

**LOCALIZACIÓN DE FUENTES SONORAS
MEDIANTE ANALISIS DE SEÑALES**

Juan Pablo Poveda Lotero

**Universidad de San Buenaventura
Facultad de Ingeniería
Programa de Sonido
Bogotá DC
2010**

**LOCALIZACIÓN DE FUENTES SONORAS
MEDIANTE ANÁLISIS DE SEÑALES**
(Análisis de señales binaurales en Matlab)

Proyecto de Grado

Presentado por:

Juan Pablo Poveda Lotero

2001179138

**Universidad de San Buenaventura
Facultad de Ingeniería
Programa de Sonido
Bogotá DC, Enero de 2010**

LOCALIZACIÓN DE FUENTES SONORAS MEDIANTE ANÁLISIS DE SEÑALES

INTRODUCCIÓN

La importancia de la labor que cumple el sistema auditivo en la vida del hombre, radica entre otras cosas, en la maravillosa cualidad que tiene para ayudar a obtener una cantidad inmensa de información, que luego el cerebro procesa a gran velocidad, como resultado de este ejercicio le es posible extraer solo las cosas que al oyente le interesa saber, como por ejemplo, distinguir entre la voz masculina y la femenina, o, entre la voz de un niño y la de un adulto, también la posibilidad de discriminar diferentes acentos o un tipo de pronunciación específica entre muchos, también la proveniencia del sonido, así como la percepción del movimiento de los objetos en el entorno y esto es solo el comienzo.

Con la motivación que inicialmente se tuvo, por esta gran cantidad de posibilidades que ofrece el asombroso sistema auditivo humano, surgió la inquietud por desarrollar un sistema artificial que esté en capacidad de simularlo, no en todas, pero si en una de las más importantes funciones que cumple a diario, “la localización de fuentes sonoras”.

Pero, ¿por qué hablar de la localización de fuentes habiendo tantas otras opciones? Precisamente, porque son muchas las tareas que realiza una persona en un solo día haciendo uso de su capacidad para localizar sonidos, en el hogar, el trabajo, la escuela o en la calle.

El desempeño de todo individuo en la sociedad, está condicionado a las posibilidades que le brindan sus sentidos y desde luego ello incluye al oído, sin embargo, un individuo promedio, la mayor parte del tiempo no es consciente de la presencia de factores que limitan o alteran su capacidad para determinar la procedencia de un sonido, como el ruido, el uso de audífonos, tapones auditivos, y cascos, solo por nombrar algunos, y que dichos factores, más que representar una incomodidad para algunos o una distracción para otros, pueden llegar a convertirse en un peligro real. Pero ¿en qué condiciones se puede hablar de incomodidad y en cuáles se debe hablar de peligro?

Por otra parte, son bien conocidos los sistemas de sonido como el Dolby 5.1 y THX que desde hace algunos años se han implementado en las salas de cine y en los hogares que pueden acceder a esta tecnología y cuyo propósito radica en proporcionar la “ilusión” de un ambiente acústico dinámico, en el cual, el espectador se encuentra inmerso y que es posible implementar gracias a la percepción binaural humana. Sin embargo, ¿están todos los asistentes a una función percibiendo el mismo ambiente acústico, que originalmente ideó el creador de la película?

Pues si bien se podría decir, que el ruido no le ha costado la vida nadie, o que en una sala de cine pocos se quejarían por la mala configuración del sistema de sonido, (lo cual no se puede asegurar del todo), también es acertado decir que se requiere de una herramienta que permita a los ingenieros y diseñadores, tanto de salas de cine como de recintos laborales, determinar los límites dentro de los cuales una persona puede determinar la proveniencia del sonido, ya sea para deleitarse en una función o exigirse en labores que requieren seguridad, agilidad, concentración y precisión.

Con el ánimo de contribuir al mejoramiento de la calidad de vida, de la población expuesta a cualquier clase de riesgo auditivo, que afecte su desempeño en un ambiente seguro, sano y confortable, así como la intención de contribuir a la evaluación de dichos riesgos, se consideró necesario elaborar una herramienta que este en capacidad de simular, mediante un análisis matemático, el fenómeno auditivo humano de la localización de fuentes, con todas las variables que ello implica, reduciendo en una medida importante la cantidad de elementos subjetivos y aumentando los de tipo objetivo.

Para adelantar el proyecto se elaboró un sistema artificial de medición, con capacidad de procesar señales acústicas de un espacio específico en forma digital, haciendo posible estudiar el fenómeno de la localización fuentes sonoras por medio de algoritmos y rutinas generadas en MATLAB, brindando al Ingeniero de Sonido una herramienta que puede simular las condiciones de estudio requeridas por el investigador.

