinde de la necesidad de una extensa cantidad de datos. Para lograr esto, se utilizó un filtro de identificación de tonos de piel en la etapa de segmentación de imágenes. Además, la técnica de correlación se empleó para comparar la similitud entre dos imágenes, una correspondiente al gesto de la mano abierta y la otra a la mano cerrada. Esto significa que no se requirió una base de datos previa de manos para construir el modelo. Esta estrategia es posible gracias al sistema de visión computacional que captura ambos gestos en cada nueva prueba realizada.

Los resultados obtenidos en las matrices de confusión resaltan que las imágenes obtuvieron mejores resultados cuando había coincidencia en el fondo de ambas. En situaciones donde no había de fondo ventanas o fuentes luminosas, se obtuvo mayor exactitud (99 %). Es importante mencionar que el cambio en la distancia entre la cámara y la mano conlleva implicaciones en los patrones mediante la transformada de Fourier, lo cual incide en problemas para la correcta identificación de patrones. A fin de evitar errores en la clasificación, es fundamental considerar los siguientes aspectos durante la realización de las pruebas: mantener la cámara en una posición fija durante toda la duración de la prueba, mantener constante la distancia entre la cámara y el participante, y evitar la presencia de ventanas en el fondo. Hacer caso omiso a estos aspectos puede llevar a una mala predicción en las imágenes de la mano abierta, esto se comprueba en la figura 17 (a), donde se obtuvo una sensibilidad de 3 %.

Por otro lado, en relación a la evaluación del audiómetro, se analizó su capacidad para emitir intensidades de sonido en un rango de -10 a 90 dB con intervalos de 5 dB. Sin embargo, debido a ciertas limitaciones, las evaluaciones se llevaron a cabo a partir de los 45 dB. De la Figura 26, se puede calcular el promedio del error absoluto para diversas frecuencias: 500 Hz con 1.5 dB, 1000 Hz con 0.9 dB, 2000 Hz con 0.8 dB, 3000 Hz con 1 dB, 4000 Hz con 1.1 dB y 6000 Hz con 1.5 dB. Las frecuencias de 1000 Hz y 2000 Hz tienen el error absoluto más bajo, 0.9 dB y 0.8 dB respectivamente, lo que indica una mayor precisión en las mediciones. En contraste, las frecuencias de 500 Hz y 6000 Hz presentan el error absoluto más alto. En general, el promedio del error absoluto se encuentra en un rango de 0.8 a 1.5 dB, lo que sugiere que las mediciones se sitúan dentro de un nivel aceptable de precisión.

En cuanto a las audiometrías realizadas por el participante AP, se observa una variación en los umbrales auditivos en las pruebas realizadas en diferentes días, tanto utilizando el audiómetro B & K como el audiómetro digital. Esta variación podría estar influenciada por factores como el ruido de fondo o el nivel de atención del participante, lo que conlleva a una ligera variabilidad en los resultados. En consecuencia,

el hecho de que las pruebas se realicen en distintos días y bajo diferentes condiciones afecta la percepción auditiva del participante.

Dos participantes más completaron la prueba. No tuvieron ninguna dificultad para entender las instrucciones y la encontraron fácil de realizar. Además, comentaron que no fue una prueba agotadora y que les tomó poco tiempo. La duración de la prueba ronda entre 10 y 12 minutos, y depende de las respuestas que vaya dando el participante.

El método de Hughson-Westlake implementado en el audiómetro se ejecutó correctamente. Sin embargo, es importante señalar que una posible fuente de error no está relacionada con la programación del método, sino con la etapa de captura de la respuesta de la mano. Si el sistema de reconocimiento de la mano comete algún error, los resultados obtenidos para determinar los umbrales podrían no ser precisos.

