

# **ESCUELA POLITÉCNICA NACIONAL**

**FACULTAD DE INGENIERÍA ELÉCTRICA Y  
ELECTRÓNICA**

**DESARROLLO DE UN TOOLBOX EN MATLAB PARA REALIZAR  
EXPERIMENTOS PSICOACÚSTICOS PARA LA LOCALIZACIÓN  
DE FUENTES SONORAS**

**TRABAJO DE INTEGRACIÓN CURRICULAR PRESENTADO COMO  
REQUISITO PARA LA OBTENCIÓN DEL TÍTULO DE INGENIERO EN  
TELECOMUNICACIONES**

**EDGAR GIOVANNI TORRES ERAS**  
edgar.torres01@epn.edu.ec

**DIRECTOR: Ph.D. FELIPE LEONEL GRIJALVA ARÉVALO**  
felipe.grijalva@epn.edu.ec

**DMQ, OCTUBRE 2022**

## CERTIFICACIONES

Yo, EDGAR GIOVANNI TORRES ERAS declaro que el trabajo de integración curricular aquí descrito es de mi autoría; que no ha sido previamente presentado para ningún grado o calificación profesional; y, que he consultado las referencias bibliográficas que se incluyen en este documento.



---

**EDGAR GIOVANNI TORRES ERAS**

Certifico que el presente trabajo de integración curricular fue desarrollado por EDGAR GIOVANNI TORRES ERAS, bajo mi supervisión.



---

**Ph.D. FELIPE LEONEL GRIJALVA ARÉVALO  
DIRECTOR DEL TRABAJO DE TITULACIÓN**

## **DECLARACIÓN DE AUTORÍA**

A través de la presente declaración, afirmamos que el trabajo de integración curricular aquí descrito, así como el producto resultante del mismo, son públicos y estarán a disposición de la comunidad a través del repositorio institucional de la Escuela Politécnica Nacional; sin embargo, la titularidad de los derechos patrimoniales nos corresponde a los autores que hemos contribuido en el desarrollo del presente trabajo; observando para el efecto las disposiciones establecidas por el órgano competente en propiedad intelectual, la normativa interna y demás normas.

EDGAR GIOVANNI TORRES ERAS

Ph.D. FELIPE LEONEL GRIJALVA ARÉVALO

## DEDICATORIA

El presente proyecto está dedicado a mi esposa Silvana y mi hijo Edguitar quienes han sido mi fortaleza y mi inspiración, son quienes me dan las fuerzas necesarias para luchar y alcanzar mis metas.

A mi madre quien en cada etapa de mi vida ha estado a mi lado apoyándome, a mi padre quien me ha forjado en un hombre de bien, enseñándome a superar las adversidades.

A mi abuelito Victoriano Eras que supo darme fuerzas para seguir adelante y no desmayar ante los problemas.

***Edgar Torres***

## **AGRADECIMIENTOS**

A Dios por haberme dado la vida y poder disfrutar de una familia maravillosa, y permitirme haber llegado hasta este momento tan importante de mi formación profesional.

A toda mi familia, amigos por darme su apoyo y el ánimo necesario para continuar y salir adelante día a día a pesar de las adversidades.

A mi director del proyecto Ph.D. FELIPE LEONEL GRIJALVA ARÉVALO quien con sus conocimientos, virtudes, paciencia y constancia me ayudo a culminar este trabajo, y sus consejos fueron siempre útiles.

***Edgar Torres***

# ÍNDICE DE CONTENIDO

CERTIFICACIONES .....	I
DECLARACIÓN DE AUTORÍA.....	II
DEDICATORIA .....	III
AGRADECIMIENTO .....	IV
ÍNDICE DE CONTENIDO.....	V
RESUMEN.....	VI
ABSTRACT .....	VIII
1. INTRODUCCIÓN .....	1
1.1. Objetivo general.....	2
1.2. Objetivos específicos .....	2
1.3. Alcance .....	2
1.4. MARCO TEÓRICO.....	4
1.4.1. Audio Espacial .....	4
1.4.2. Sistema de Coordenadas .....	4
1.4.3. Percepción auditiva .....	5
1.4.4. HRTF .....	7
1.4.5. Características en el dominio del tiempo de las HRIRs .....	10
1.4.6. Características en el dominio de la frecuencia de las HRTFs.....	11
1.4.7. Características en el dominio de la frecuencia debido al pabellón auricular .	11
1.4.8. Formatos para bases de datos de HRTFs .....	13
1.4.9. Tipos de Experimentos Psicoacústicos .....	17

2. METODOLOGÍA.....	18
2.1. Lectura de HRTFs .....	18
2.1.1. Interfaz inicial .....	19
2.2. Procesamiento .....	21
2.3. Selección del experimento psicoacústico .....	22
2.3.1. Experimentos de localización absoluta .....	22
2.3.2. Experimentos de localización relativa .....	24
2.4. Sintetizador de audio espacial.....	25
2.5. Interfaz gráfica de usuarios.....	27
3. RESULTADOS, CONCLUSIONES Y RECOMENDACIONES .....	29
3.1. RESULTADOS .....	29
3.1.1. Experimentos de localización absoluta .....	29
3.1.2. Experimentos de localización absoluta .....	30
3.1.3. Estadísticas .....	33
3.2. CONCLUSIONES.....	36
3.3. RECOMENDACIONES.....	37
4. REFERENCIAS BIBLIOGRÁFICAS .....	38
5. ANEXOS.....	40

# RESUMEN

El audio espacial binaural tiene como objetivo simular fuentes de sonidos en una determinada ubicación mediante el uso de Funciones de Transferencia Relacionadas a la Cabeza ("*Head-Related Transfer Function*, HRTF") causado por la dispersión del sonido alrededor de la cabeza, el pabellón auricular y el torso antes que ingrese al canal auditivo. Las HRTFs son diferentes para cada persona por lo que, si se aplican a un cierto número de personas pueden obtener distintos resultados e información errónea. Para obtener HRTFs efectivas para cada persona es necesario medirlas, para ellos se han desarrollado varios métodos, unos usando equipamiento como la cámara anecoica que consiste en un arreglo específico de parlantes y micrófonos, otros mediante modelos que intentan simular la forma de la cabeza humana otros con métodos numéricos y varios más. Esta tarea resulta compleja y costosa y se vuelve poco accesible para las personas que quieran incursionar en este campo.

Un método adicional de medición de HRTFs es a través de la escucha y obteniendo datos, por esta razón el presente trabajo de titulación desarrolla un *toolbox* en Matlab con una interfaz gráfica para realizar los experimentos psicoacústicos con diferentes HRTFs que garanticen audio 3D de alta calidad. El *toolbox* utiliza archivos en formato SOFA (*Spatially Oriented Format for Acoustics*) que son de libre acceso y se encuentran disponibles en los repositorios oficiales. El *toolbox* está diseñado para realizar 4 experimentos: 2 de localización relativa y 2 de localización absoluta en los planos horizontal, mediano, frontal. Este *toolbox* tiene una interfaz gráfica que permite interactuar con el usuario y su uso no es complejo por lo que cualquier persona puede trabajar con él. Una vez realizados los experimentos éste permite obtener datos estadísticos e información de los usuarios para su debido análisis ya sea en forma gráfica o en hojas de cálculo.

**PALABRAS CLAVE:** Audio Espacial, HRTF, MATLAB, Interfaz gráfica, Toolbox, SOFA



## ABSTRACT

Binaural spatial audio aims to simulate sound sources in a certain location by using Head-Related Transfer Functions (*HRTF*) caused by the scattering of sound around the head, pinna, and torso before it enters the ear canal. HRTFs are different for each person so if they are applied to a certain number of people they can get different results and misinformation. To obtain effective HRTFs for each person it is necessary to measure them, for them several methods have been developed, some using equipment such as the aneconic camera that consists of a specific arrangement of speakers and microphones, others through models that try to simulate the shape of the human head others with numerical methods and several more. This task is complex and expensive and becomes inaccessible to people who want to venture into this field.

One additional method of measuring HRTFs is through listening and obtaining data, for this reason the present titling work develops a *toolbox* in Matlab with a graphical interface to perform psychoacoustic experiments with different HRTFs that guarantee high quality 3D audio. The *toolbox* uses files in SOFA (Spatially Oriented Format for Acoustics) format that are freely accessible and available in the official repositories. The *toolbox* is designed to perform 4 experiments: 2 of relative localization and 2 of absolute location in the horizontal, medium, frontal planes. This *toolbox* has a graphical interface that allows you to interact with the user and its use is not complex so anyone can work with it. Once the experiments have been carried out, it allows to obtain statistical data and information from the users for their due analysis either in graphic form or in spreadsheets.

**KEYWORDS:** Spatial Audio, HRTF, Matlab, Graphical Interface, Toolbox, SOFA

# 1. INTRODUCCIÓN

Las funciones de transferencia Relacionadas a la Cabeza *HRTFs*, codifican los tiempos y las señales de audio de tal manera que logran simular la localización del sonido debido a la diferencia de tiempo causada por la dispersión de éste alrededor de la cabeza, pabellón auricular y el torso antes que llegue al tímpano.

Las *HRTFs* difieren entre cada individuo, es decir son únicas y no se pueden utilizar para otra persona que no haya sido para la que se obtuvo, esto debido a las características morfológicas de cada persona. Por esta razón surge la necesidad de personalizarlas garantizando audio de alta calidad, ya que si no se personalizan dificultan la precisión de la localización del sonido causando confusiones en las distintas direcciones de la que éste puede provenir. Existen procedimientos para personalizar las *HRTFs*, una de ellas es a través de mediciones directas, sin embargo es un proceso lento, complejo, costoso y no escalable. A razón de esto se han propuesto varios métodos alternativos para evitar las mediciones incluyendo la estimación de *HRTFs* a partir de un pequeño conjunto de mediciones, además, hay modelos teóricos (por ejemplo, modelo de cabeza esférica, modelo de muñeco de nieve, modelos estructurales que intentan aproximarse a la compleja anatomía humana. [1]

Adicionalmente existen otros métodos numéricos (por ejemplo: método de elementos de contorno, método de diferencias finitas en el dominio del tiempo) las cuales requieren utilización de equipos costosos de alto rendimiento y cuyo uso es complejo. Tomando una dirección diferente, varios autores han propuesto métodos basados en la percepción, donde los sujetos eligen sus *HRTF* a través de pruebas de escucha ajustando algunos parámetros hasta que alcancen una precisión espacial aceptable [1]. Es en este campo donde el presente trabajo de titulación desarrollará un *toolbox* en Matlab que, mediante una interfaz gráfica realizarán distintos experimentos psicoacústicos utilizando las *HRTFs* en formato SOFA, se obtendrán datos mediante la escucha por parte de distintos usuarios y posteriormente serán analizados, ya que este *toolbox* permite obtener resultados distribuidos en tablas con todos los datos de los experimentos realizados, adicionalmente permitirá interpretar los datos obtenidos a través de gráficos.

## 1.1. Objetivo general

Desarrollar un *toolbox* en Matlab para realizar experimentos psicoacústicos de localización de fuentes sonoras.

## 1.2. Objetivos específicos

- Estudiar los conceptos básicos de audio espacial y de las HRTFs.
- Implementar los módulos del *toolbox* en Matlab para la realización de experimentos psicoacústicos de localización absoluta y relativa de fuentes de sonoras.
- Realizar pruebas de validación de los experimentos psicoacústicos usando el *toolbox* en un grupo de personas.

## 1.3. Alcance

Se desarrollará un *toolbox* en Matlab que permitirá realizar experimentos psicoacústicos de localización absoluta y relativa para fuentes sonoras, los mismos que serán validados en un grupo de personas (Ver Figura 1.1). Específicamente, se tiene los siguientes experimentos.

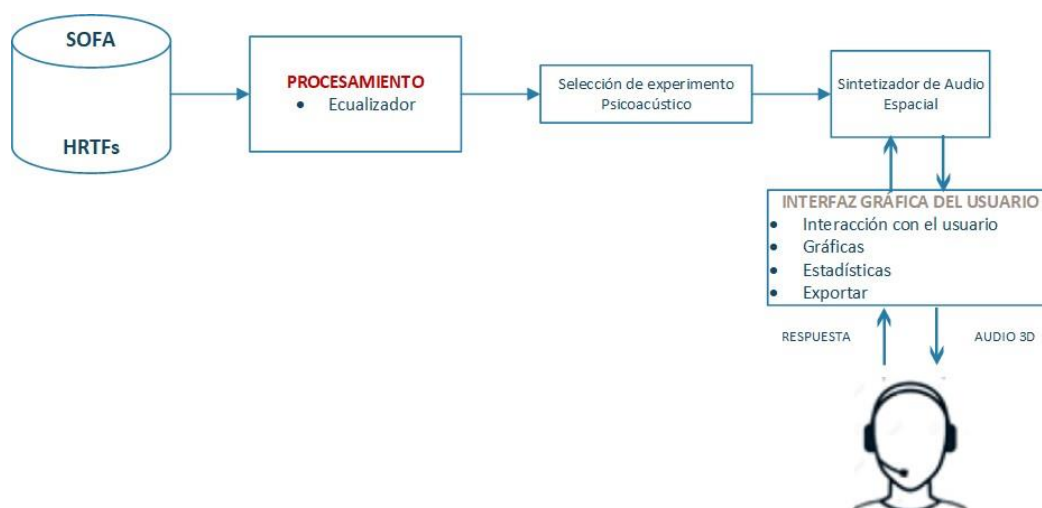
### 1. Localización Absoluta

- a) En este experimento el programa conocerá la posición en dónde se genera el sonido pero el usuario no. El oyente reportará en dónde escuchó el estímulo auditivo.
- b) En este experimento, el usuario dispondrá de una posición conocida de la fuente de audio. El programa presentará una serie de estímulos auditivos y el usuario seleccionará el estímulo que más se acerque a la posición conocida.