TABLA DE CONTENIDO

1	PLANTEAMIENTO DEL PROBLEMA	7
1.1	ANTECEDENTES	7
1.2	Descripción del problema	8
1.3	Justificación	9
1.4	OBJETIVOS DE LA INVESTIGACIÓN	10
1.4.1	Objetivo General	10
1.4.2	Objetivos Específicos	10
1.5	Alcances y Limitaciones	10
1.5.1	Alcances	10
1.5.2	Limitaciones	11
2	MARCO DE REFERENCIA	12
2.1	Marco Conceptual	12
2.2	Marco Teórico	14
3	METODOLOGÍA	18
3.1	Enfoque de La Investigación	18
3.2	Línea, Sub línea y Campo de la Investigación	18
3.3	Técnicas de Recolección de Información	19
3.4	Población y Muestra	20
3.5	Hipótesis	20
3.6	Variables	21
3.6.1	Variables Independientes	21
3.6.2	Variables Dependientes	21
4	DESARROLLO INGENIERIL	22
4.1	Consideraciones teóricas previas para el diseño de la cabeza binaural con pabellones y torso.....	22
4.1.1	Evaluación del Sistema con un micrófono.....	22
4.2	Evaluación del Sistema con dos micrófonos.....	31
4.2.1	Descripción General del Sistema.....	32
4.2.1.1	Partes del Sistema de Localización Binaural La Unidad de Captura	33
4.2.1.2	La Interfaz Análoga Digital / Digital Análoga	35
4.2.1.3	La Unidad de Procesamiento	35
4.2.2	La Aplicación de software y consideraciones generales de su diseño.....	36
4.2.3	Arquitectura de la Aplicación para Localización de Fuentes Sonoras.....	37
4.2.3.1	Etapas de la Aplicación para Localización de Fuentes Sonoras.....	38
4.2.3.2	Los Módulos de La Aplicación.....	43
4.2.4	Descripción de la Aplicación paso a paso.....	44

4.3	La Rutina de Calibración.....	55
4.4	Arquitectura General del Sistema de Localización.....	58
5	PRESENTACIÓN DE ANÁLISIS Y RESULTADOS	60
5.1	Datos Generales de la Simulación.....	60
5.2	Resultados Obtenidos de las simulaciones.....	60
5.3	Análisis del Ejercicio de Simulación y de los resultados obtenidos.....	69
5.3.1	Los Resultados	69
5.3.1.1	Funcionamiento del Sistema	71
5.3.1.2	Eficacia del Sistema	71
5.3.2	Resultados del Sistema USB- Radar y los antecedentes Documentados.....	74
6	CONCLUSIONES	76
7	RECOMENDACIONES	78
8	BIBLIOGRAFÍA	79
9	GLOSARIO	81

1.1 ANTECEDENTES.

En el año 2004 un grupo de estudiantes de la Universidad de San Buenaventura, adelantó un proyecto de investigación titulado “Aplicaciones y especificación del modelo de captación binaural utilizando un material en su relleno que proporcione un desempeño natural”¹, trabajo mediante el cual se realizó una propuesta para hacer evaluación de recintos con la ayuda de un arreglo de dos micrófonos dispuestos dentro de un modelo de cabeza binaural. La investigación permitió evaluar el desempeño de una cabeza binaural de forma esférica construida en madera con un relleno de fibra de vidrio, para la toma de registros estereofónicos en recintos.

Cuatro años después, dos nuevos grupos de investigación hicieron aportes importantes en el área de la psicoacústica, el primero de ellos tituló su proyecto: “Emulador de un oído humano para determinar factores de riesgo auditivo por el uso prolongado de dispositivo de audio portátil”² y el segundo se llamó “Neuro - estimulador auditivo binaural para tratamiento y optimización de estados cerebrales inducidos de vigilia y concentración”³, en ambos casos se refleja la preocupación por adelantar investigaciones orientadas a entender los procesos relacionados con la audición humana haciendo una evaluación de los riesgos a los que se está expuesto tanto fisiológica como neuronalmente.

No obstante, las diferentes propuestas de investigación ya publicadas por la Universidad de San Buenaventura sede Bogotá y los esfuerzos de estos grupos de estudiantes, no existen referencias documentadas que abarquen de forma puntual o parcial la temática del presente proyecto de investigación.

En lo que respecta a otras entidades o individuos ajenos a la Universidad de San Buenaventura sede Bogotá, existe una gran cantidad de publicaciones en la Internet, que hablan del proceso de ecolocalización, no solo en humanos sino también en ciertos animales que han sido objeto de estudio en este aspecto, como los murciélagos o los búhos. Adicionalmente, se encuentran publicaciones que hacen referencia al estudio del sonido envolvente y a sistemas para reproducción

¹ Arango Giraldo, María Isabel, Cánter Ríos, Nelson Javier, Reyes Monguí, Ricardo Andrés. T 76.2004 / A662a . Bogotá, Universidad de San Buenaventura, Facultad de Ingeniería, Ingeniería de Sonido, 2004

² Zuñiga Libreros, Álvaro José. CD T.IS 760094-08 / Z93e. Bogotá, Universidad de San Buenaventura, Facultad de Ingeniería, Ingeniería de Sonido, 2008

³ Puentes Moreno, David Humberto. CD T.IS 760093-08 / P976n. Bogotá, Universidad de San Buenaventura, Facultad de Ingeniería, Ingeniería de Sonido, 2008

de sonido 3d, sin embargo, la gran mayoría de estos trabajos carecen de la rigurosidad metodológica y científica suficiente para ser referenciados en el presente proyecto. El uso de cabezas binaurales es en la actualidad ampliamente conocido y algunas de sus aplicaciones más destacadas son: la evaluación de sistemas de sonido estéreo, estudios de exposición al ruido y evaluación de recintos en general.

Existen otras publicaciones relacionadas con el tema de la investigación en el banco de documentos de la AES y están disponibles para quienes tengan acceso por suscripción a los mismos, no obstante, se ha mencionado en el presente trabajo, la investigación desarrollada por la Universidad de Aalborg, en la que se adelantaron experimentos de localización con diferentes tipos de unidades de captación binaural, como se indica en la sección 5.3.2 Resultados del sistema USB Radar y los antecedentes documentados.