Capítulo 6. Conclusiones

En este estudio, se ha diseñado y desarrollado un audiómetro digital automatizado que mide los umbrales auditivos sin la necesidad de la intervención de un experto en el campo que manipule el audiómetro. Esto cumple con el objetivo general de la tesis. Los resultados mostraron que la parte del sistema de visión computacional para reconocer gestos de la mano, es un buen modelo bajo ciertas condiciones, para evitar errores en la clasificación se toma en cuenta los aspectos como mantener fija la cámara, evitar la presencia de ventanas en el fondo y mantener constante la distancia entre el participante y la cámara. Así mismo el sistema permite guiar a los participantes durante la prueba para poder llevar a cabo la audiometría. Respecto al audiómetro, la generación de tonos se emite correctamente a las frecuencias audiométricas programadas y las intensidades tienen errores aceptables entre 0.8 y 1.5 dB.

Se ha descrito el proceso de diseño y desarrollo del audiómetro digital automatizado, abarcando aspectos como el método de calibración utilizado, generación de tonos y método para encontrar los umbrales auditivos. Este método se basó en los resultados de dos audiometrías: una realizada con un audiómetro calibrado y otra con el audiómetro precalibrado, donde para cada frecuencia se realiza un ajuste para igualar los resultados.

Durante las audiometrías realizadas, no hubo interacción directa con los participantes. El método implementado para determinar los umbrales implica que las pruebas se realicen de manera estandarizada y pueda ser replicado.

La contribución de un audiómetro digital automatizado tiene como ventaja que puede ser accesible y portátil, ya que puede estar disponible en sistemas de cómputo que tengan una cámara web. Además, el resultado digital puede ser almacenado en el expediente clínico electrónico. Debido a que se trata de información de salud, es necesario cumplir con los requisitos de la NOM-024, la cual garantiza la privacidad y la integridad de los datos.

La automatización de las pruebas también podría ahorrar tiempo a los profesionales de la salud al no requerir su presencia directa durante la realización de las pruebas auditivas. Es importante recalcar que la interpretación de los resultados y el diagnóstico definitivo siguen siendo tareas de los especialistas. Su experiencia y conocimiento son importantes para analizar los datos obtenidos, considerar el historial clínico del paciente y brindar un diagnóstico preciso.

6.1. Limitaciones del estudio

Se han identificado algunas limitaciones en este estudio. Una limitación fue el uso de un sonómetro diseñado para mediciones en campo, en lugar de uno específicamente diseñado para entornos de laboratorio. A pesar de esto, se logró medir la intensidad del sonido emitido por los audífonos con cambios en intervalos de 5 dB. Además, el sonómetro utilizado fue capaz de medir las intensidades entre 45 y 90 dB. Otra limitación es la viabilidad económica para realizar evaluaciones auditivas con audiómetros comerciales para todos los participantes, las cuales permitan comparar los resultados con aquellos obtenidos con el audiómetro desarrollado en este proyecto. El costo de una audiometría oscila entre los 600 y 1,500 pesos, por lo que es poco probable encontrar personas que se hayan realizado este estudio, aunado a la poca información y/o prevención sobre el cuidado de la salud auditiva. En ese sentido, no es posible generalizar la eficiencia del sistema con los resultados de un solo participante. Además, el audiómetro no ha sido evaluado por ningún especialista, en ese sentido está limitado ya que no se cuenta con la validación médica requerida para este tipo de desarrollos.

6.2. Trabajo a futuro

Para futuros trabajos, se recomienda llevar a cabo un análisis más exhaustivo y realizar validaciones adicionales para confirmar la precisión y confiabilidad del audiómetro en diferentes condiciones clínicas. Sería beneficioso incluir un mayor número de participantes en las investigaciones futuras. Además, se sugiere diseñar una oreja artificial que permita medir intensidades de sonido más bajas, lo cual sería útil para el proceso de calibración. A pesar de que las mediciones son en su mayoría precisas, las frecuencias de 500 Hz y 6000 Hz parecen ser más propensas a errores en comparación con las demás. Esto podría indicar la necesidad de realizar ajustes o mejoras específicas en la capacidad de emisión de sonido en estas frecuencias para garantizar mediciones más precisas en el futuro.