### 2. Localización Relativa

- a) El programa presentará dos estímulos auditivos y solicitará al oyente que identifique la posición de la segunda fuente de sonido con respecto a la primera (e.g. la posición del segundo estímulo está a la izquierda del primero).
- b) Experimento para determinar el mínimo ángulo perceptible. Para desarrollar este experimento, el programa presentará dos estímulos auditivos. El usuario responderá si escuchó o no los estímulos en la misma posición. En el caso de que

responda que si, el programa nuevamente generará dos sonidos separados un ángulo ligeramente mayor respecto a los dos anteriores y nuevamente preguntará al oyente si escuchó o no los estímulos en la misma posición. Este proceso realizará hasta que el usuario responda que dichos estímulos provienen de diferentes posiciones.



**Figura 1.1:** Diagrama de bloques.

El toolbox tendrá las siguientes características:

- Se implementará una interfaz gráfica para interactuar con el usuario durante el experimento y configurar los parámetros del mismo.
- Se utilizará bases de datos de HRTFs en formato SOFA.
- Se permitirá ecualizar la HRTF a través de una función externa. Cabe mencionar que el diseño del ecualizador no es parte del alcance de este proyecto. Sin embargo, el usuario podrá ingresar su ecualizador de acuerdo a la función prototipo especificada en el diseño del programa. De esta manera, el usuario no requerirá conocer detalles de implementación del *toolbox*.
- Se obtendrá estadísticas y gráficas de dispersión así como histogramas de los resultados de los experimentos.
- Finalmente, los componentes del toolbox serán diseñados de manera modular para permitir la reutilización de código así como la implementación de futuras características (e.g. nuevos experimentos psicoacústicos) sin que se requiera conocer detalles de programación internos del mismo.

## 1.4. MARCO TEÓRICO

### 1.4.1. Audio Espacial

Desde los últimos 50 años el método más simple para proveer sonido espacial es el sonido estéreo. Con el pasar de los años aparecieron los sistemas de sonido envolvente (*surround*) que intentan mejorar la sensación sonido espacial al agregar más canales que el estéreo como ((Dolby Surround, Dolby Digital, DTS, SDDS y otros 5.1, 6.1 y 7.1) pero estos sistemas no proporcionan la sensación de sonido 3D. Además el campo de acción de estos sistemas solo es útil mientras se encuentra dentro del punto central del círculo que forman los altavoces degradándose la percepción fuera del centro. [2]

Los avances en el procesamiento de señales acústicas hizo posible la creación de sonido envolvente más realista que consiste en reproducir el sonido directamente en los oídos y logra crear experiencias de escucha de sonido envolvente, así como intentos de crear experiencias auditivas binaurales de la vida real utilizando altavoces estéreo y auriculares tradicionales. Esta experiencia acústica logra simular que la persona estuviese en el espacio acústico. La fidelidad de este sonido depende de la sensación que obtenga dicho oyente; comúnmente se lo conoce como señal binaural<sup>1</sup>. [3]

El término audio espacial se refiere al conjunto de técnicas que modelan la influencia de la dirección dependiente de los oídos, cabeza y torso sobre la incidencia del campo sonoro usando filtros digitales para generar sonido virtual. Si se filtra un audio a través de dicho filtro, se obtiene sonido virtual similar al de una ubicación espacial específica.

El objetivo de los sistemas de audio 3D es manipular la percepción espacial de audio del oyente teniendo en cuenta las especificaciones de ingeniería, así como las consideraciones psicoacústicas [4].

### 1.4.2. Sistema de Coordenadas

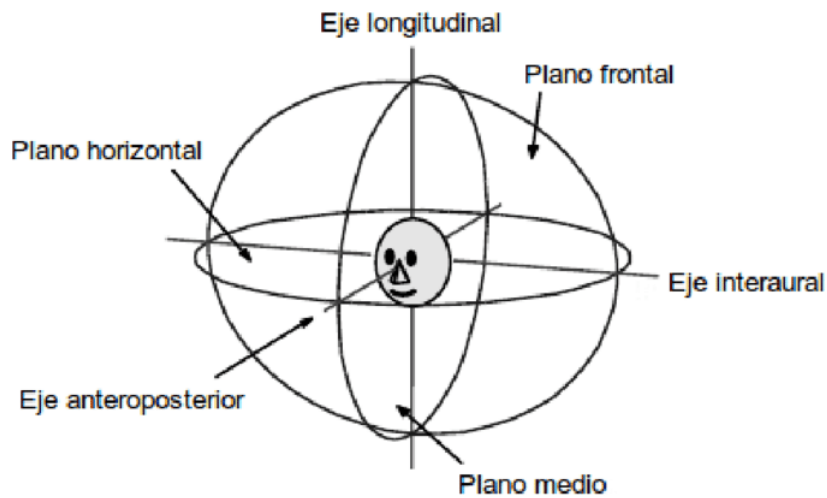
En audio espacial se debe especificar la posición de la fuente de sonido con respecto al centro de la cabeza del oyente. En la Figura 1.2 se muestra los tres importantes planos de referencia para localizar la fuente:

- Plano Mediano: plano y-z

- Plano Frontal: plano x-z

<sup>1</sup>El sonido binaural grabado con dos micrófonos en una cabeza artificial intenta crear una sensación de sonido inmersivo similar a la de estar físicamente en el lugar.

- Plano Horizontal: plano x-y



**Figura 1.2:** Planos de referencia [4]

El eje interaural es la línea que conecta los dos oídos y se puede observar en la Figura 1.2. Hay dos sistemas de coordenadas ampliamente utilizados en audio espacial: el sistema de coordenadas esféricas y el sistema de coordenadas polar interaural. En el sistema de coordenadas esféricas la posición de la fuente de sonido está dado por  $(r; \theta; \phi)$ . La distancia  $0 \leq r \leq \infty$  está medido con respecto al origen. El ángulo  $-90^\circ \leq \theta \leq +90^\circ$  corresponde la elevación de la fuente sonora encima y por debajo del oyente siendo  $0^\circ$  la posición en el plano horizontal. El ángulo  $0^\circ \leq \phi \leq 360^\circ$  representa el azimut donde  $0^\circ$ ,  $90^\circ$ ,  $180^\circ$  y  $270^\circ$  corresponde a las posiciones frente, derecha, atrás e izquierda respectivamente.

El sistema de coordenadas polar interaural está definida por  $(r, \Theta, \phi)$  donde la distancia con respecto al origen es  $0 \leq r \leq \infty$ . Los ángulos de azimut y elevación están comprendidos entre  $-90^\circ \leq \Theta \leq +90^\circ$  y  $-90^\circ \leq \phi \leq 270^\circ$  respectivamente. Así en este sistema, las coordenadas

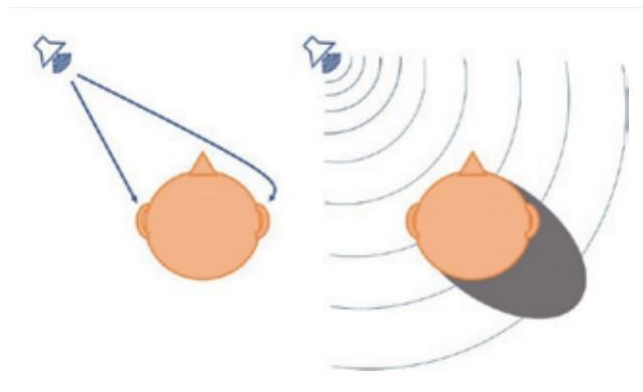
$$(\Theta, \phi) = (0^\circ, 0^\circ), (0^\circ, 90^\circ), (0^\circ, 270^\circ), (90^\circ, 0^\circ), (-90^\circ, 0^\circ)$$

corresponden a las posiciones frente, arriba, atrás, debajo, derecha e izquierda respectivamente. [4]

### 1.4.3. Percepción auditiva

Para explicar el audio espacial se debe conocer las técnicas que utiliza el oído humano para identificar la localización de la fuente sonora en un determinado espacio.

Las señales binaurales son las más importantes para localizar una fuente de sonido en el plano horizontal. Implican las diferencias relativas entre las ondas que llegan a ambos oídos de la siguiente manera: la cabeza funciona como barrera acústica de tal modo que la señal auditiva que proviene de una fuente que se encuentra fuera del plano lateral llegue a cada oído con distintas intensidades y espaciadas por un lapso temporal como se muestra en la Figura 1.3. Así, el sonido producido por una fuente alcanzará al oído más cercano a la fuente más pronto que al oído más alejado, esta es la Diferencia Interaural de Tiempo (*ITD*).



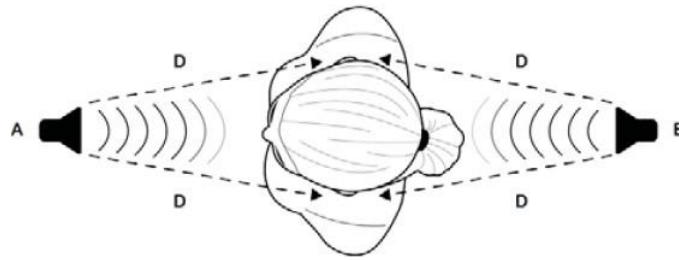
**Figura 1.3:** Representación de la trayectoria que el sonido recorre para llegar a los oídos permitiendo identificar la diferencia interaural [5]

Además, la cabeza funciona como un panel acústico absorbente que hace que el sonido llegue con distintas intensidades a cada oído, si la fuente sonora se encuentra fuera del plano lateral, lo que se conoce como Diferencia de Intensidad Interaural (*ILD*) [6].

Estas diferencias se conocen como diferencias interaurales y se definen como la diferencia en el nivel sonoro de una misma señal recibida en ambos oídos y permiten al cerebro procesar las señales para localizar la fuente sonora. [5] De esta manera la persona puede saber de dónde proviene la fuente sonora en su entorno, mediante las diferencias que se producen en las ondas acústicas que llegan a cada oído, siempre y cuando dicha fuente se encuentre fuera del plano lateral.

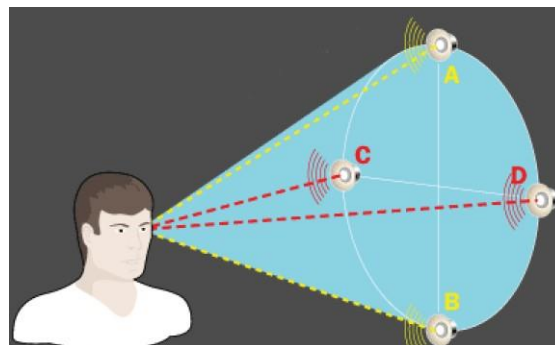
Sin embargo, cuando la fuente se encuentra en el plano lateral ya sea adelante o atrás del oyente, la localización de la fuente es más difícil ya que las diferencias pueden ser nulas. En la Figura 1.4 se puede observar la ejemplificación de la dificultad para distinguir la localización de la fuente sonora ya que se encuentra a la misma distancia de los oídos lo que da diferencias idénticas de intensidad y tiempo. [6]

Para la localización del sonido también comprenden los factores de elevación y distancia debido a que existe una zona en que la ITD e ILD no permiten distinguir una sola posición, esta zona se la conoce como *cono de confusión*.



**Figura 1.4:** Confusión frente-atrás [6]

En la Figura 1.5 se puede observar la zona que forma dicho cono de confusión.



**Figura 1.5:** Cono de confusión donde (A y B) y (C y D) tienen las mismas ITD e ILD. [6]

La elevación es otro factor para poder determinar la posición de la fuente sonora, sin embargo, ésta es netamente monoaural, es decir, se puede percibir la elevación con un oído debido a las características de resonancia del mismo.

#### 1.4.4. HRTF

Las HRTF vienen de las siglas Head Related Transfer Function (*Funciones de Transferencia Relacionada a la Cabeza*) y describe cómo el oído recibe el sonido que llega de una fuente sonora. El sonido viaja a través del espacio en todas las direcciones describiendo una esfera de rápida expansión. Estos sonidos van chocando con los objetos y muros presentes por lo tanto, el sonido se va refractando y reflejando y llega al oído humano desde varias direcciones. Cuando las ondas sonoras llegan al oyente, son transformadas por el cuerpo del oyente, las orejas, la cabeza, los hombros e incluso el torso contribuyen a la HRTF. Especialmente, el tamaño y masa de la cabeza, la forma del oído, la longitud y diámetro del canal auditivo, las dimensiones de cavidades orales y sinusales manipulan las ondas sonoras entrantes aumentando algunas frecuencias y disminuyendo otras. Estos cambios es lo que crean una percepción única para el oyente y ayuda a identificar la ubicación de la



fuente sonora [7].

Las HRTFs son componentes clave para analizar y sintetizar la señal de audio espacial renderizado binauralmente, porque al filtrar un sonido a través de una par de HRTFs, es decir, uno para cada oído es posible colocar un sonido en posiciones específicas en el espacio [4].

### Definición

Las HRTFs pueden definir como: el sonido emitido por una fuente de sonido en el campo libre llega a los oídos después de interactuar con las características anatómicas de una persona (es decir, cabeza, torso y pabellón auricular). El resultado contiene varias señales como el ILD e ITD las cuales se modelan a través de las HRTFs. El par de HRTFs para el oído izquierdo y derecho,  $H_L$  y  $H_R$  respectivamente está definido por:

$$\begin{aligned} H_L(r, \Theta, \Phi, f, a) &= \frac{P_L(r, \vartheta, \Phi, f, a)}{P_0(r, f)} \\ H_R(r, \Theta, \Phi, f, a) &= \frac{P_R(r, \vartheta, \Phi, f, a)}{P_0(r, f)} \end{aligned} \quad (1.1)$$

Donde,  $P_L$  y  $P_R$  representan las presiones sonoras en el dominio de frecuencia en el oído izquierdo y derecho respectivamente, y  $P_0$  representa el la presión sonora en el campo libre en el dominio de la frecuencia en el centro de la cabeza con ésta ausente. Se debe tomar en cuenta que  $P_0$  debe ser generado con la misma potencia que en las mediciones de  $P_L$  y  $P_R$ . La variable  $a$  depende de las características anatómicas de cada persona y normalmente se representa por un conjunto de medidas antropométricas de la cabeza, el torso y el pabellón auricular. Si la distancia es aproximadamente  $r > 1$  m, las HRTFs son independientes de la distancia y son conocidas como HRTFs de campo lejano. En cambio, si las HRTFs dependen de la distancia se denominan de campo cercano [4].