1.2 DESCRIPCIÓN Y FORMULACIÓN DEL PROBLEMA.

Existe un estancamiento en la generación de nuevas tecnologías y recursos técnicos que brinden al ingeniero de sonido herramientas de valoración objetiva en el estudio de recintos y fenómenos acústicos. En la actualidad la facultad de ingeniería de sonido carece de ciertos elementos de medición que se precisan para adelantar investigaciones en diferentes ramas de la acústica, psicoacústica y la electroacústica, un ejemplo de ello es la necesidad de una cámara anecoica, una sonda intensimétrica, un acelerómetro y una unidad de captación binaural que esté provista no solo de un material de relleno adecuado, sino también de un soporte estructural eficiente y una aplicación de software apropiada para el procesamiento y manipulación de la información.

Es posible descubrir los alcances reales asociados a la acústica, que puede brindar un software como Matlab, en la elaboración de herramientas que permitan brindarle a la unidad de captación binaural de la universidad de san Buenaventura Bogotá, un entorno digital que este en capacidad de procesar la información que al ingeniero de sonido le interese conocer, para el caso concreto del presente proyecto, la proveniencia del sonido.

Tomando en cuenta lo anteriormente dicho habría que formular la siguiente pregunta:

¿Cómo podría ajustarse una herramienta como Matlab, para simular artificialmente la función de localización de una determinada fuente sonora?

1.3 JUSTIFICACIÓN

Las razones para llevar a cabo el proyecto son las siguientes:

El proyecto se plantea como una posible solución ante las limitaciones técnicas actuales de la facultad de Ingeniería de Sonido en el área de la acústica, porque con la implementación de este sistema, se supliría una necesidad que permitiría poner en funcionamiento un elemento como la cabeza de captación binaural para realizar estudios sobre el desempeño acústico de recintos específicos, como por ejemplo, auditorios, salas de mezcla en estéreo, salas de mezcla y reproducción para sistemas 5.1 y de acuerdo con los antecedentes que existen respecto a este proyecto también prestaría una utilidad importante en el estudio de ruido al interior de vehículos y ambientes laborales. Por otra parte y atendiendo al creciente interés de las empresas por evaluar las condiciones de seguridad al interior de sus instalaciones, se contempla la posibilidad de que este tipo de estudios contribuyan a determinar el nivel de riesgo al que puede estar sometida una población específica, cuando en su medio laboral el ruido u otro factor pueda alterar la capacidad de determinar la proveniencia de los sonidos.

La implementación de una cabeza binaural como las que se encuentran en el mercado requiere una inversión de capital importante, si se incurriera en una inversión de este tipo, cada vez que se necesitara algún tipo de actualización, refacción, o mantenimiento del sistema, el tiempo y los costos asociados a estos procesos que son inevitables, lo harían un sistema poco eficiente y bastante dependiente de los recursos del proveedor. Por otra parte, un sistema diseñado y sometido a evaluación por estudiantes de la misma facultad, representa no solo una reducción de los costos, sino que además, incrementa el nivel académico y permite desarrollar nuevas aplicaciones para un software como matlab, que tiene librerías especializadas para aplicaciones como el procesamiento de imágenes pero que carece de unas librerías adecuadas para el estudio de fenómenos relacionados con acústica y psicoacústica.

Varios elementos que aportan a la realización exitosa del proyecto son: la disponibilidad abundante de fuentes de información (biblioteca USB, bibliotecas externas), trabajos relacionados con el tema de investigación, publicados en Internet, recurso humano capacitado (representado en el cuerpo docente) para orientar el desarrollo del proyecto, igualmente, equipos de cómputo y software, así como laboratorios e instrumentos de medición para adelantar las mediciones y ensayos necesarios tanto dentro como afuera de la universidad.

1.4 OBJETIVOS DE LA INVESTIGACIÓN

1.4.1 Objetivo General:

- Desarrollar un sistema que permita determinar la ubicación de una fuente sonora real, mediante el análisis de señales digitales en Matlab, haciendo uso de una unidad de captación binaural adecuada.

1.4.2 Objetivos Específicos:

- Diseñar y/o modificar una unidad de captura de sonido adecuada, teniendo en cuenta la función que cumplen los pabellones auditivos y el dispositivo de soporte en el proceso de localización de fuentes sonoras.
- Desarrollar una aplicación en Matlab que permita procesar las señales provenientes de la unidad capturadora, para determinar la ubicación de la fuente.
- Elaborar un informe que contenga detalles de las mediciones y pruebas realizadas al sistema, indicando los avances conseguidos, las dificultades más significativas, así como las conclusiones y sugerencias.

1.5 ALCANCES Y LIMITACIONES.

1.5.1 Alcances

El alcance del proyecto llega hasta el diseño y la puesta en marcha de un sistema artificial que esté en capacidad de determinar el origen de una fuente sonora.

Para ello, se precisa un diseño especial de una unidad capturadora de sonido (cabeza binaural o en su defecto un arreglo de micrófonos que simulen una cabeza artificial), la implementación o adecuación de una interfaz que convierta las señales de micrófono en señales digitales y una aplicación en Matlab que procese los datos pertinentes (no necesariamente en tiempo real) y suministre información suficiente para determinar la proveniencia del sonido.

El proyecto permite ampliar el campo de acción de herramientas como Matlab, que a pesar de ser un software de alto nivel para ingeniería, no posee librerías exclusivas para el estudio en acústica y psicoacústica, como si las tiene por ejemplo para el procesamiento de imágenes.