Literatura citada

- Abburu, S. & Golla, S. B. (2015). Satellite image classification methods and techniques: A review. *International journal of computer applications*, 119(8), 20–25. <https://doi.org/10.5120/21088-3779>.
- American Speech-Language-Hearing Association (1990). Audiometric symbols. *ASHA*. <https://doi.org/10.1044/policy.GL1990-00006>.
- American Speech-Language-Hearing Association (2005). *Guidelines for manual pure-tone threshold audiometry*. ASHA. <https://doi.org/10.1044/policy.GL2005-00014>.
- Amma, K., Yaguchi, Y., Niitsuma, Y., Matsuzaki, T., & Oka, R. (2013). A comparative study of gesture recognition between rgb and hsv colors using time-space continuous dynamic programming. In *2013 International Joint Conference on Awareness Science and Technology Ubi-Media Computing (iCAST 2013 UMEDIA 2013)*, 185–191. <https://doi.org/10.1109/ICAwST.2013.6765431>.
- Bassett, I. G. & Eastmond, E. J. (1964). Echolocation: Measurement of pitch versus distance for sounds reflected from a flat surface. *The Journal of the Acoustical Society of America*, 36, 911–916. <https://doi.org/10.1121/1.1919117>.
- Beaudoin, N. & Beauchemin, S. S. (2002). An accurate discrete fourier transformation for image processing. *Proceedings - International Conference on Pattern Recognition*, 16, 935–939. <https://doi.org/10.1109/ICPR.2002.1048189>.
- Borja-Robalino, R., Monleon-Getino, A., & Rodellar, J. (2020). Estandarización de métricas de rendimiento para clasificadores machine y deep learning. *Revista Ibérica de Sistemas e Tecnologías de Informação*, (E30), 184–196. https://www.researchgate.net/publication/339943922_Estandarizacion_de_Metricas_de_Rendimiento_para_Clasificadores_Machine_y_Deep_Learning.
- Brigham, E. O. (1988). *The fast Fourier transform and its applications*. Prentice Hall.
- Cardier, M., Manrique, R., Huarte, A., Valencia, M. L., Borro, D., Calavia, D., & Manrique, M. (2016). Telemedicina. estado actual y perspectivas futuras en audiología y otología. *Revista Médica Clínica Las Condes*, 27, 840–847. <https://doi.org/10.1016/J.RMCLC.2016.11.016>.
- Carhart, R. & Jerger, J. F. (1959). Preferred method for clinical determination of pure-tone thresholds. *Journal of Speech and Hearing Disorders*, 24, 330–345. <https://doi.org/10.1044/JSHD.2404.330>.
- Castañeda, P. (1999). El lenguaje verbal del niño. *Lima: Fondo Editorial de la UNMSM*, Recuperado de <https://gc.scalahed.com/recursos/files/r161r/w23056w/E1%20lenguaje%20verbal%20del%20nino.pdf>.
- Chan, V. & Perlas, A. (2011). Basics of ultrasound imaging. *Atlas of Ultrasound-Guided Procedures in Interventional Pain Management*, 13–19. https://doi.org/10.1007/978-1-4419-1681-5_2.
- Chen, X. & Yuille, A. L. (2014). Articulated pose estimation by a graphical model with image dependent pairwise relations. In *Advances in Neural Information Processing Systems*, volumen 27. <https://api.semanticscholar.org/CorpusID:6619926>.
- Clark, J. (1981). Uses and abuses of hearing loss classification. *ASHA*, 23(7), 493–500. Recuperado de https://www.researchgate.net/publication/16145943_Uses_and_abuses_of_hearing_loss_classification.