La ventaja de este método es que preserva el modelo complejo de las diferencias interaurales sobre el espectro completo de los estímulos, así capturando los efectos de filtrado por el pabellón auricular, cabeza, hombros y torso [8].

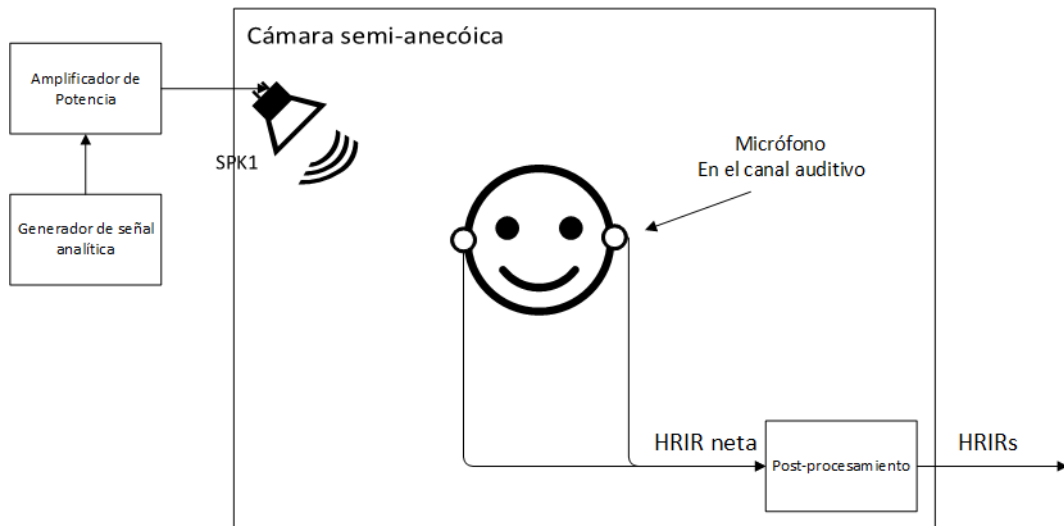
### Mediciones de HRTF

La HRTF de un individuo es obtenida reproduciendo una señal sintética conocida (señal analítica) en la dirección deseada con una distancia de al menos de 1m de la fuente (HRTF de campo lejano) y luego midiendo el HRIR<sup>2</sup> total usando micrófonos ubicados cerca del

<sup>2</sup>Las HRIR (Head-Related Impulse Response, Respuesta impulsiva relacionada con la cabeza) son la contraparte en tiempo de las HRTFs [4]

canal auditivo.

Teóricamente las HRTFs son funciones continuas como se muestra en la Ecuación 1.1 pero en la práctica, las HRTFs se miden para posiciones discretas en el espacio y para posiciones no medidas podrían obtenerse mediante *interpolación espacial* [9] [10] [11]. La Figura 1.6



**Figura 1.6:** Diagrama de bloques para el procedimiento de medición de HRIRs

muestra el diagrama de bloques para la mediciones de las HRTF las cuales se detallan a continuación:

- **Generador de señal analítica:** debido al ruido presente es necesario que las características de la señal de entrada ayuden a mejorar la relación señal a ruido (SNR) aumentando la potencia de la señal evitando la distorsión de la misma. En el caso de las HRIR se han empleado varios tipos de señales deterministas (por ejemplo: impulsos, sinusoidales, señales aleatorias: ruido blanco, ruido rosa) y señales pseudoaleatorias aunque en la actualidad los barridos sinusoidales son las más usados ya que generan HRIRs con un SNR mayor [4].
- **Posición del micrófono:** el cual se coloca en un punto cercano a la entrada del canal bloqueando o desbloqueando el mismo o incluso cerca del tímpano. [4].
- **Individuos:** las HRIRs son funciones que dependen de las características anatómicas que varían de persona a persona por lo que es complicado tener un conjunto de HRIRs generales por lo tanto, deben ser medidas para varios sujetos. Debido a que es difícil mantener a una persona quieta por largos períodos de tiempo ya que éstas tienden a mover la cabeza y el cuerpo, se usan comúnmente maniqués KEMAR (Knowles Electronics Manikin For Acoustics Research) [12].

- **Cámara (semi)anecóica:** comúnmente usada para las mediciones de las HRIRs para eliminar los reflejos del entorno, sin embargo debido al alto costo de estas cámaras se usan las cámaras semianecóicas las cuales están equipadas con material absorbente acústico.
- **Post-procesamiento de HRIRs netas:** las HRIRs netas deben ser ecualizadas para compensar las distorsiones del espectro causadas por el sistema electroacústico (por ejemplo: micrófonos, parlantes, amplificadores) y minimizar los cambios subjetivos en el timbre de las señales de entrada [4].

### Bases de datos de HRTFs

Usando el procedimiento descrito en la sección 1.4.4 o similares, los equipos de investigación han creado bases de datos de HRTFs disponibles para el público. Por ejemplo, una de las más completas es la base de datos CIPIC [13] porque tiene el número de individuos más grande. Se realizaron 1250 mediciones en coordenadas interaurales para 25 azimut uniformemente no distribuidas con una resolución máxima de  $5^\circ$  para ubicaciones cercanas al plano mediano y 50 elevaciones con una resolución  $5.625^\circ$  en el rango  $-45^\circ$  a  $230.625^\circ$ . La desventaja de esta base de datos es su baja resolución en la dirección lateral, es decir, de  $15^\circ$  a  $20^\circ$ . Otra popular es la base de datos ARI cuya principal ventaja sobre la CIPIC que ofrece HRTFs de más sujetos (alrededor de 150) y tiene mejor resolución espacial [14]. También hay bases de datos que solo midieron las HRTF para KEMAR [12] o maniqués similares. Para facilitar el intercambio de datos entre usuarios e investigadores la Asociación de Ingeniería de Audio (*AES, Audio Engineering Society*) ha estandarizado el formato SOFA (Spatially Oriented Format for Acoustics) [15], el cual es un formato de archivo para almacenar e intercambiar datos acústicos orientados espacialmente como las HRTFs. Actualmente las bases de datos más populares están disponibles públicamente para su descarga en formato SOFA incluidas las bases de datos CIPIC y ARI [4].

#### 1.4.5. Características en el dominio del tiempo de las HRIRs

La señal de localización más importante en el tiempo es la ITD. Aunque hay varios métodos para estimar la ITD la manera más directa es calcular la diferencia entre los retrasos de inicio entre la HRIR izquierda y derecha, así, la ITD se define como la diferencia entre inicio de retraso  $t_L$  y  $t_R$ :

$$ITD(\Theta, \Phi) = t_L - t_R \quad (1.2)$$

Se debe tomar en cuenta que para determinar un retraso es necesario elegir un umbral. Los HRIRs se muestrean a menudo antes del cálculo del ITD para aumentar la resolución en tiempo.

#### **1.4.6. Características en el dominio de la frecuencia de las HRTFs**

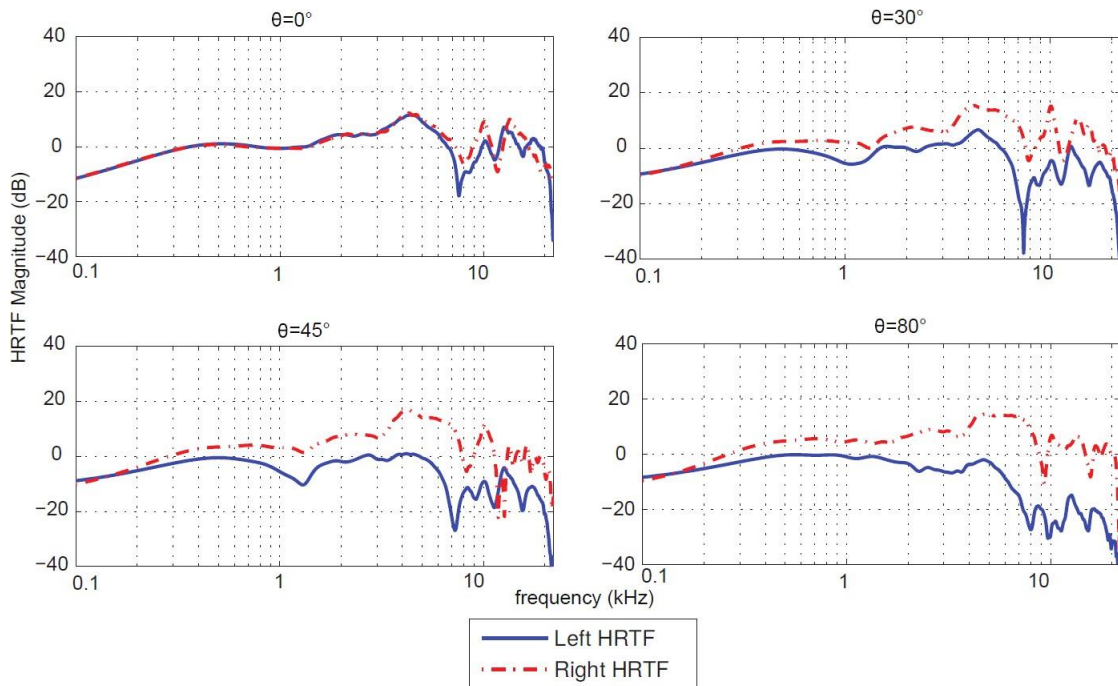
Para analizar las características en frecuencia de las HRTFs, en la Figura 1.9 se muestra las magnitudes de las HRTF para cuatro azimuts en el plano horizontal donde [4]:

1. Por debajo de los 200Hz existe una caída de nivel de señal debido a la respuesta en frecuencia de los altavoces utilizados, ya que estos suelen estar limitados a bajas frecuencias.
2. A frecuencias bajo los 0.4-0.5 KHz, el efecto de atenuación en la cabeza es insignificante de modo que las magnitudes de las HRTFs para ambos oídos son alrededor de 0dB y son independientes de la frecuencia.
3. A medida que la frecuencia aumenta por encima de los 1.5KHz las diferencias de nivel interaural se vuelven más evidentes.
4. Aunque para  $\Theta=0^\circ$  las HRTF izquierda y derecha son similares, las pequeñas diferencias entre ellas vienen del hecho que la cabeza no es perfectamente simétrica.
5. A altas frecuencias, en la banda de 5-6KHz las magnitudes de las HRTFs varían en una manera compleja presentando varios picos y bajos.
6. El pico cerca de los 4KHz se debe a la resonancia del canal auditivo.

#### **1.4.7. Características en el dominio de la frecuencia debido al pabellón auricular**

A frecuencia sobre los 3KHz, cuando el tamaño del pabellón auricular es comparable a la longitud de onda de la fuente, las características del pabellón auricular se vuelven indispensables para resolver la parte frente-atrás y las confusiones arriba-abajo así como la precisión de localización del sonido.

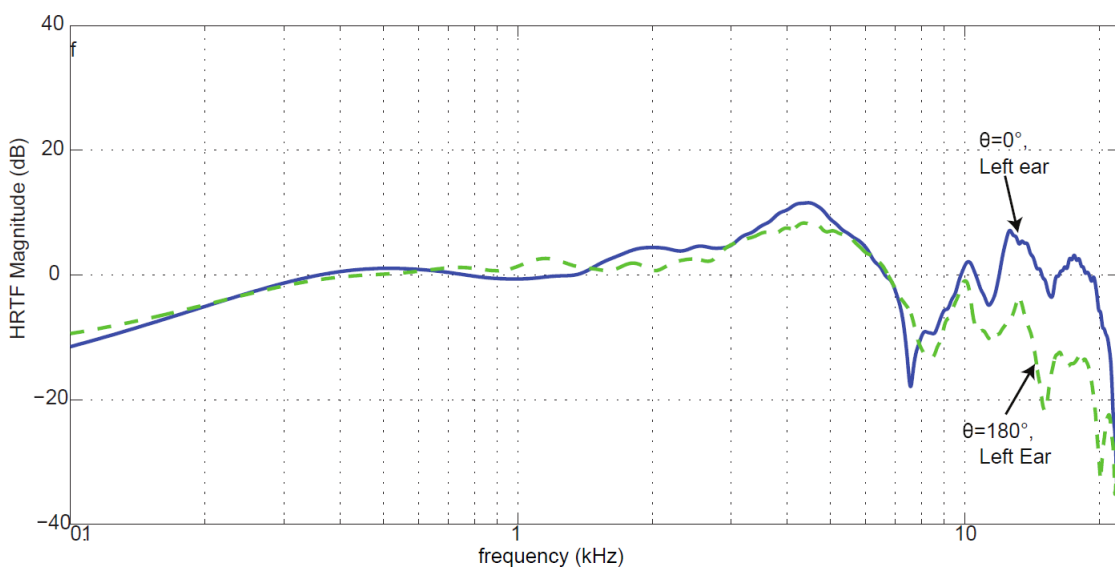
Debido a que la ITD e ILD similares causan ambigüedades con la confusión frente-atrás, las cuales se resuelven a las características de alta frecuencia debido al pabellón auricular. Para comprender cómo el pabellón auricular provee información espectral para resolver estas ambigüedades, en la Figura se muestra las magnitudes de la HRTF en el plano horizontal



**Figura 1.7:** Magnitudes de HRTFs de KEMAR de la base de datos CIPIC para varios azimuts en el plano horizontal [4]

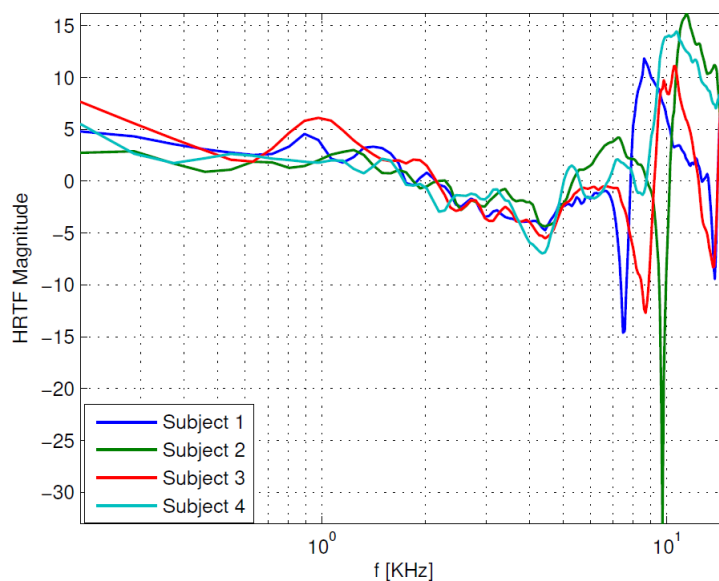
para el mismo oído a  $\Theta=0^\circ$  (es decir al frente) y  $\Theta=180^\circ$  (es decir atrás) [4].