El proyecto concluye con la elaboración de un informe final que resume el proceso de investigación en sus etapas de diseño, puesta en marcha y evaluación del sistema, que conducirán a conclusiones con respecto a la funcionalidad de Matlab en el estudio de fenómenos referentes a la acústica y la psicoacústica, así como la confiabilidad de los algoritmos, el criterio de localización y las herramientas de hardware empleadas en el proceso.

1.5.2 Limitaciones

La principal limitación, para este proyecto se encuentra en la fidelidad que suministrarán los resultados, como consecuencia del estado actual del modelo de cabeza binaural con que cuenta la universidad. Del mismo modo, no se puede determinar la variación que podría presentarse en los resultados si se adelantaran las simulaciones en una cámara anecoica ante la imposibilidad actual de acceder a este recurso.

2.1 MARCO CONCEPTUAL

El Método de Análisis de la Aplicación: Para el desarrollo de la presente investigación, se han implementado una serie de procedimientos de los cuales no se tiene antecedentes en el desarrollo de sistemas de localización binaural al interior de la Universidad. Estos procedimientos, están basados en deducciones que inicialmente se hicieron intuitivamente y en una segunda instancia se reafirmaron con la evaluación de los resultados obtenidos en las pruebas realizadas al sistema. **El Método de Convolución**, por ejemplo, es un procedimiento de comparación entre las señales del oído izquierdo y derecho, aplicando entre éstas una convolución real, en donde se asume que una de las señales es la entrada a un sistema ($x(t)$) y la otra es la respuesta al impulso ($h(t)$). Teniendo en cuenta la propiedad conmutativa de la convolución, la determinación de cuál de las dos señales es $x(t)$ o $h(t)$, no es de importancia.

Existe un segundo método, denominado **Función de Transferencia**, que no debe confundirse, con el modo de operación del sistema que lleva su mismo nombre. Este procedimiento es complementario al anteriormente explicado y hace uso de una relación matemática Entrada – Salida para determinar la ubicación de la fuente⁴.

Los Modos de Operación del Sistema: Para el funcionamiento del sistema de Localización, se desarrollaron dos modos de operación, que dependen del tipo de fuente que se emplee para hacer el análisis binaural. El método de **fuentes desconocidas** es un procedimiento de localización para una fuente arbitraria, en tanto que el modo **función de transferencia** es para la localización de una fuente conocida porque se utiliza una señal reproducida por el mismo sistema. Este último método, fue tenido en cuenta por la necesidad de trabajar con señales fácilmente identificables ante la presencia de ruido y la posibilidad que ofrece matlab de enviar y recibir señales de modo simultáneo.

⁴ Véase sección 4.1 Consideraciones teóricas previas para el diseño de la cabeza binaural con pabellones y torso.

Matriz de Comparación: tal como su nombre lo indica se trata de una matriz generada en matlab, que puede tener dimensiones de n filas por n columnas, pero su tamaño se limita por los recursos de hardware disponibles en el equipo. Esta matriz almacena en cada una de sus celdas una función de transferencia específica asociada a una posición de la fuente con respecto a la cabeza. Se dice que es de comparación porque cuando el software hace la convolución entre las señales provenientes de los micrófonos y determina una función de transferencia para éstas, la compara con cada una de las que se encuentra almacenada en la matriz.

Programa: Este término, en el caso particular de la presente investigación, obedece a un conjunto de pasos que se realizan y que en su totalidad constituyen un procedimiento; dicho procedimiento obedece a una necesidad o un conjunto de necesidades identificadas durante la puesta en marcha del sistema. Por ejemplo, el proceso de calibración del sistema obedece a la necesidad de balancear los diferentes componentes de mismo con respecto a una referencia, luego, esto se constituye como un programa en el cual se debe contar con unos elementos específicos que intervienen en una secuencia también determinada. Bajo estas consideraciones, el procedimiento de localización representaría otro programa diferente al de calibración y también diferente al programa para cargar las funciones de transferencia a la matriz de comparación.

Software, Aplicación de Software y Algoritmo: son tres términos que conviene aclarar para el desarrollo del presente trabajo. En primer lugar, se considera como algoritmo a cada secuencia de instrucciones escritas en lenguaje de programación, que tiene como fin realizar una tarea específica con los recursos lógicos de matlab. La Aplicación de Software, es el conjunto de algoritmos que hacen parte de un paquete de instrucciones desarrolladas con el fin de ejecutar varias tareas de forma simultánea o secuencial, automática o asistida, dentro de parámetros de diseño claros y con recursos específicos de hardware. El software es la plataforma sobre la cual se desarrolla una aplicación con el propósito de satisfacer una necesidad específica. Luego el software empleado en el presente proyecto es Matlab, la aplicación desarrollada es un localizador binaural basado en análisis de señales, que para ello cuenta con una serie de algoritmos que controlan las entradas-Salidas de una tarjeta de sonido y la unidad aritmeticológica del computador.

Relación Señal / Ruido: como en cualquier situación en la que se realice una medición acústica, indica por medio de una expresión matemática, la relación en decibeles, de la energía percibida por la señal deseada con respecto al nivel de ruido presente en el lugar de la medición.

En el caso de la aplicación de software para localización, la relación Señal/Ruido, se debe tener en cuenta en las tres condiciones de trabajo del sistema que son: la calibración, la alimentación de la matriz de comparación y la captura para localización.