- Contreras, D. A. A. & Gallegos, F. J. F. (2016). Comparación de dos técnicas propuestas hs-cbcr y hs-ab para el modelado de color de piel en imágenes. *Research in Computing Science*, 114, 33–44. Recuperado el 24 de julio del 2023 de https://rcs.cic.ipn.mx/2016_114/Comparacion%20de%20dos%20tecnicas%20propuestas%20HS--CbCr%20y%20HS--ab%20para%20el%20modelado%20de%20color%20de%20piel.pdf.
- Duhamel, D. (1996). Efficient calculation of the three-dimensional sound pressure field around a noise barrier. *Journal of Sound and Vibration*, 197(5), 547–571. <https://doi.org/10.1006/jsvi.1996.0548>.
- Durrant, J. D., Lovrinic, J. H., & Vernon, J. (1985). Bases of Hearing Science. *The Journal of the Acoustical Society of America*, 77(5), 1973–1974. <https://doi.org/10.1121/1.391810>.
- D’Orazio, T. & Leo, M. (2010). A review of vision-based systems for soccer video analysis. *Pattern Recognition*, 43(8), 2911–2926. <https://doi.org/10.1016/j.patcog.2010.03.009>.
- Epstein, A. (1960). Xi variables involved in automatic audiometry. *Annals of Otology, Rhinology Laryngology*, 69, 137–141. https://doi.org/10.1177/000348946006900111/ASSET/000348946006900111.FP.PNG_V03.
- Fairman, H. S., Brill, M. H., & Hemmendinger, H. (1997). How the cie 1931 color-matching functions were derived from wright-guild data. *Color Research & Application*, 22(1), 11–23. [https://doi.org/10.1002/\(SICI\)1520-6378\(199702\)22:1<11::AID-COL4>3.0.CO;2-7](https://doi.org/10.1002/(SICI)1520-6378(199702)22:1<11::AID-COL4>3.0.CO;2-7).
- Ferrari, S., Silva, M., Guarino, M., & Berckmans, D. (2008). Monitoring of swarming sounds in bee hives for early detection of the swarming period. *Computers and Electronics in Agriculture*, 64, 72–77. <https://doi.org/10.1016/J.COMPAG.2008.05.010>.
- Fletcher, H. & Munson, W. A. (1933). Loudness, its definition, measurement and calculation*. *Bell System Technical Journal*, 12, 377–430. <https://doi.org/10.1002/J.1538-7305.1933.TB00403.X>.
- Fletcher, S. H. (1992). *Harvey Fletcher 1884—1981*. National academy of sciences. <https://www.nasonline.org/publications/biographical-memoirs/memoir-pdfs/fletcher-harvey.pdf>.
- Godfrey, K. (1980). Correlation methods. *Automatica*, 16(5), 527–534. [https://doi.org/10.1016/0005-1098\(80\)90076-X](https://doi.org/10.1016/0005-1098(80)90076-X).
- Gonzalez, R. C. & Woods, R. E. (2007). *Digital Image Processing (3.^a ed.)*. Pearson.
- Hongo, H., Ohya, M., Yasumoto, M., & Yamamoto, K. (2000). Face and hand gesture recognition for human-computer interaction. 2, 921–924. <https://doi.org/10.1109/ICPR.2000.906224>.
- Instituto de Salud para el Bienestar (2023). Día mundial de la audición. <http://www.gob.mx/insabi/articulos/dia-mundial-de-la-audicion>.
- Jerger, J. (2017). C.c. bunch: The first audiologist. *Audiology today*. <https://www.audiology.org/news-and-publications/audiology-today/articles/c-c-bunch-the-first-audiologist/>.
- Kakani, V., Nguyen, V. H., Kumar, B. P., Kim, H., & Pasupuleti, V. R. (2020). A critical review on computer vision and artificial intelligence in food industry. *Journal of Agriculture and Food Research*, 2, 100033. <https://doi.org/10.1016/J.JAFR.2020.100033>.
- Kapul, A. A., Zubova, E. I., Torgaev, S. N., & Drobchik, V. V. (2017). Pure-tone audiometer. *Journal of Physics: Conference Series*, 881(1). <https://doi.org/10.1088/1742-6596/881/1/012010>.