Se debe tomar en cuenta que a altas frecuencias, las diferencias de magnitud desde ambas direcciones son evidentes. Estas diferencias causadas por la asimetría de la cabeza, la posición del oído y la forma del pabellón auricular permiten resolver las confusiones frente-atrás [4].



**Figura 1.8:** Magnitudes de HRTF para azimuts =  $0^\circ$  (es decir, frente) y  $180^\circ$  (es decir, atrás) en el plano horizontal [4]

Adicionalmente, las magnitudes de las HRTF se caracterizan por presentar picos y hondonadas a frecuencias superiores a aproximadamente 5KHz. La frecuencia central de la primera de estas hondonadas es considerada una señal importante para en la ubicación vertical. Otro factor importante respecto a la frecuencia de la primera hondonada es la variabilidad entre diferentes personas debido a las características antropométricas del pabellón auricular [13]. La Figura muestra el HRTF izquierdo en una ubicación específica para cuatro sujetos diferentes en los cuales la variabilidad entre ellos es notable a partir de los 4KHz donde las señales monoaurales introducidas por el pabellón auricular son más prominentes. Por esta razón, el pabellón auricular se considera como la huella dactilar acústica de una persona. Varios estudios muestran una disminución en la precisión de la localización debido a las HRTFs no individualizadas, que a menudo producen confusiones frente-atrás, mala externalización del sonido y percepción incorrecta de la elevación. Por esta razón es necesario personalizar las HRTFs para garantizar la percepción del sonido 3D de alta calidad [4].



**Figura 1.9:** HRTF izquierda en una localización específica para cuatro personas diferentes) en el plano horizontal [4]

#### 1.4.8. Formatos para bases de datos de HRTFs

Las HRTFs han sido medidas por cada laboratorio y almacenadas en el formato preferido de cada laboratorio. Sin embargo, puede resultar complicado al momento de compartir con otros investigadores y usuarios debido a la incompatibilidad entre formatos. Por esta razón se han propuesto formatos que se han estandarizado para almacenar HRTFs los cuales, algunas se mencionan a continuación: [6]

## **OpenDaff**

OpenDaff (Open Directional Audio File Format) es un formato libre y de código abierto para contenido de audio direccional como directividades de fuentes de sonido (por ejemplo: altavoces, instrumentos musicales) y receptores de sonido (por ejemplo: micrófonos, HRI-R/HRTF). OpenDaff permite el intercambio y representación eficiente de datos en un solo archivo de formato ".daff" [16].

## **Listen**

Consiste en una matriz HRTF y matrices adicionales que describen la dirección de la HRTF correspondiente lo que permite representar la HRTF desde cualquier dirección. En este formato las HRTFs de cada oyente se almacenan en un archivo separado [17].

## **SOFA**

El presente proyecto de titulación realiza el toolbox en Matlab utilizando el formato SOFA por lo que se describirá una forma detallada.

SOFA viene como propuesta para almacenar datos acústicos enfocado a la intercambiabilidad y extensibilidad. Es un formato de archivos para almacenar o representar datos acústicos orientados espacialmente y permite almacenar además de HRTFs y HRIRs, datos más complejos como respuestas direccionales de impulso de sala (DRIR) medidas con una matriz de micrófonos multicanal excitada por una matriz de altavoces. SOFA ha sido estandarizado por la Audio Engineering Society (AES) como AES69-2015 [17].

SOFA tiene como objetivo cumplir los siguientes requisitos:

- Describe una configuración de medición de geometría arbitraria, es decir, no se limita solo casos como cuadrícula regular o una distancia constante.
- Toda la información requerida sobre la configuración de la medición debe proporcionarse como metadatos en el archivo.
- Flexibilidad para escribir datos de múltiples condiciones (oyentes, distancias, etc.) en un solo archivo.
- Soporte parcial de archivos y redes.
- Disponible como archivo binario con compresión de datos para un eficiente almacenamiento y transferencia.

- Convenciones de descripción predefinidas para las configuraciones de medición más comunes.

La configuración de la medición se describe por varios objetos y sus relaciones. Se considera una medición como una observación discreta muestreada en un momento específico y bajo una condición específica. Cada medición consta de datos y se describe por sus correspondientes dimensiones y metadatos. Todas las mediciones se almacenan en una sola estructura de datos. Las convenciones SOFA proporcionan una descripción coherente de las configuraciones de medición.

El formato SOFA soporta dos sistemas de coordenadas: cartesianas y esféricas en la representación azimut-elevación [17].

### **Archivos SOFA**

El repositorio principal de SOFA tiene como objetivo recopilar las HRTF, BRIR, DRIR y otros datos relacionados con SOFA disponibles para todos, los cuales pueden acceder y descargar. Al momento no está disponible la descarga parcial, acceso a metadatos y la búsqueda en la base de datos, sin embargo, se está investigando la posibilidad de usar OPeNDAP<sup>3</sup> para los repositorios SOFA [17].

### **Bases de datos de campo libre de propósito general**

Las bases de datos incluyen:

HRTFs estándar (en el canal auditivo) de humanos [17]:

- ARI: HRTF de la base de datos ARI. HRTF y DTF en el oído para más de 220 oyentes.
- ARI (ALTB): HRTFs de la base de datos ARI. Son mediciones para algunos de los oyentes de la base de datos ARI las cuales fueron repetidas y evaluadas unos años más tarde.
- CIPIC: HRTF de la base de datos CIPIC con 45 oyentes y datos antropométricos parcialmente disponibles.
- RIEC: HRTFs de campo lejano de la base de datos RIEC del Advanced Acoustic Information Systems Laboratory con más de 100 oyentes humanos.

---

<sup>3</sup>OPeNDAP es un framework que simplifica la creación de redes de datos científicos que permite que dichos datos sean accesibles desde cualquier ubicación remota y se encuentra disponible gratuitamente [18].



- Aachen: HRTFs de la base de datos de Aachen combinados con datos antropométricos y un modelo 3D de oídos de 48 oyentes.
- Aachen: HRTF de alta resolución de la base de datos Aachen combinados con un modelo 3D de un oyente humano.
- HUTBUS: HRTFs de la base de datos HUTUBS que contiene datos antropométricos, respuestas de impulso de auriculares y modelos de cabeza 3D de 96 oyentes.
- CHEDAR: HRTF calculadas numéricamente (.sofa) con mallas 3D de la cabeza y pabellón auricular (.ply) y datos antropométricos (.mat) proporcionados.
- 3D3A: HRTF medidas en el oído (.sofa) de 38 personas y escaneos 3D de cabeza y torso de 31 personas de la base de datos 3D3A de Princeton.

A continuación se presentan HRTFs de cabezas artificiales [17]:

- MIT-KEMAR: HRTF de la cabeza artificial de MIT-KEMAR utilizada en muchas publicaciones.
- ARI (ARTIFICIAL): HRTFs de maniqués (cabezas ficticias) medidas en ARI utilizando la misma configuración para oyentes humanos.
- THK/HRIR: HRTFs de varios maniqués provistos por el Technische Hochschule Köln.
- SCUT: HRTF de campo cercano de la base de datos SCUT de KEMAR.
- TU-Berlin: HRTF de TU-Berlin de la cabeza ficticia de KEMAR y también de FABIAN para varias distancias mayores a 0.5 metros.
- Club Fritz: HRTFs de Neumann KU, 100 mediciones como parte del proyecto Club Fritz donde varios institutos midieron exactamente la misma cabeza artificial.
- VIKING: HRTF de esfera completa (.sofa) de la base de datos Viking con el maniquí KEMAR con 20 pares diferentes de pabellones auriculares artificiales de silicona. También se incluyen escaneos 3D de pabellones auriculares.
- Aachen: HRTF de alta resolución de la cabeza artificial ficticia KEMAR de la base de datos de Aquisgran, combinados con un modelo 3D de un individuo.
- PKU-IOA: Base de datos de HRTFs de alta resolución espacial en la oreja de la cabeza ficticia de KEMAR a distancias de 20, 30, 40, 50, 75 100 130 160 cm.

HRTFs especiales [17]:

- ARI (BTE): HRTF y DTF detrás de la oreja de la base de datos ARI.

Bases de datos de respuestas de impulso de sala (DRIR, SRIR, SRIR) [17]:

- Oldenburg: DRIRs de Oldenburg grabadas en una oficina bajo varias condiciones.
- THK/DRIR y THK/BRIR: DRIRs y BRIRs medidos en los estudios de difusión WDR con varias matrices de micrófonos proporcionados por la Technische Hochschule Köln (TH Köln).

#### **1.4.9. Tipos de Experimentos Psicoacústicos**

##### **Localización absoluta**

La localización absoluta presenta un reto mayor ya que no se cuenta con la información de la posición de la fuente sonora, por lo tanto se debe conocer la ubicación de dicha fuente [19].

##### **Localización relativa**

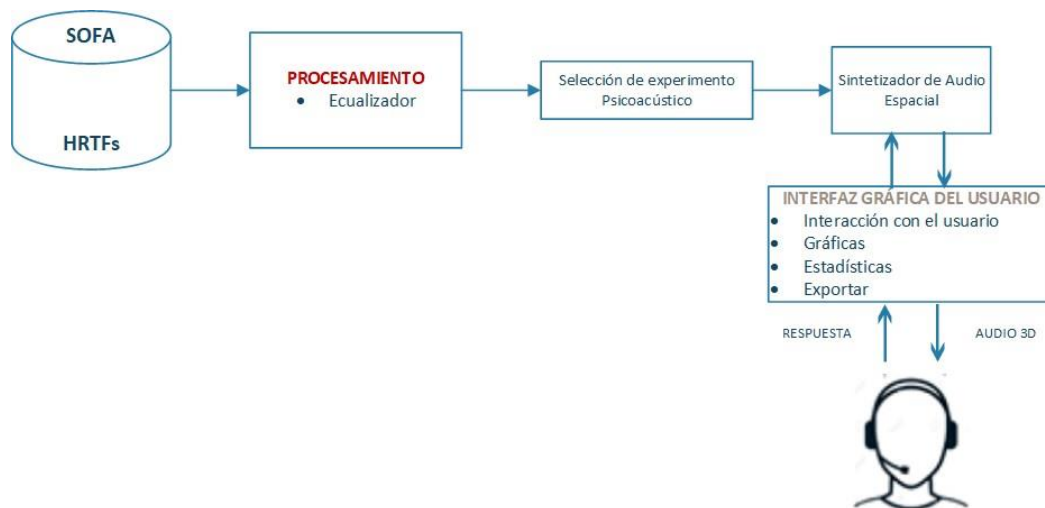
Para la localización relativa, se conoce la posición inicial de la fuente sonora, y luego se reproduce un segundo estímulo para determinar la localización de éste respecto a la posición inicial de la fuente sonora [19].

El presente trabajo de integración realizará el *toolbox* en Matlab para realizar los experimentos psicoacústicos de localización absoluta y relativa, en total son cuatro y se encuentran detallados en el alcance del presente documento en la sección 1.3.

## 2. METODOLOGÍA

En esta sección se detalla el funcionamiento del *toolbox*, para ello, la Figura 2.1 indica el proceso que se sigue para su diseño y se puede evidenciar las diferentes fases que consisten en lectura de las HRTFs a través de la carga de archivos SOFA y creación de bases de datos, procesamiento en el que interviene la ecualización, selección del experimento psicoacústico que lo conforman cuatro experimentos: dos de localización absoluta en donde el usuario no conoce la localización del sonido y dos localización relativa en donde el usuario conoce la ubicación del primer sonido pero no del segundo sonido, sintetizador de audio espacial que junto con el archivo SOFA logra reproducir un sonido en una ubicación específica y la obtención de estadísticas que contiene la información de todos los experimentos realizados.

Una vez mostrado el esquema se procederá a detallar cada fase y cómo se realizó en base al esquema de la Figura 2.1.



**Figura 2.1:** Diagrama de bloques

Es importante mencionar que el código se encuentra disponible como anexo digital en el presente trabajo de titulación.

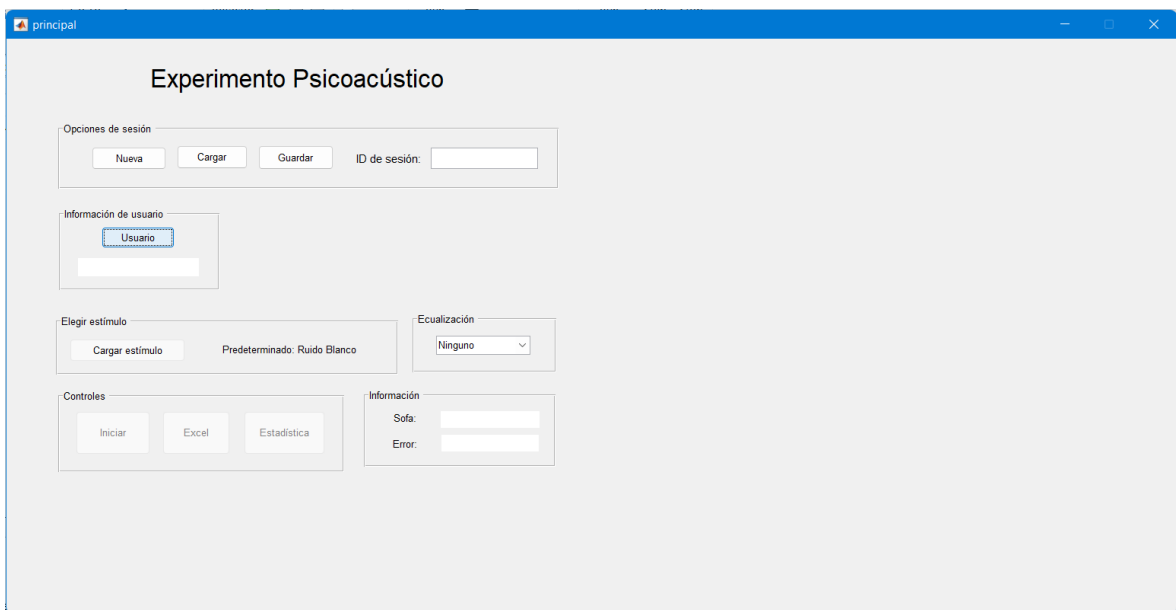
### 2.1. Lectura de HRTFs

En esta fase se realiza la creación de base de datos, usuarios, lecturas de las HRTFs por medio de los archivos SOFA previamente descargados de los repositorios disponibles y demás información necesaria para la correcta ejecución del programa. Para empezar con la creación de bases de datos y usuarios, el *toolbox* que cuenta con una interfaz gráfica la

cual está conformada por grupos de botones que se muestra en la Figura 2.2 los cuales cumplen una determinada función. Estos grupos son para las opciones de inicio de sesión, información de usuario y elegir estímulo. Este último es opcional y por defecto ya tiene precargado un audio en formato WAV pero el usuario puede optar por elegir su propio archivo de audio que también debe estar en formato WAV.