La relación Señal/Ruido es una variable que influye de manera importante en la realización del proceso de localización.

2.2 MARCO TEÓRICO

La importancia de los pabellones auditivos y el diseño de la unidad de captura.

El estudio de sonido envolvente para aplicaciones de reproducción en 3D, pone de manifiesto la enorme importancia de la función que cumplen los pabellones auditivos en la localización espacial de fuentes sonoras.

“Tanto la cabeza, pero principalmente el pabellón auditivo, modifican el espectro de los sonidos en dependencia del ángulo de incidencia del sonido con respecto a la cabeza” (García A. 2005, p11)

Su tarea se extiende más allá de los intereses de la acústica puesto que estos órganos intervienen en la modificación del espectro percibido por el sistema auditivo y por ello influye en la interpretación que se hace de la información en el cerebro, lo cual hace parte de los intereses de otra disciplina conocida como neuropsicología.

Los pabellones auditivos conforman un dispositivo auxiliar que complementa la información reproducida por el tímpano, considerando que esta estructura cartilaginosa está en capacidad de atenuar y/o reforzar ciertas frecuencias bajo determinadas condiciones de procedencia.

“El pabellón auditivo actúa como una cavidad de resonancia, donde amplifica algunas frecuencias, y su geometría genera ciertos efectos de interferencia que atenúan otras frecuencias. Dicho de otra forma, su respuesta en frecuencia depende de la dirección” (García A., p12)

El fenómeno neuro biológico de localizar fuentes sonoras es producto de un proceso que no solo involucra al sistema auditivo como se le conoce usualmente, oído externo, medio, interno y nervio auditivo; al proceso se vincula una importante función del cerebro asociada con funciones de

memoria que están en capacidad de comparar sensaciones audibles pasadas con las presentes.

“En realidad, el cerebro se ha entrenando toda la vida, y tiene memorizados una serie de parámetros basados en la experiencia y personalizados para cada uno” (García A., p4)

Lo anterior no es del todo desconocido para el ingeniero de sonido, si se considera que el caso del ser humano, se puede tratar como un sistema con dos terminales de entrada, para las cuales se genera una única respuesta al impulso por cada posición de fuente, como resultado de la correlación cruzada entre las señales percibas por los oídos izquierdo y derecho.

Pero no solamente es importante el pabellón auditivo, también cuenta la influencia que ejercen otros elementos como el torso y la misma cabeza, como se menciona a continuación:

“La acústica binaural estudia una estereofonía determinada, que es la que reproduce la audición humana. Se realiza con maniquís con torso, cabeza y orejas para simular un cuerpo humano. En la entrada del canal auditivo se sitúan micrófonos de medida para poder registrar las señales justamente antes de su paso por el canal. Estas señales han sufrido un modelado ("spectral shapping") debido a las reflexiones en el torso, cabeza y pabellón auditivo”. (Ordoyo A., P3)

En el anterior fragmento del citado trabajo se hace énfasis en la necesidad de simular de la mejor manera posible las condiciones de audición humanas, desde luego, existen ciertas mediciones que no requieren un grado de similitud elevado entre la cabeza artificial y una cabeza humana promedio, como por ejemplo **la fracción de energía lateral** de un determinado punto de un recinto, pero para mediciones en donde se requiera evaluar el manejo de la espacialidad de un determinado sistema de amplificación, las mayores similitudes entre el modelo artificial y un individuo real arrojarán resultados más veraces.

Con respecto a la capacidad de localización vertical está claramente identificada la labor que cumplen los pabellones auditivos, como se cita a continuación:

“el contorno del pabellón auditivo crea una serie de reflexiones que son conducidas al oído interno. Estas reflexiones causan coloración del sonido en su recorrido al oído interno. La identificación tonal de estas reflexiones está codificada en el cerebro como una especie de mapa que no provee del sentido de una imagen sonora (Everest, 1994, p.54). Pareciera que el

sistema Oídos/Ojos/Cerebro ha aprendido con el tiempo a hallar la correlación entre las imágenes visuales y las imágenes sonoras y que esta coloración está registrada como un mapa en la memoria. Esta respuesta, Head Related Transfer Function (HRTF) opera independientemente para cada oído” (McCarthy, 2007, p.154).

El aporte sustancial de este autor a la investigación radica en la afirmación que hace asegurando que con cada oído se realiza una determinación vertical independiente de la fuente.

“La posición horizontal de una fuente sonora aislada puede ser localizada con facilidad y alta precisión. Cuando el oído es estimulado por múltiples señales que arriban el proceso de localización se hace más complejo. Si las señales están suficientemente cercanas en tiempo, las señal se mezclarán en el oído y la HRTF será modificada como una suma de múltiples fuentes. Cada oído independientemente determina la localización vertical. Esto contrasta con la localización horizontal la cual está basada en la diferencia entre los oídos (Everest, 1994,pp.51-53). Como resultado las fuentes que están desplazadas horizontal y verticalmente tendrán conflictos para su localización vertical conjuntamente con la localización y diferenciación horizontal.” (McCarthy, 2007, pp.156-157).

Podemos considerar cada oído por aparte o, en el caso especial en el que se tienen múltiples fuentes dispuestas verticalmente pero ubicadas exactamente sobre la línea central horizontal.