- Kjeldsen, R. & Kender, J. (1996). Finding skin in color images. *Proceedings of the International Conference on Automatic Face and Gesture Recognition*, 312–317. <https://doi.org/10.1109/AFGR.1996.557283>.
- Kumar, N., Belhumeur, P. N., Biswas, A., Jacobs, D. W., Kress, W. J., Lopez, I. C., & Soares, J. V. (2012). Leafsnap: A computer vision system for automatic plant species identification. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 7573 LNCS, 502–516. https://doi.org/10.1007/978-3-642-33709-3_36.
- Kuruville, J., Sukumaran, D., Sankar, A., & Joy, S. P. (2016). A review on image processing and image segmentation. *Proceedings of 2016 International Conference on Data Mining and Advanced Computing, SAPIENCE 2016*, 198–203. <https://doi.org/10.1109/SAPIENCE.2016.7684170>.
- Liu, J., Sun, J., & Wang, S. (2006). Pattern recognition: An overview. *IJCSNS International Journal of Computer Science and Network Security*, 6(6), 57–61. http://paper.ijcsns.org/07_book/200606/200606A10.pdf.
- Liu, Q. & zheng Peng, G. (2010). A robust skin color based face detection algorithm. *2010 2nd International Asia Conference on Informatics in Control, Automation and Robotics (CAR 2010)*, 2, 525–528. <https://doi.org/10.1109/CAR.2010.5456614>.
- Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 3431–3440. https://openaccess.thecvf.com/content_cvpr_2015/papers/Long_Fully_Convolutional_Networks_2015_CVPR_paper.pdf.
- Maglogiannis, I. & Doukas, C. N. (2009). Overview of advanced computer vision systems for skin lesions characterization. *IEEE Transactions on Information Technology in Biomedicine*, 13, 721–733. <https://doi.org/10.1109/TITB.2009.2017529>.
- Mairal, J., Bach, F., Ponce, J., Sapiro, G., & Zisserman, A. (2008). Discriminative learned dictionaries for local image analysis. *26th IEEE Conference on Computer Vision and Pattern Recognition, CVPR*. <https://doi.org/10.1109/CVPR.2008.4587652>.
- Meyer-Bischoff, C. (2005). Les chiffres du bruit. *médecine/sciences*, 21, 546–550. <https://doi.org/10.1051/MEDSCI/2005215546>.
- Nasirahmadi, A., Edwards, S. A., & Sturm, B. (2017). Implementation of machine vision for detecting behaviour of cattle and pigs. *Livestock Science*, 202, 25–38. <https://doi.org/10.1016/j.livsci.2017.05.014>.
- Nasteski, V. (2017). An overview of the supervised machine learning methods. *Horizons. b*, 4, 51–62. Recuperado de https://www.researchgate.net/publication/328146111_An_overview_of_the_supervised_machine_learning_methods.
- Neupane, D. & Seok, J. (2020). A review on deep learning-based approaches for automatic sonar target recognition. *Electronics 2020, Vol. 9, Page 1972*, 9, 1972. <https://doi.org/10.3390/ELECTRONICS9111972>.
- Organización Mundial de la Salud (2023). Sordera y pérdida de la audición. Recuperado de <https://www.who.int/es/news-room/fact-sheets/detail/deafness-and-hearing-loss>.
- Ouyang, W., Zeng, X., Wang, X., Qiu, S., Luo, P., Tian, Y., Li, H., Yang, S., Wang, Z., Li, H., Wang, K., Yan, J., Loy, C. C., & Tang, X. (2017). Deepid-net: Object detection with deformable part based convolutional neural networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39, 1320–1334. <https://doi.org/10.1109/TPAMI.2016.2587642>.