### 2.1.1. Interfaz inicial

La interfaz gráfica que se muestra al ejecutar el programa comprende la ventana que se muestra en la Figura 2.2.



**Figura 2.2:** Interfaz inicial del toolbox

Al ejecutar la función "principal.m" se despliega la interfaz de la Figura 2.2. Esta función, como su nombre lo indica es la principal función del *toolbox* y es la que contiene todas las funciones (búsqueda de archivos y carpetas, guardar, cargar, reproducción de audio, estadísticas), además invoca a programas externos (Microsoft Excel) para mostrar datos y estadísticas. En la mencionada interfaz se encuentran los grupos de botones para la ejecución del *toolbox* los cuales se describen a continuación:

#### **Opciones de inicio de sesión:**

Este grupo contiene los botones para crear, cargar o guardar una base de datos. A continuación se describen los botones que presenta el grupo:

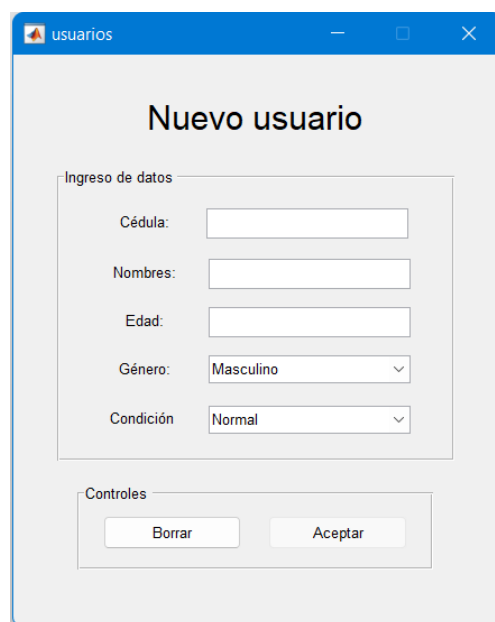
- Nueva: Para borrar todos los datos ingresados y crear una nueva base de datos.

- Cargar: Para cargar una base de datos en el caso que ya se haya guardado una anteriormente.
- Guardar: para guardar una base datos con todas las configuraciones y experimentos realizados. Para poder guardar se debe colocar un nombre en el campo ID de sesión.

El código de estos tres botones se encuentra en la función principal y se los puede localizar con el mismo nombre que se muestran en la interfaz.

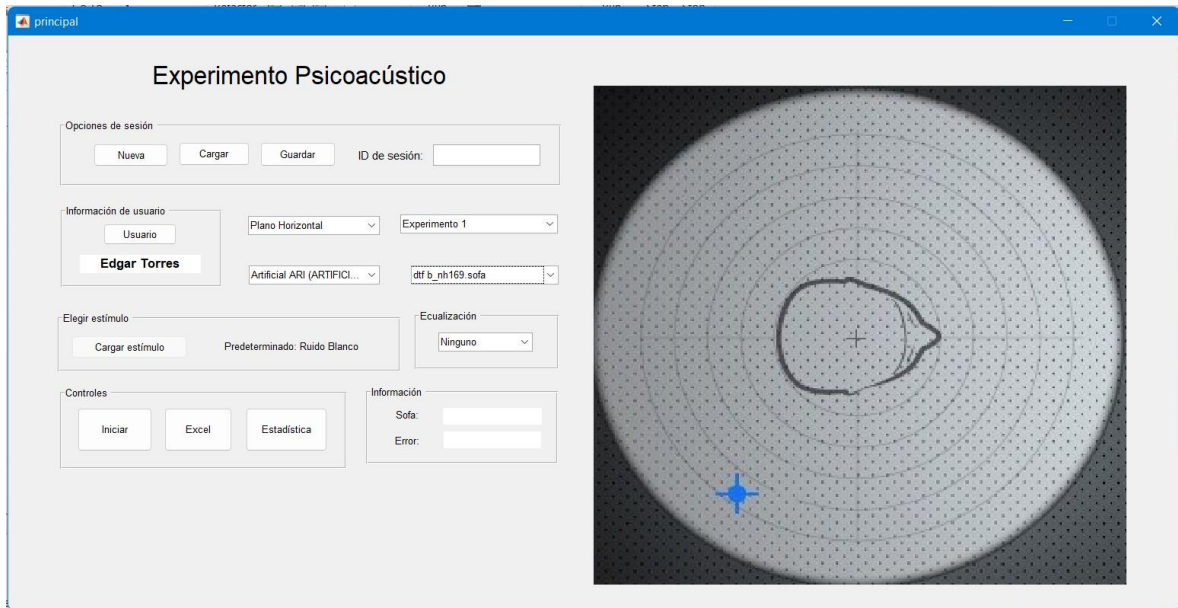
### Información de usuario

Este grupo tiene la función de agregar usuarios a la base de datos. Al presionar el botón "Usuarios" se despliega una interfaz mostrada en la Figura 2.3 donde se deberá ingresar los datos del usuario que va a realizar los experimentos.



**Figura 2.3:** Interfaz para agregar usuarios

En la interfaz de la Figura 2.3 se deberá completar los campos solicitados, al presionar en Aceptar se regresará a la interfaz inicial y el nombre de usuario se mostrará en el cuadro de texto dentro del grupo Información. Una vez ingresados los datos aparecerán las listas desplegables que permiten escoger, el plano, el experimento, el archivo SOFA además se mostrará el gráfico correspondiente del plano escogido, como se muestra en la Figura 2.4.



**Figura 2.4:** Interfaz para escoger el plano, tipo de experimento y archivo SOFA

### Elegir estímulo

En el programa, por defecto se encuentra cargado un archivo de sonido denominado "Ruido Blanco", pero el usuario puede optar por cargar su propio archivo de sonido audio que debe estar en formato WAV presionando el botón "Cargar estímulo".

## 2.2. Procesamiento

En esta fase, el *toolbox* realiza un procesamiento del archivo de sonido. He aquí donde la interfaz presenta una lista desplegable llamada Ecualización donde se pueden elegir los ecualizadores disponibles.

### Ecualización

Permite elegir un archivo de ecualización que se encuentra en la carpeta de "ecualizadores". Por defecto está configurado en ninguno por lo que el audio se reproducirá sin ninguna modificación. Al elegir otra opción de la lista desplegable podrá obtener el sonido modificado según la función que se haya elegido.

Esta función tiene como entradas el archivo de audio en formato WAV, no es necesario que éste sea estéreo o monofónico ya que el programa tomará un solo canal para el procesamiento. Como salida tendrá el archivo de audio ecualizado o no, dependiendo de la elección del usuario.

Cabe resaltar que el usuario puede añadir sus propias funciones ".m" de ecualización a la carpeta ecualizadores y estos archivos se mostrarán cuando se ejecute nuevamente el programa principal. Para que las funciones colocadas en dicha carpeta sean correctamente procesadas por el *toolbox*, éstas deben tener la forma:

$$y = EQ(x)$$

Donde

- *y* es el archivo de audio ecualizado que retorna de la función principal.
- *x* es el argumento de entrada, en este caso el archivo WAV.
- *EQ* es el nombre de la función. Por lo general esta toma el nombre que se le ha colocado en la carpeta de ecualizadores. Para el presente ejemplo sería "EQ.m".

De este modo con el formato indicado, el usuario podrá colocar las funciones que desee en formato ".m".

### **2.3. Selección del experimento psicoacústico**

En esta sección el usuario podrá elegir entre los cuatro experimentos que presenta el *toolbox* de los cuales dos son dedicados para localización relativa y dos para localización absoluta. A continuación se detalla cada experimento:

#### **2.3.1. Experimentos de localización absoluta**

Como se ha detallado en la sección 1.3, estos experimentos consisten en que el usuario no conoce la localización de la fuente sonora por lo que deberá seleccionar la ubicación de donde proviene el sonido.

##### **Experimento 1**

En este experimento el usuario debe reportar dónde escuchó el sonido. Para ello el procedimiento para ejecutar experimento en el programa es el siguiente:

1. Elegir el plano (Horizontal, Mediano o Frontal) en el que se desea realizar el experimento para que se habiliten las siguientes listas desplegadas.

2. Elegir el Experimento 1 en la siguiente lista desplegable para que se habilite las listas desplegables de los archivos SOFA.
3. Escoger entre las diferentes archivos SOFA que el programa le ofrecerá al usuario.
4. Cargar un archivo WAV para escuchar (opcional). Por defecto el programa tiene pre-cargado ruido blanco.
5. Elegir un ecualizador (opcional), por defecto está configurado en Ninguno.
6. Presionar el botón Iniciar para que empiece a reproducirse el sonido.
7. Al ejecutar el programa, éste le mostrará todas las posiciones en el plano seleccionado del que puede provenir el sonido para que el usuario elija la posible ubicación.
8. Una vez el usuario señale la ubicación, el audio se detendrá y se le mostrará con un color la posición señalada y con un color diferente la posición real.

## **Experimento 2**

En este experimento el usuario podrá seleccionar la ubicación de la fuente de audio y el programa le presentará tres sonidos en localizaciones cercanas a la posición elegida. El usuario deberá elegir el audio que más se acerque a la posición seleccionada. Para ello el procedimiento para ejecutar el segundo experimento en el programa es el siguiente:

1. Elegir el plano en el que se desea realizar el experimento. De igual manera las listas desplegables se habilitarán una vez elegido el plano.
2. Elegir el Experimento 2 y se habilitarán las listas para elegir el archivo SOFA.
3. Elegir el archivo SOFA deseado.
4. Igualmente se puede elegir un archivo WAV y ecualizador. Este paso es opcional.
5. Presionar el botón "Iniciar" y el programa dibujará todas las posiciones posibles en el plano donde el usuario debe elegir la ubicación de la fuente de sonido.
6. El usuario elegirá una posición y el programa pintará de diferente color las tres posiciones cercanas a la seleccionada.
7. El programa le reproducirá tres sonidos en orden aleatorio de las tres posiciones y luego se desplegará una cuadro de diálogo con la pregunta para elegir la posición que más se acercó a la seleccionada.



8. El usuario elegirá entre las tres opciones y el programa pintará de un color la posición elegida y de otro color la posición real.

### **2.3.2. Experimentos de localización relativa**

Como se ha detallado en la sección 1.3 los experimentos de localización relativa consisten en ubicar la fuente sonora respecto a una posición conocida.

Para estos experimentos, al igual que los Experimentos 1 y 2 de la sección 2.3.1, los pasos de elegir el plano, agregar archivos WAV y elegir un ecualizador se repiten en los dos siguientes experimentos por lo que dichos pasos serán omitidos de los procedimientos que se describen a continuación.

### **Experimento 3**

En este experimento el usuario deberá identificar la posición de un segundo audio respecto al primero audio cuya posición es conocida. El procedimiento es el siguiente:

1. Seleccionar el Experimento 3 de la lista desplegable para que se habilite las siguientes listas para elegir el archivo SOFA.
2. Presionar el botón "Iniciar" y el programa reproducirá un primer sonido y se dibujará la posición del primer sonido.
3. Luego reproducirá un segundo sonido y le preguntará al usuario dónde escuchó el segundo sonido con respecto al primero.
4. El usuario deberá elegir la posición dependiendo del plano: en el plano horizontal le mostrará las opciones: izquierda o derecha. Para los planos frontal y mediano se le mostrará las opciones: arriba o abajo como se muestra en la Figura 3.3.
5. Al elegir una posición, se mostrará la posición real de un color y la posición elegida en otro color. En el campo de información se mostrará el archivo SOFA escogido y el error porcentual, es decir, si acertó en la posición el error es 0 % y si no acertó el error es 100 %.

### **Experimento 4**

En este experimento se reproducirán dos sonidos consecutivos y el usuario deberá informar si los escuchó en el mismo lugar o no. En el caso de una respuesta afirmativa el proceso se

repetirá con los sonidos separados un ángulo cada vez mayor en cada repetición hasta que la respuesta sea negativa. El proceso se detalla a continuación:

1. Seleccionar el Experimento 4 de la lista desplegable y al activarse las siguientes opciones se elegirá los correspondientes archivos SOFA y presionar el botón "Iniciar" para ejecutar el programa.
2. El programa reproducirá dos sonidos separados ligeramente por un ángulo . El valor de este ángulo proviene del archivo SOFA y puede variar entre diferentes archivos. El programa pintará de un color todas las posibles posiciones, aunque para este experimento la posición inicial siempre será al frente del usuario para los planos horizontal y mediano. Para el plano frontal la posición inicial será el punto más arriba posible que el archivo SOFA permita ubicar. El segundo sonido se reproducirá en la siguiente posición disponible y se pintará de un color diferente.
3. El programa le preguntará si escuchó en la misma posición ambos sonidos. Si la respuesta es afirmativa se repite el proceso, donde el primer sonido sigue siendo al frente o arriba del usuario dependiendo del plano escogido y el segundo sonido se desplazará a la siguiente posición disponible.
4. Cuando el usuario responda negativamente a la pregunta que el programa le muestra, el proceso se detendrá y se indicará el ángulo de apertura en la sección de Información además del archivo SOFA utilizado. Las posiciones disponibles, la posición inicial y las posiciones desplazadas se encuentran pintan de diferentes colores.