Si la señal de dos fuentes verticalmente desplazadas arriban con niveles iguales a los valores de la HRTF presentan un conflicto que involucra la posición de las fuentes en medio de estas. Esto no es experimentado como una fuente puntual, como ocurriría si se tuviera un solo parlante en el centro. En lugar de esto, la imagen sonora es experimentada como una prolongación espacial de la fuente sobre el plano vertical que se encuentra entre los parlantes.

Si las fuentes presentan niveles desiguales, el parlante dominante va a conservar la imagen sonora. Esto podría deberse a la existencia de una HRTF reconocida para la señal que arriba con más fuerza.

Si los tiempos de llegada son desiguales, no se tiene un mecanismo para discernir cual señal arriba primero. Las primeras reflexiones usualmente caen dentro de esta categoría. Tan prolongado como sea el sonido dominante sobre las reflexiones la imagen favorecerá al sonido directo. Solo cuando las señales se encuentran suficientemente desiguales en tiempo son percibidas como sonidos distintos, cada uno con una ubicación de fuente también diferente.

El dominio del nivel sobre el tiempo contrasta con el mecanismo horizontal, el cual usa un sistema de canal dual para la localización. En el sistema horizontal de 2 canales, el tiempo de arribo y el nivel relativo es el criterio dominante.

Localización Adelante / Atrás

El rol secundario de los pabellones auditivos es el de proveer localización frontal / trasera (Everest, 1994,p.52) El mecanismo para esto radica en la anatomía del pabellón el cual provee algo de direccionalidad en alta frecuencia y añade otro tipo de HRTF a la ecuación.

Localización horizontal

“El mecanismo horizontal es más sensible que el vertical. La disposición y separación de los oídos en los costados opuestos de la cabeza provee automáticamente un sensor espacial. Cualquier posición en horizonte puede ser encontrada por triangulación, el sonido es escuchado en dos sitios y la fuente es detectada por la diferencia en sus arribos”

La ubicación percibida de la imagen sonora horizontal depende tanto de la diferencia de tiempo como de nivel en los oídos. El tiempo relativo es denominado Inter-aural time difference (ITD), mientras que el nivel relativo es conocido como Inter-aural level difference (ILD). El ITD es el factor dominante en la localización en baja frecuencia donde las longitudes de onda son suficientemente grandes como para bordear la cabeza sin sufrir pérdida de nivel aún estando ligeramente corridas temporalmente.

La sincronización del ITD está mapeada en el cerebro como una representación polar para la ubicación de la fuente. Las altas frecuencias son menos capaces de refractarse sin pérdida de nivel alrededor de la cabeza y por ello las diferencias de nivel son sustancialmente suficientes para proveer la detección necesaria”. (McCarthy, 2007, pp157-158).

3.

METODOLOGIA

3.1 ENFOQUE DE LA INVESTIGACIÓN

El enfoque de la presente investigación es de tipo empírico analítico, puesto que el interés del proyecto está orientado hacia la interpretación y procesamiento de señales digitales haciendo uso de los recursos técnicos disponibles en la universidad, como lo son, la cabeza binaural y el software Matlab, para determinar la localización de una determinada fuente sonora.

3.2 LÍNEA, SUBLÍNEA Y CAMPO TEMÁTICO DE LA INVESTIGACIÓN.

Campo temático del Programa: Diseño de Sistemas de Sonido

La razón por la cual se considera que el campo temático se encuentra dentro de esta área radica en el planteamiento del objetivo general del proyecto, “Desarrollar un sistema que permita determinar la ubicación de una fuente sonora real, mediante el análisis de señales digitales en Matlab”, para lo cual se hace necesario manipular las herramientas matemáticas, electroacústicas y de software con que cuenta la universidad, en aras de llevar a cabo una investigación que permita comprender el fenómeno psicoacústico de la localización de sonidos, con el diseño de un sistema artificial.

Sublínea de la Facultad: Procesamiento de señales digitales y/o analógicas

Tomando en cuenta lo anterior, la sublínea que se ajusta a las características y al perfil de la investigación es la correspondiente al procesamiento de señales digitales y/o analógicas, ya que existen dentro del sistema que se pretende desarrollar, dos etapas claramente identificadas,

la primera, corresponde a la captura y manipulación de las señales analógicas con ayuda de una unidad de captura como lo puede ser una cabeza binaural y a continuación, una etapa dedicada al procesamiento y posterior interpretación de las señales muestreadas haciendo uso de software.

Línea de investigación de la Universidad: Tecnologías actuales y sociedad

Teniendo como soporte la sublínea anterior, se puede identificar la línea de investigación que mejor se ajusta a los intereses del proyecto, que por su evidente enfoque técnico y científico, corresponde a Tecnologías actuales y sociedad.

3.3 TÉCNICAS DE RECOLECCIÓN DE INFORMACIÓN

Se plantean dos técnicas de recolección de información, como se enuncia a continuación:

Técnica inicial de fuentes virtuales:

Corresponde a las etapas de planeación y diseño del sistema, durante las cuales se trabajará con fuentes disponibles principalmente en la Internet. De allí se escogerán y seleccionarán los documentos que aporten a los intereses de la investigación, para lo cual se evaluarán tres aspectos a cada documento, así:

Pertinencia del documento: Es decir, que la información contenida en el mismo atañe a alguno de los siguientes tópicos: electroacústica aplicada a simulación de espacios, procesamiento análogo y/o digital de señales, captación y/o simulación binaural,

Validez científica del documento: Se requiere que toda fuente que aporte información a la investigación tenga un respaldo científico sólido, para lo cual, todo documento debe estar respaldado por una entidad acreditada o reconocida, de modo que a pesar de la existencia de sitios web tales como foros y páginas personales, solo se tendrán en cuenta las que tengan el respaldo de facultades y escuelas de ingeniería, empresas privadas relacionadas con acústica

y electroacústica, así como entidades dedicadas a la investigación tales como IEEE y AES.