- Phung, S. L., Bouzerdoum, A., & Chai, D. (2005). Skin segmentation using color pixel classification: Analysis and comparison. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27, 148–154. <https://doi.org/10.1109/TPAMI.2005.17>.
- Real Academia Española (2022). Diccionario de la lengua española. (versión 23.6 en línea). recuperado el 13 de junio del 2023 en <https://dle.rae.es/diapas%C3%B3n>.
- Resnick, R., Halliday, D., & Krane, K. (2005). *Física*. (5ª ed., Vol. 1). Patria.
- Salud auditiva integral (2018). *Audiometría*. Recuperado de https://www.facebook.com/SaludAuditivaIntegral/photos/audiometrial-audiometr%C3%ADa-cl%C3%ADnica-o-audiometr%C3%ADa-tonal-determina-el-grado-de-audi/1844691232266091/?paipv=0&eav=AfYgVmYXv0bXC1C_7HPMcasmjcZip09oMD4bt3WBnTrHdu8FmP90iioquSotQU5Wq1c&_rdr.
- Schomberg, H. & Timmer, J. (1995). The gridding method for image reconstruction by fourier transformation. *IEEE Transactions on Medical Imaging*, 14(3), 596–607. <https://doi.org/10.1109/42.414625>.
- Secretaría de Salud (2021). 530. con discapacidad auditiva, 2.3 millones de personas: Instituto Nacional de Rehabilitación. <https://www.gob.mx/salud/prensa/530-con-discapacidad-auditiva-2-3-millones-de-personas-instituto-nacional-de-rehabilitacion?idiom=es>.
- Sharma, P. & Kaur, M. (2013). Classification in pattern recognition: A review. *International Journal of Advanced Research in Computer Science and Software Engineering*, 3(4), 298–306. Recuperado de <https://www.utm.mx/--jjf/rp/RPA3.pdf>.
- Shearer, E., Hildebrand, M. S., Schaefer, A. M., & Smith, R. J. H. (1999). Genetic hearing loss overview. *GeneReviews*. <https://www.ncbi.nlm.nih.gov/books/NBK1434/>.
- Sucar, L. E. & Gómez, G. (2011). Visión computacional. *Instituto Nacional de Astrofísica, Óptica y Electrónica. Puebla, México*, Recuperado de https://www.researchgate.net/publication/267295870_Vision_Computacional.
- Suzuki, Y. & Takeshima, H. (2004). Equal-loudness-level contours for pure tones. *The Journal of the Acoustical Society of America*, 116, 918–933. <https://doi.org/10.1121/1.1763601>.
- Tajbakhsh, N., Shin, J. Y., Gurudu, S. R., Hurst, R. T., Kendall, C. B., Gotway, M. B., & Liang, J. (2016). Convolutional neural networks for medical image analysis: Full training or fine tuning? *IEEE Transactions on Medical Imaging*, 35(5), 1299–1312. <https://doi.org/10.1109/TMI.2016.2535302>.
- Tao, R., Deng, B., & Wang, Y. (2006). Research progress of the fractional fourier transform in signal processing. *Science in China Series F*, 49, 1–25. <https://doi.org/10.1007/s11432-005-0240-y>.
- Tippens, P. E. (2007). *Física: conceptos y aplicaciones*. McGraw-Hill.
- van Tonder, J., Swanepoel, D. W., Mahomed-Asmail, F., Myburgh, H., & Eikelboom, R. H. (2017). Automated smartphone threshold audiometry: validity and time efficiency. *Journal of the American Academy of Audiology*, 28, 200–208. <https://doi.org/10.3766/jaaa.16002>.
- Vera, M., Huérfano, Y., Contreras, J., Valbuena, O., Chacón, J., Vera, M., et al. (2015). Técnicas de preprocesamiento de imágenes cardíacas: fundamentos y alcances. *Revista Latinoamericana de Hipertensión*, 11(3), 60–66. Recuperado de <https://search.ebscohost.com/login.aspx?direct=true&db=lth&AN=122248459&lang=es&site=eh--ost--live>.

- Wright, J., Ma, Y., Mairal, J., Sapiro, G., Huang, T. S., & Yan, S. (2010). Sparse representation for computer vision and pattern recognition. *Proceedings of the IEEE*, 98, 1031–1044. <https://doi.org/10.1109/JPROC.2010.2044470>.
- Xiao, D. & Ohya, J. (2007). Contrast enhancement of color images based on wavelet transform and human visual system. In *Proceedings of the IASTED International Conference on Graphics and Visualization in Engineering*, 58–63. Recuperado de <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=32b8d7fe317ed9defebf49e19d875f768e2397dd>.
- Yousefi, J. (2011). Image binarization using otsu thresholding algorithm. *University of Guelph*. <https://doi.org/10.13140/RG.2.1.4758.9284>.
- Zhang, B., Wu, X., You, J., Li, Q., & Karray, F. (2010). Detection of microaneurysms using multi-scale correlation coefficients. *Pattern Recognition*, 43(6), 2237–2248. <https://doi.org/https://doi.org/10.1016/j.patcog.2009.12.017>.