Cabe aclarar que las posiciones desplazadas para los tres planos empiezan desde un punto fijo y se van desplazando como se muestra a continuación:

- Para el plano horizontal se dirigen desde el frente hacia la izquierda en sentido anti-horario.
- Para el plano mediano se dirigen desde al frente hacia arriba.
- Para el plano frontal desde arriba hacia la derecha y bajando en sentido horario.

#### **2.4. Sintetizador de audio espacial**

Esta sección comprende la ejecución del programa y es la parte principal ya que aquí se sintetiza el audio espacial, en conjunto con la ecualización y el plano elegido. Para la ejecución de esta parte de código se hace el llamado a una función denominada "iniciosofa.m"

para todos los experimentos y en cualquier plano. Esta función tiene como entradas, el archivo SOFA seleccionado, las coordenadas de un punto del plano elegido (dependiendo del experimento y plano este valor puede variar) y como salidas se tiene el valor escogido por el usuario, es decir, las coordenadas que señaló en el experimento y el valor real. El formato se detalla a continuación:

$$[valor\_real, cartesian\_coord] = iniciosofa(archivowav, fullpath, posicion\_h, handles)$$

Donde:

- *archivowav* es el archivo de audio en formato WAV. Es indiferente si el archivo es estéreo o monofónico ya que la función toma un solo canal para procesarlo, tal como se describió en el ecualizador. Esta variable es una matriz de datos numéricos.
- *fullpath* es el directorio completo donde se encuentra el archivo SOFA elegido. Es un vector de tipo String.
- *posicion\_h* es el punto de coordenadas del plano seleccionado. Como se mencionó anteriormente este valor puede variar dependiendo del tipo de experimento y plano. Esta variable es de tipo flotante.
- *handles* es una función en Matlab que permite pasar varios parámetros a otra función. Para este caso la función es `''iniciosofa.m''`.
- *valor\_real* y *cartesian\_coord* son valores que retornan a la función principal y son las coordenadas de las posiciones del valor real y valor elegido por el usuario.

Cabe mencionar que esta función no retorna el vector de audio espacial porque en ésta se realiza la reproducción del mismo. Esta función se encuentra programada en el grupo de botones denominado Controles que se describen a continuación.

## Controles

Este grupo permite ejecutar el programa y desplegar una interfaz para obtener estadísticas. Los botones que contiene se describe a continuación:

- Iniciar: Permite ejecutar el programa con el archivo SOFA, ecualizador, tipo de experimento escogido en el respectivo plano.

- Excel: permite ejecutar el programa de Microsoft Excel que contiene las tablas con la información de los experimentos realizados. Cabe aclarar que para que se ejecute el programa de Excel, la hoja de cálculo toma el nombre de ID de sesión por lo que este nombre debe estar colocado.
- Estadísticas: despliega una interfaz gráfica que permite obtener datos estadísticos de los experimentos realizados. Esta interfaz se la explicará en la sección 2.5.

Cabe aclarar que cada vez que el usuario realiza un experimento, es decir presiona el botón "Iniciar", todos los datos obtenidos se van almacenando en un vector. Al repetir el proceso este vector se va concatenando con el anterior formando una matriz con los datos cada que se presione el botón de inicio. Por lo tanto, la matriz resultante tendrá un tamaño proporcional al número de repeticiones o experimentos realizados.

## **2.5. Interfaz gráfica de usuarios**

Esta fase se encuentra presente en todas las anteriores, debido a que desde la elección del archivo SOFA, los planos, experimento, procesamiento y sintetizador del audio espacial se lo ha realizado mediante una interfaz gráfica que permite interactuar con el usuario. En segundo plano se han ejecutado archivos ".m" que han realizado el trabajo de las diferentes fases pero dicha ejecución es invisible hacia el usuario ya que solo observa la interfaz gráfica.

En esta fase se obtienen, las gráficas, estadísticas y la exportación de la información hacia otros formatos. A continuación se las detalla:

### **Información**

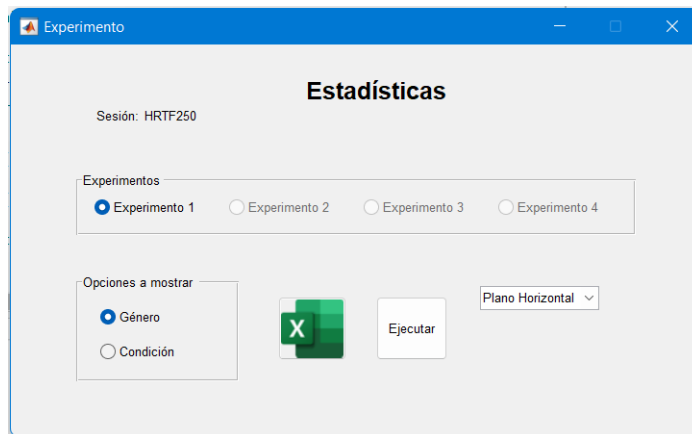
Son dos cuadros de texto que no son editables cuyo objetivo es mostrar la información del archivo SOFA utilizado y el error ya sea absoluto o relativo resultado del experimento realizado.

### **Tabla de resultados**

Aunque en la interfaz inicial no se muestra, ésta aparece cuando ya se ejecuta el programa y se obtienen los resultados de los experimentos. Esta tabla muestra los mismos datos que la hoja de cálculo cuando se presiona el botón "Excel" del grupo Controles.

## Interfaz Estadísticas

Esta interfaz se activa cuando se presiona el botón "Estadísticas" y permite obtener datos estadísticos siempre y cuando se hayan realizado un número suficiente de experimentos. Ésta permite clasificar los datos por experimentos, por género e incluso por condición auditiva. Esta interfaz se muestra en la Figura 2.5.



**Figura 2.5:** Interfaz para realizar estadísticas

Como indica la Figura 2.5, solo se encuentran habilitados los campos cuyos experimentos y planos hayan sido realizados. Esta es una función que se denomina "experimentos.m" y es la única función que puede ejecutarse tanto desde la función "principal.m" como por si sola, es decir se la puede ejecutar sin necesidad de llamarla desde la función "principal". Para poder realizar las estadísticas esta función tiene como entrada la matriz con absolutamente todos los datos obtenidos al realizar los experimentos. Por esta razón puede ejecutarse sola, ya que puede buscar la sesión guardada (base de datos) y obtener todos los datos necesarios para su ejecución.

### 3. RESULTADOS, CONCLUSIONES Y RECOMENDACIONES

#### 3.1. RESULTADOS

Una vez se haya realizado el diseño de las interfaces gráficas se procede a realizar los experimentos. Dos experimentos para localización absoluta y dos para localización relativa en los planos horizontal, frontal y mediano. Antes de realizar cualquiera de los cuatro experimentos se debe ingresar el usuario, colocar un nombre en el ID de sesión para que los resultados mostrados se vayan guardando en la base de datos con ese nombre. Con la información ingresada se procede a realizar los experimentos que se describen a continuación:

##### 3.1.1. Experimentos de localización absoluta

###### Experimento 1

En la sección 2.3.1 se ha explicado el procedimiento del Experimento 1 donde el usuario no conoce la ubicación del sonido y debe reportar dónde lo escuchó. En la Figura 3.1 se muestra el resultado de la realización del experimento.

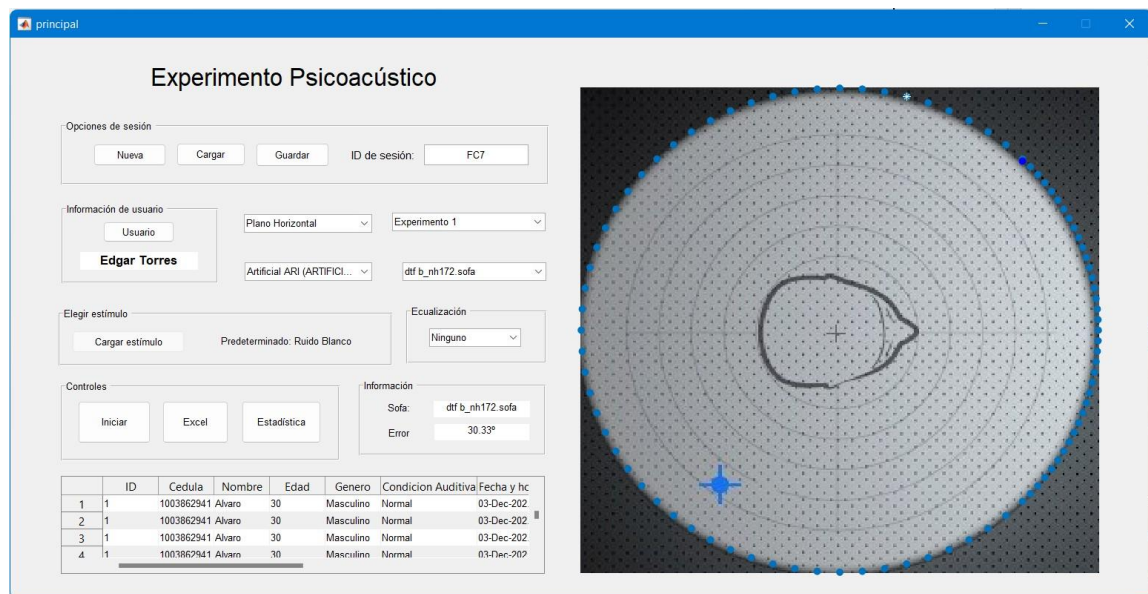


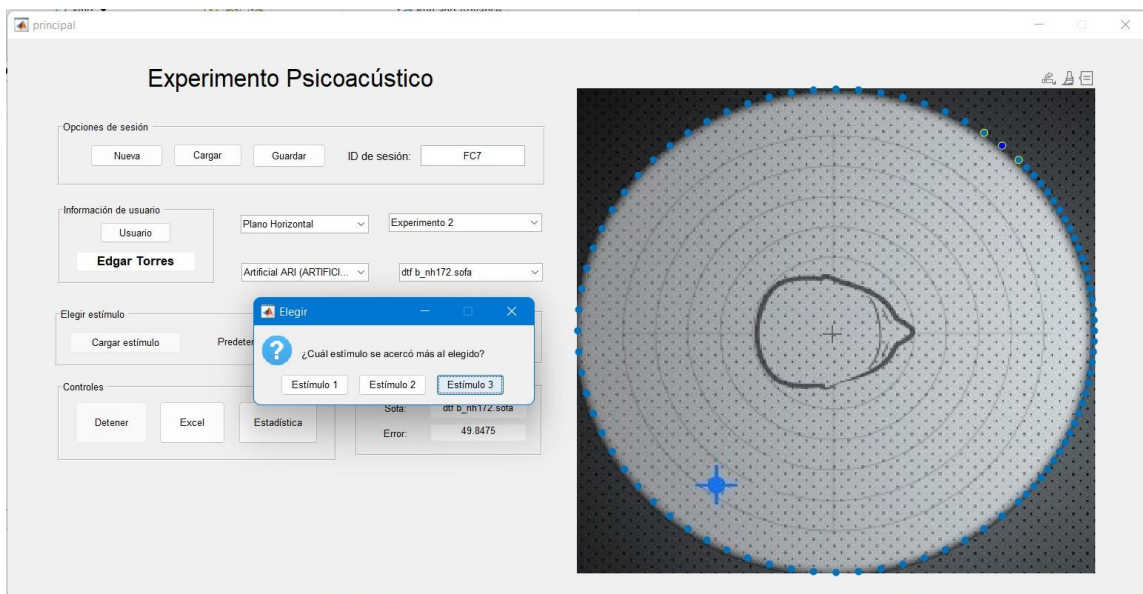
Figura 3.1: Experimento 1 en el plano horizontal

En la Imagen 3.1 se muestra de color celeste las posiciones que puede tomar la fuente sonora, de color azul la posición que el usuario percibió el sonido y de color blanco la posición real de la fuente sonora. Adicionalmente, se puede observar en el campo "Información" el archivo SOFA elegido y el error en grados (°) resultado entre el valor real y el valor percibido. Mientras se vayan realizando los experimentos, la información de éstos aparecerán en la

tabla inferior mencionada en la sección 2.5 con los datos del usuario y experimento realizado.

## Experimento 2

En la sección 2.3.1 se indica el procedimiento de este experimento en el cual el usuario podrá seleccionar la ubicación de la fuente de audio y el programa le presentará tres sonidos en localizaciones cercanas a la posición elegida y el usuario deberá elegir el audio que más se acerque a la posición seleccionada. En la Figura 3.2 se muestra el resultado del experimento realizado.