Soporte bibliográfico: Toda publicación tiene un respaldo bibliográfico, incluso las que se hallan en la red, pero existiendo la posibilidad de que se encuentren documentos incompletos o que por alguna razón no contengan una relación de las fuentes utilizadas, dicho tipo de documentos tendrán que ser omitidos ante la imposibilidad de confrontar la información aportada por los mismos.

Técnica Complementaria de medición:

Esta técnica corresponde a la segunda etapa, en la cual se evalúa el desempeño del sistema (una vez que está en funcionamiento), con respecto a los objetivos propuestos por la investigación.

3.4 POBLACIÓN Y MUESTRA

Debido a que las conclusiones del proyecto, estarán sujetas a los resultados de las mediciones que se hagan del sistema, se hace necesario un tratamiento estadístico de dichos resultados. Lo anterior, no solo facilitará la comprensión de la información, sino que además revelará tendencias y márgenes de error presentes en el diseño.

3.5 HIPÓTESIS

Es posible adecuar una herramienta de ingeniería como Matlab, para determinar la localización de una fuente sonora en un espacio específico. Para conseguirlo, es necesario diseñar un sistema que en primera instancia esté en capacidad de capturar una señal acústica haciendo uso de técnicas especiales de microfónica para simular la captación binaural humana, posteriormente, se requiere de un algoritmo debidamente elaborado en Matlab que tomando las señales provenientes de los micrófonos, realice procesos de comparación y estimación espectral y temporal entre las señales, haciendo uso de técnicas de convolución y correlación. Los datos arrojados por el programa deben conducir a afirmaciones objetivas con respecto a la ubicación de la fuente.

3.6 VARIABLES.

3.6.1 Variables Independientes

La posibilidad de contar con un recinto de condiciones anecoicas puede reducir al máximo el margen de error en los dictámenes que se hagan con la cabeza binaural, no solo porque se disminuye el efecto reverberante sobre las señales a registrar sino que además permitiría trabajar con un aporte mínimo de ruido ambiente.

Por otra parte, las condiciones de igualdad mecánica y eléctrica de los micrófonos empleados, lo cual es de especial importancia para el correcto funcionamiento del modelo de captación binaural, pueden verse afectadas teniendo en cuenta la sensibilidad y fragilidad de estos elementos y sus componentes ante el uso continuo. Por lo anterior y ante la posibilidad de ser expuestos a movimientos bruscos y niveles de presión inapropiados, aumenta la posibilidad de que los micrófonos lleguen a presentar diferencias en su respuesta en frecuencia alejando al sistema de su condición ideal de funcionamiento.

3.6.2 Variables Dependientes

Según la versión de matlab que se utilice, se pueden ver limitados algunos de los recursos gráficos y/o del simulador que están diseñados para configurar y manipular los recursos físicos del computador.

El sistema operativo bajo el cual se trabaje puede disminuir o no, las posibilidades que ofrece matlab en cada una de sus versiones, de configurar y manipular a voluntad los recursos de la tarjeta de sonido. Igualmente, el rendimiento del programa podría verse afectado, si los recursos de memoria y procesador que se requieren para el tratamiento de las señales están comprometidos por las necesidades del sistema operativo para atender otras tareas como funciones gráficas.

4. DESARROLLO INGENIERIL

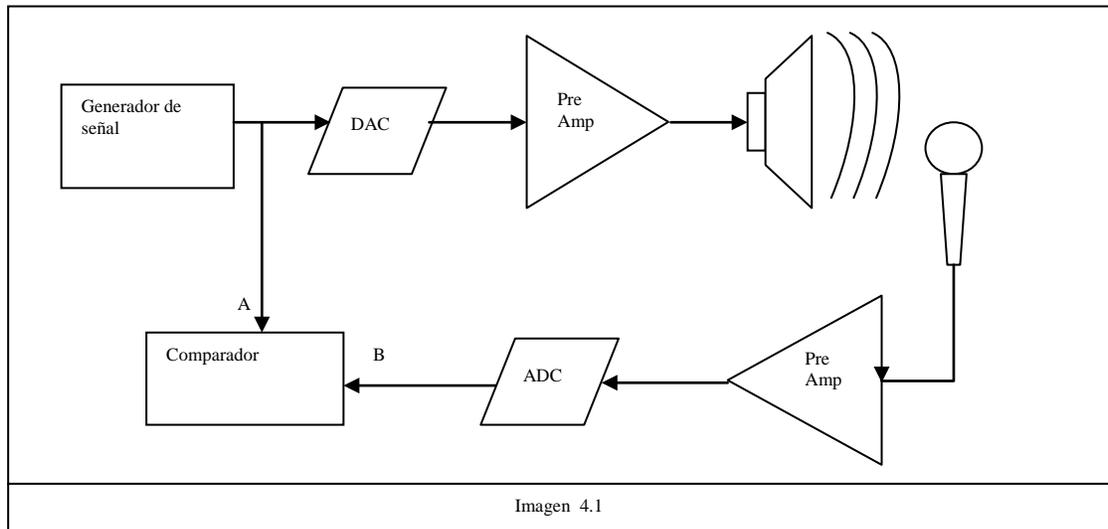
4.1 CONSIDERACIONES TEÓRICAS PREVIAS PARA EL DISEÑO DE LA CABEZA BINAURAL CON PABELLONES Y TORSO

4.1.1 Evaluación del Sistema con un micrófono.

Descripción del Sistema bajo el concepto de Función de Transferencia.