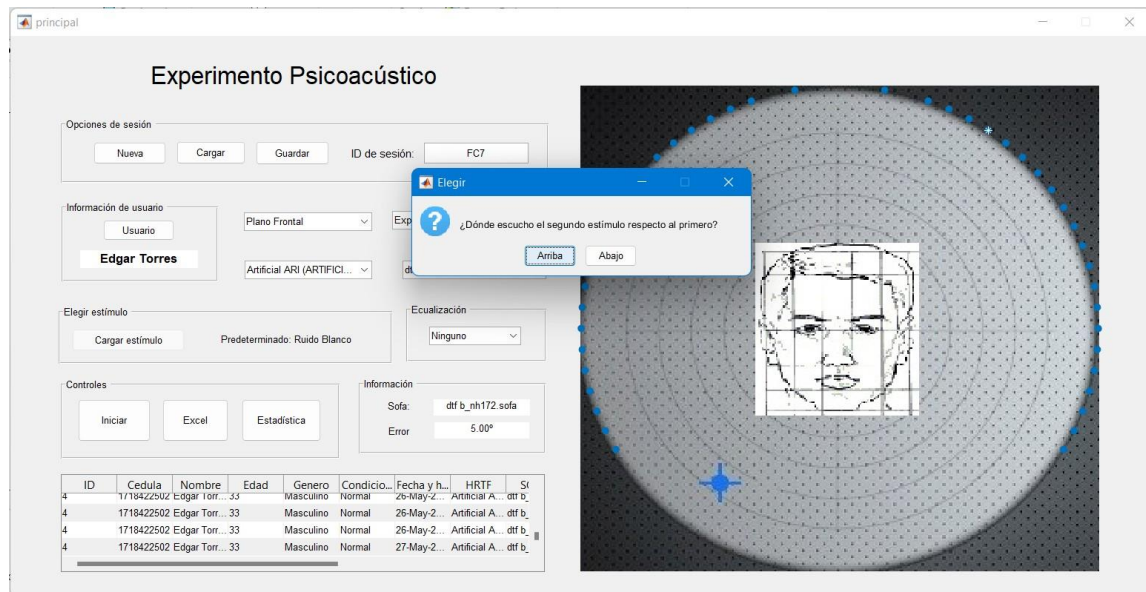
**Figura 3.2:** Experimento 2 en el plano horizontal

Como se puede observar en la Figura 3.2 de color celeste se muestran todas las posiciones que el usuario puede elegir la ubicación de la fuente de sonido. Los círculos amarillos muestran las tres posiciones, de color azul la posición seleccionada por el usuario y de color blanco la posición real. De igual manera en el campo de "Información" se observa el archivo SOFA utilizado y el error en grados ( $^{\circ}$ ) obtenido entre el valor real y el valor escogido.

### 3.1.2. Experimentos de localización absoluta

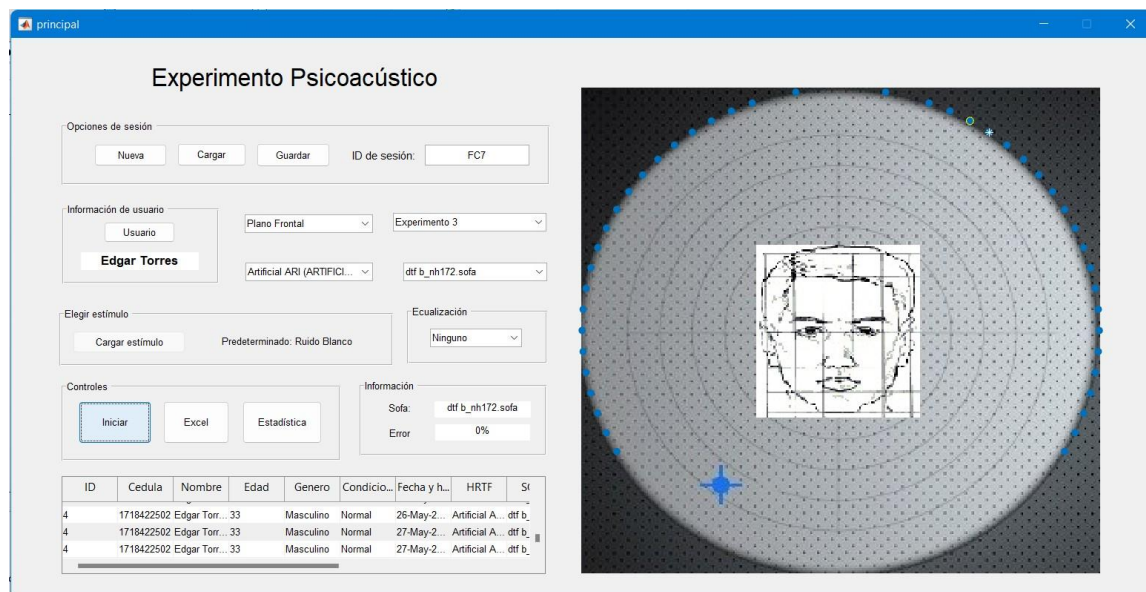
## Experimento 3

Como se indicó en la sección 2.3.2 en este experimento el usuario deberá identificar la posición de un segundo audio respecto al primero audio cuya posición es conocida. En la Figura 3.3 se muestra el cuadro de diálogo que se muestra al usuario para preguntar dónde escuchó el segundo sonido con respecto al primero.



**Figura 3.3:** Experimento 3 en el plano frontal

Una vez que el usuario haya elegido la posición según su criterio el *toolbox* le mostrará el resultado de este experimento y se muestra en la Figura 3.4.



**Figura 3.4:** Resultado del experimento 3

En la Figura 3.4 la posición inicial se muestra de color blanco y la posición elegida se muestra con un círculo amarillo. Además, los datos se muestran en el cuadro de Información y en la tabla de datos ubicada en la parte inferior de la interfaz. Para la presente explicación, en el experimento realizado se ha acertado con la posición, por lo tanto, el valor del error es 0%.



## Experimento 4

Igualmente como se explicó en la sección 2.3.2 en este experimento se reproducirán dos sonidos consecutivos y el usuario deberá informar si los escuchó en el mismo lugar o no. En la Figura 3.5 se muestra el cuadro de diálogo que le aparece al usuario.

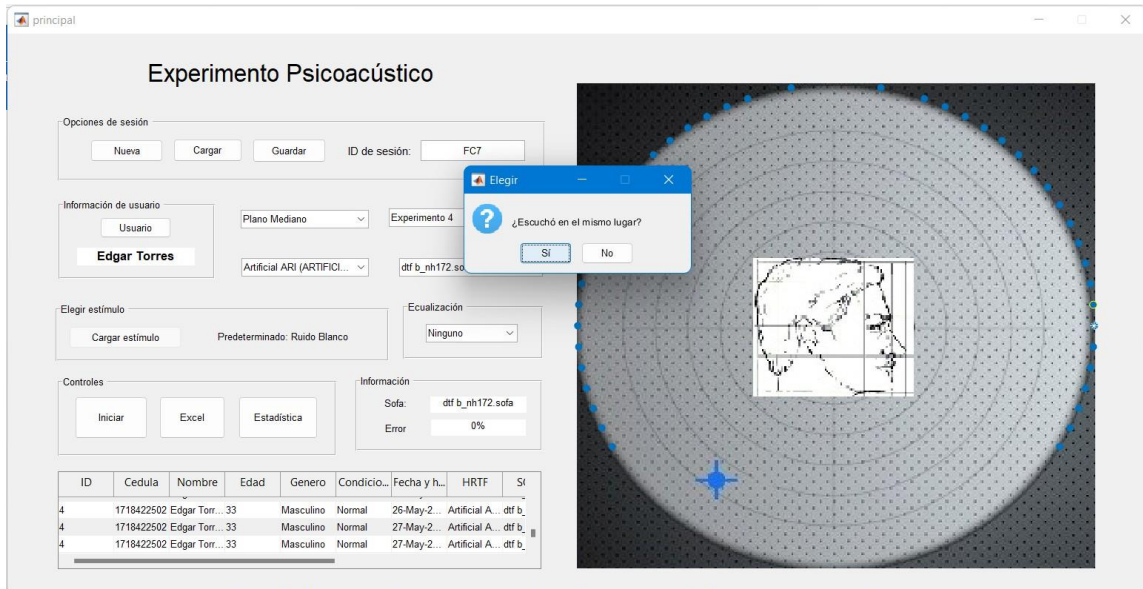


Figura 3.5: Experimento 4 en el plano mediano

Solamente cuando el usuario responda negativamente el proceso se detendrá. Las posiciones disponibles, la posición inicial y las posiciones desplazadas se encuentran pintan de diferentes colores como se muestra en la Figura 3.6.

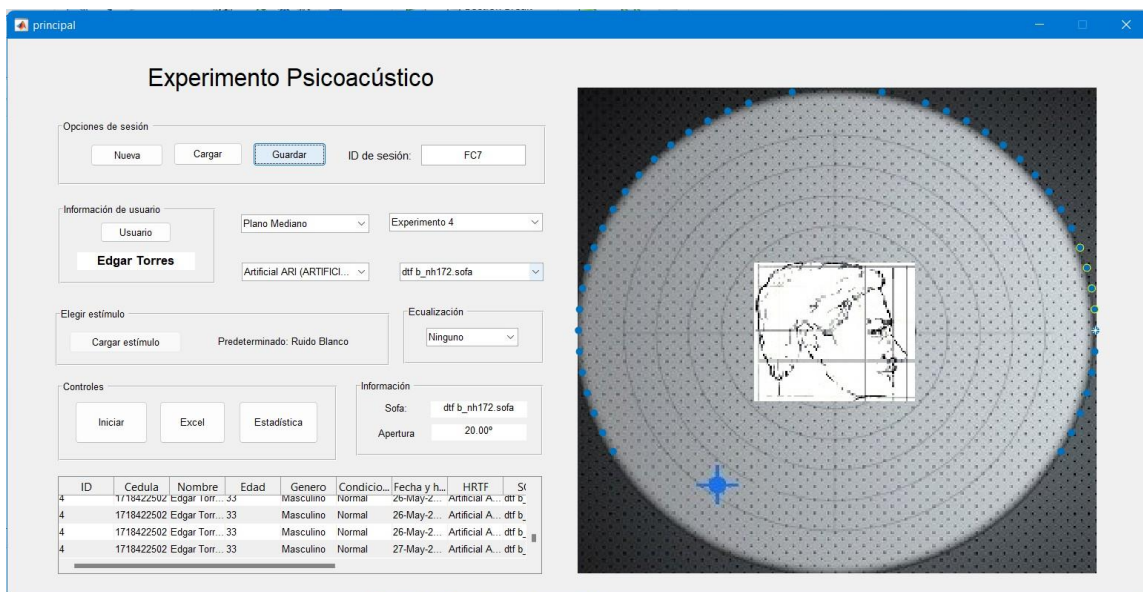


Figura 3.6: Resultado del experimento 4

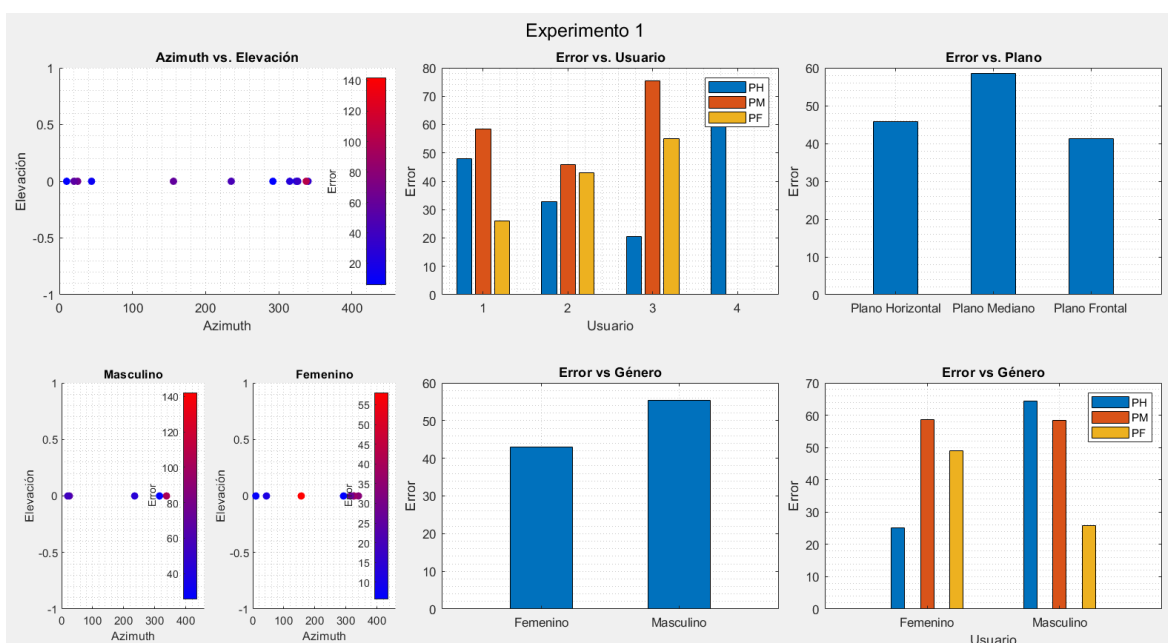
En la Figura 3.6 se observa que, para este ejemplo la posición inicial se encuentra coloreada de blanco, las posiciones disponibles se encuentran de color celeste y las posiciones desplazadas se pintan de un círculo amarillo. Adicionalmente, en la sección de Información se muestra el ángulo de apertura de  $20^\circ$  y en la tabla de datos se va mostrando todos los experimentos antes realizados.

### 3.1.3. Estadísticas

Una vez se hayan realizado suficientes experimentos se pueden obtener datos estadísticos. Al presionar el botón "Estadísticas" descrito en la sección 2.4 se despliega la interfaz de la Figura 2.5. Con esta interfaz se puede obtener gráficos de los resultados los cuales se pueden filtrar por experimento, por plano, por género y por condición.

Para el presente trabajo de titulación se ha realizado distintos experimentos para mostrar los datos estadísticos. Al presionar el botón "Ejecutar" aparecerán seis gráficos de barras y dispersión los cuales se pueden exportar a una hoja de cálculo presionando el logo de Excel, luego se ejecutará el programa y se mostrará dicha hoja de cálculo con seis páginas con los datos correspondientes a cada gráfico.

En la Figura 3.7 se muestra un ejemplo de algunos experimentos realizados y para la obtención de estos gráficos que para este ejemplo se ha filtrado por género:



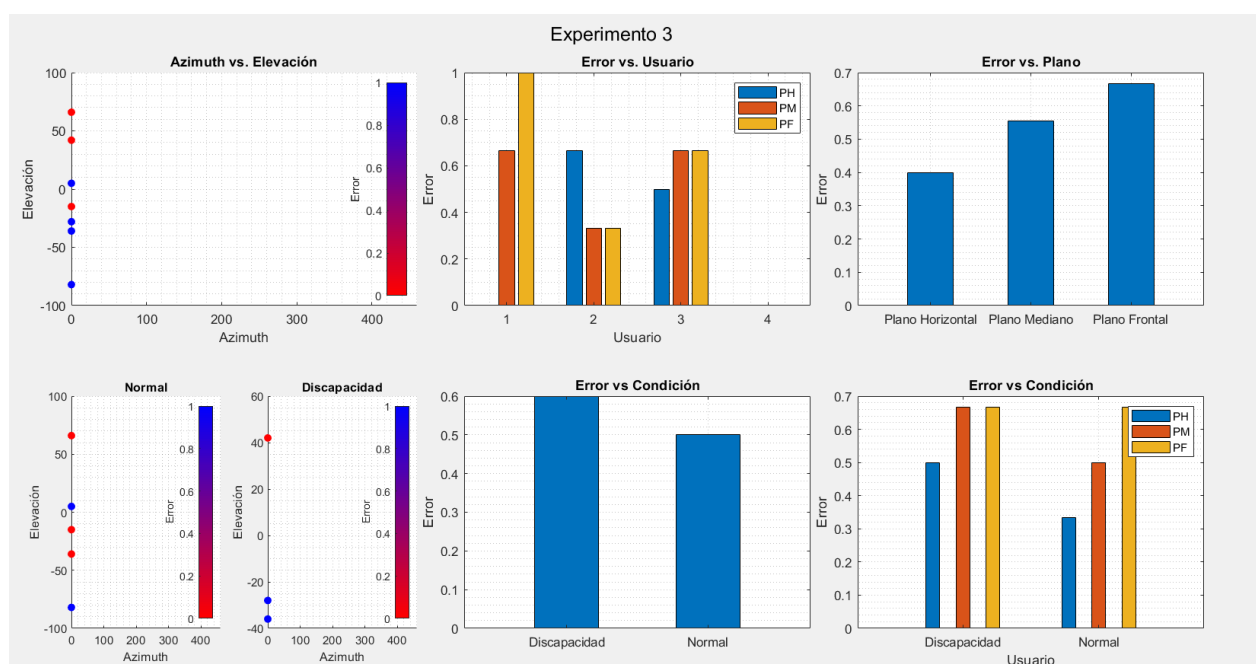
**Figura 3.7:** Estadísticas para el Experimento 1 en el plano horizontal filtrado por género

En este ejercicio se realizaron los experimentos en los tres planos con 4 usuarios entre hombres y mujeres. La interfaz para datos estadísticos que se muestra en la Figura 3.7 está

filtrada por género, en el plano horizontal y para el experimento 1. La Figura 3.7 muestra seis gráficos que se explican a continuación:

1. Azimut vs Elevación: muestra las posiciones del experimento 1 donde el eje X es el ázimut y el eje Y es la elevación. El color indica el error: mientras más cerca del valor real se vuelve más azul y mientras más alejado se encuentra se torna de color rojo.
2. Error vs. Usuario: indica el error generado en la percepción del sonido en los tres planos (Plano Horizontal, Plano Mediano, Plano Frontal: PH, PM, PF respectivamente) por usuario. Para el presente ejemplo se realizaron los con cuatro usuarios por lo que, en el gráfico muestra 4 resultados.
3. Error vs. Plano: muestra el error promediado de todos los usuarios por cada plano.
4. Azimut vs. Elevación por género: muestra las posiciones del experimentos pero divididos por género en dos gráficos. Al igual que el primer gráfico el color indica el valor del error.
5. Error vs. Género: muestra el error promediado en todos los planos para cada género.
6. Error vs. Género: muestra el error promediado para ambos géneros para cada plano.

En la Figura 3.8 se muestra los gráficos estadísticos pero esta vez se ha filtrado por condición auditiva y se ha elegido otro plano y experimento.



**Figura 3.8:** Estadísticas para el Experimento 3 en el plano horizontal filtrado por condición auditiva

Al igual que en la Figura 3.7, en la Figura 3.8 se muestra 6 gráficos los cuales se explican a continuación:

1. Azimut vs Elevación: muestra las posiciones del experimento 3 donde el eje X es el azimut y el eje Y es la elevación. El color indica el error: mientras más cerca del valor real se vuelve más azul y mientras más alejado se encuentra se torna de color rojo.
2. Error vs. Usuario: indica el error generado en la percepción del sonido en los tres planos (Plano Horizontal, Plano Mediano, Plano Frontal: PH, PM, PF respectivamente) por usuario. Para el presente ejemplo se realizaron los experimentos con cuatro usuarios por lo que, en el gráfico muestra 4 resultados.
3. Error vs. Plano: muestra el error promediado de todos los usuarios por cada plano.
4. Azimut vs. Elevación por género: muestra las posiciones del experimentos pero divididos por género en dos gráficos. Al igual que el primer gráfico el color indica el valor del error.
5. Error vs. Condición: muestra el error promediado en todos los planos para la condición auditiva.
6. Error vs. Condición: muestra el error promediado para ambos géneros para la condición auditiva.

En la interfaz mostrada en la Figura 2.5, al presionar el ícono de Excel se despliega una hoja de cálculo con los datos tal como se muestra en la Figura 3.9 cuya parte inferior se observa que cuenta con seis hojas, las cuales pertenecen a los seis gráficos respectivamente.

ID	Plano	Experimento	GroupCount	mean_Valor de referencia	mean_Valor real	mean_Metrica
1	Plano Horizontal	Experimento 1	3	199,1666667	213,0053	47,89602667
3	Plano Mediano	Experimento 1	3	-5	77,78475	58,421678
4	Plano Frontal	Experimento 1	3	43,33333333	20,21256667	25,90776667
5	Plano Horizontal	Experimento 1	3	168,6666667	38,24018	32,75275333
6	Plano Mediano	Experimento 1	4	-5,5	236,5333	45,8813345
7	Plano Frontal	Experimento 1	3	-4,666666667	38,32003333	42,9867
8	Plano Horizontal	Experimento 1	5	260,2	208,6716212	20,4716148
9	Plano Mediano	Experimento 1	3	29,33333333	63,5166	75,46833333
10	Plano Frontal	Experimento 1	3	-56,66666667	185,8771067	55,05239967
11	Plano Horizontal	Experimento 1	6	211,6666667	194,1138	72,51804167

**Figura 3.9:** Hoja de cálculo con los valores mostrados en los gráficos

### 3.2. CONCLUSIONES

- Las HRTFs no personalizadas pueden funcionar o no en varias personas, esto se debe a las características morfológicas únicas que tiene cada ser humano, esto es, la forma y tamaño de la cabeza, torso, pabellón auricular, tamaño de las orejas, incluso la ubicación de los oídos en la cabeza.
- Se puede obtener sonido 3D sin necesidad de contar con varios altavoces al rededor del usuario, solamente con dos altavoces o audífonos estéreo, esto se logra con el cambio de frecuencias del audio proveniente.
- El sonido binaural es lo que más se asemeja al audio 3D inmersivo, ya que estos sonidos se obtienen grabando los sonidos colocando los micrófonos en el canal auditivo del individuo.
- Existen varios institutos de investigación que utilizan las cabezas artificiales para obtener las HRTFs, en las pruebas realizadas usando las HRTFs de cabezas artificiales en el presente proyecto de titulación se han obtenido resultados satisfactorios, lo que significa que este tipo de HRTFs se acercan cada vez más a una cabeza humana real.
- En la actualidad es más fácil trabajar con las HRTFs para implementar APIs o programas como el del presente proyecto de titulación ya que se pueden obtener fácilmente de los repositorios oficiales de manera libre.
- La funcionalidad GUIDE que posee Matlab para elaborar programas con interfaces gráficas permite al usuario crear programas de manera intuitiva, aunque muy pronto será reemplazada por la nueva funcionalidad App Designer. Una vez GUIDE sea removida de Matlab, las aplicaciones hechas en GUIDE no se podrán editar aunque sí se podrán ejecutar.
- El presente proyecto de titulación cuenta con un gran número de funciones (archivos ".m") que fueron hechas específicamente para el programa, sin embargo éstas pueden usarse para otros fines, incluso queda como base para futuros proyectos de titulación.
- Al realizar las pruebas con usuarios y en las investigaciones realizadas se concluye que el pabellón auricular del oído tiene un papel fundamental en la localización de la fuente sonora ya que permite distinguir el sonido proveniente de la fuente sonora ubicada ya sea adelante o atrás del usuario permitiendo mitigar la confusión frente-atrás descrito en el presente proyecto de titulación.

### 3.3. RECOMENDACIONES

- Debido a que este programa fue realizado en Matlab, y que, para futuros usos del mismo requiere tener instalado el programa de Matlab en el computador, se recomienda convertir este software a un archivo ejecutable (En Windows: ".exe" ) de tal manera que se pueda ejecutar sin la necesidad de una instalación del programa. Adicionalmente, para poder ejecutar el programa sin la necesidad de una instalación de Matlab es necesario descargar el paquete de herramientas de Matlab en el computador y para ello es necesario dirigirse a la página de Mathworks [20] y descargar el paquete correspondiente a la versión, que, para el presente proyecto de titulación, fue realizado en la versión R2022a.
- Se recomienda realizar los experimentos usando auriculares que cubran completamente al oído y tenga una amplia respuesta en frecuencia para percibir todo el rango de frecuencias del archivo WAV de esta manera se evita cualquier distracción y ruidos externos que afecten la percepción exacta del sonido.
- Para la correcta ejecución del programa se recomienda colocar los archivos SOFA en carpetas, ya que el programa fue diseñado para buscar los archivos en carpetas y no en la raíz del directorio. De igual manera los archivos de ecualización se recomienda colocarlos en la carpeta destinada ya que es el único directorio donde el programa busca los ecualizadores.
- Para la correcta ejecución del programa se recomienda crear un directorio exclusivo para los archivos SOFA y funciones del programa, tal como se presentó en el presente proyecto de titulación debido a que el programa busca el directorio principal donde están las funciones, archivos (audios WAV e imágenes de los planos) y guarda los resultados, archivos de hojas de cálculo y bases de datos.

## 4. REFERENCIAS BIBLIOGRÁFICAS

- [1] F. Grijalva, L. Martini, D. Florencio, and S. Goldenstein, "A manifold learning approach for personalizing hrtfs from anthropometric features," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 24, no. 3, pp. 559–570, 2016.
- [2] B. Pueo and V. Tur-Viñes, "Sonido espacial para una inmersión audiovisual de alto realismo," 2009-10-15.
- [3] L. Savioja, A. Ando, R. Duraiswami, E. A. P. Habets, and S. Spors, "Introduction to the issue on spatial audio," *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 5, pp. 767–769, 2015.
- [4] F. Grijalva, "Manifold learning for spatial audio rendering," pp. 20–107, 2018.
- [5] F. M. San Victoriano and E. A. López-Poveda, "La audición binaural y el uso de audífonos e implantes cocleares," *Revista de acústica*, vol. 52, no. 3, pp. 3–14, 2021.
- [6] M. S. L. Cerda, "Preparación de setup para experimento de localización de audio 3d en tiempo real," *Informe Proyecto de Título de Ingeniero Civil Electrónico*, pp. 3–14, 2019.
- [7] 3DIO. (19 de Junio de 2019) What is HRTF? [Online]. Available: <https://3diosound.com/blogs/learn-about-sound/what-is-hrtf>
- [8] E. M. Wenzel, M. Arruda, D. J. Kistler, and F. L. Wightman, "Localization using nonindividualized head-related transfer functions," *The Journal of the Acoustical Society of America*, vol. 94, no. 1, pp. 111–123, 1993. [Online]. Available: <https://doi.org/10.1121/1.407089>
- [9] G. Ramos and M. Cobos, "Parametric head-related transfer function modeling and interpolation for cost-efficient binaural sound applications," *The Journal of the Acoustical Society of America*, vol. 134, no. 3, pp. 1735–1738, 2013.
- [10] G. Ramos, M. Cobos, B. Bank, and J. A. Belloch, "A parallel approach to hrtf approximation and interpolation based on a parametric filter model," *IEEE Signal Processing Letters*, vol. 24, no. 10, pp. 1507–1511, 2017.
- [11] H. Gamper, "Head-related transfer function interpolation in azimuth, elevation, and distance," *The Journal of the Acoustical Society of America*, vol. 134, no. 6, pp. EL547–EL553, 2013.

- [12] W. G. Gardner and K. D. Martin, "Hrtf measurements of a kemar," *The Journal of the Acoustical Society of America*, vol. 97, no. 6, pp. 3907–3908, 1995.
- [13] V. R. Algazi, R. O. Duda, D. M. Thompson, and C. Avendano, "The cipic hrtf database," in *Proceedings of the 2001 IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics (Cat. No. 01TH8575)*. IEEE, 2001, pp. 99–102.
- [14] K. Iida, "Comparison of hrtf databases," in *Head-Related Transfer Function and Acoustic Virtual Reality*. Springer, 2019, pp. 171–177.
- [15] P. Majdak, Y. Iwaya, T. Carpentier, R. Nicol, M. Parmentier, A. Roginska, Y. Suzuki, K. Watanabe, H. Wierstorf, H. Ziegelwanger *et al.*, "Spatially oriented format for acoustics: A data exchange format representing head-related transfer functions," in *Audio Engineering Society Convention 134*. Audio Engineering Society, 2013.
- [16] F. Wefers, "A free, open-source software package for directional audio data," in *Proceedings of the 36th German Annual Conference on Acoustics (DAGA 2010), Berlin, Germany, 2010*.
- [17] A. R. Institute. Sofa (spatially oriented format for acoustics). [Online]. Available: [https://www.sofaconventions.org/mediawiki/index.php/SOFA\\_\(Spatially\\_Oriented\\_Format\\_for\\_Acoustics\)](https://www.sofaconventions.org/mediawiki/index.php/SOFA_(Spatially_Oriented_Format_for_Acoustics))
- [18] OPeNDAP. Advanced software for remote data retrieval. [Online]. Available: <https://www.opendap.org/>
- [19] C. Gaspar-Valle, E. Aranda-Bricaire, and M. Velasco-Villa, "Localización absoluta de objetos mediante minimización de errores de mediciones ultrasónicas."
- [20] Mathworks. Matlab runtime. [Online]. Available: <https://la.mathworks.com/products/compiler/matlab-runtime.html>



## **5. ANEXOS**

Dado que el programa fue realizado en el programa Matlab y algunos de los scripts fueron incluidos en el presente proyecto de titulación, se adjunta como anexo digital dichos scripts en extensión ".m" y ".fig" para cada etapa del trabajo. Además se adjunta la base de datos utilizada que se encuentra con la extensión ".mat".