Para llegar a analizar el sistema de localización binaural humano en su totalidad, primero se pretende partir de una situación hipotética en la que una persona cuente con solo uno de los oídos, para estudiar las limitaciones y los alcances de cada uno de los Subsistemas que componen el Sistema general de localización.

Considérese entonces un sistema artificial como el de la imagen 4.1 y ahora supóngase que se pretende determinar con los elementos allí señalados, la localización de la fuente con respecto al micrófono, haciendo uso del principio de función de transferencia.



Puesta en marcha del sistema

El generador emite una señal determinada, tal como un tono de 1 kHz., ésta es conducida a un conversor digital análogo y posteriormente a un preamplificador para su respectiva radiación, la señal (B) que es capturada por un micrófono, se lleva a un segundo preamplificador, luego a un conversor análogo digital y de allí a una unidad de comparación donde también es recibida la señal (A) que está siendo emitida directamente por el generador.

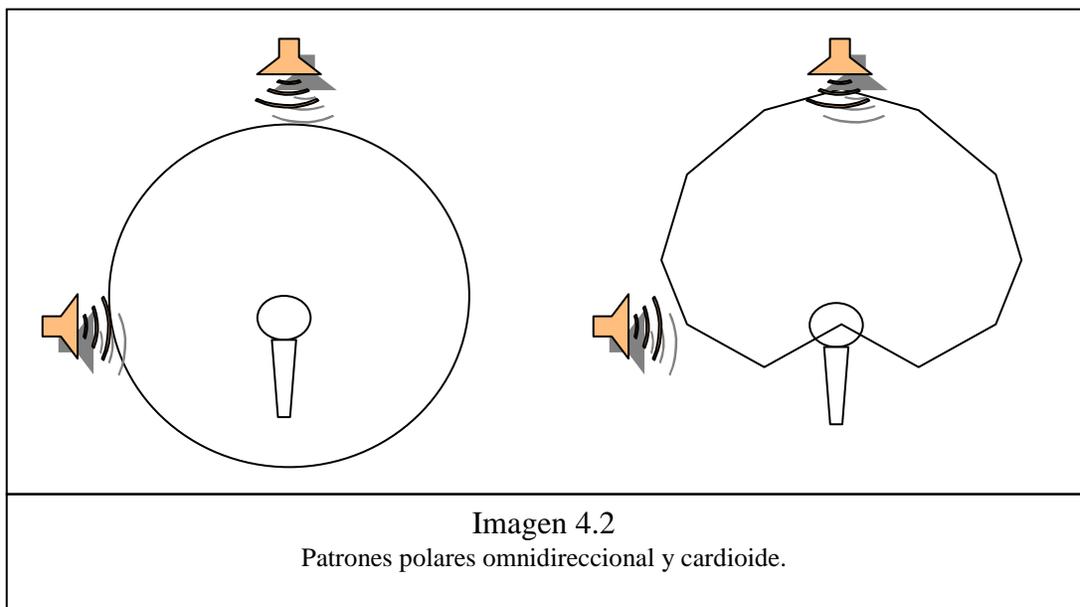
El comparador examina la diferencia de fase y potencia de las señales A y B para luego emitir un dictamen sobre la localización de la fuente. ¿Sería posible afirmar que este sistema tal como ha sido descrito, es eficaz y cumple con el objetivo? Seguramente no, sin embargo, al preguntarse si el sistema es completamente ineficaz, la respuesta también es negativa, ¿por qué?

La razón, que, es una de las causas que motivó a trabajar bajo el concepto de función de transferencia, es la siguiente:

Si se manipula correctamente la unidad de captura (el micrófono) y se hace uso de algunos principios del funcionamiento de la misma, se puede implementar un mecanismo capaz de determinar de forma parcial la localización de la fuente, lo cual da pie para hablar más adelante de la importancia de los pabellones, el torso y la geometría de la cabeza humana.

Considérese nuevamente el sistema de la imagen 4.1, pero esta vez en dos condiciones diferentes, la primera con un micrófono de patrón de captación omnidireccional y la segunda con un micrófono cardioide, como se muestra en la imagen 4.2.

También considérense dos posiciones de fuente diferentes en cada caso, la primera con la fuente a cero grados sobre el eje del micrófono y la segunda a 90 grados sobre el mismo eje y en ambas situaciones se utiliza una señal de un tono de 1 KHz a una distancia y nivel de potencia de radiación constantes.



El criterio de localización que se utiliza en este caso es la relación de nivel de potencia sonora entre las señales A y B, teniendo en cuenta que para una distancia constante y un nivel de radiación también constante, las variaciones de nivel de la señal B con respecto a la referencia A, será una variable en función del ángulo de ubicación de la fuente con respecto al micrófono.

Cuando el sistema tiene el micrófono omnidireccional la variación de nivel que se da es nula para cualquier ubicación de la fuente, sin embargo, para el sistema provisto del micrófono cardioide, si es posible determinar o asociar un determinado ángulo con una relación de nivel de presión sonora específica, es decir, aplica el concepto de función de transferencia para ubicar la fuente.

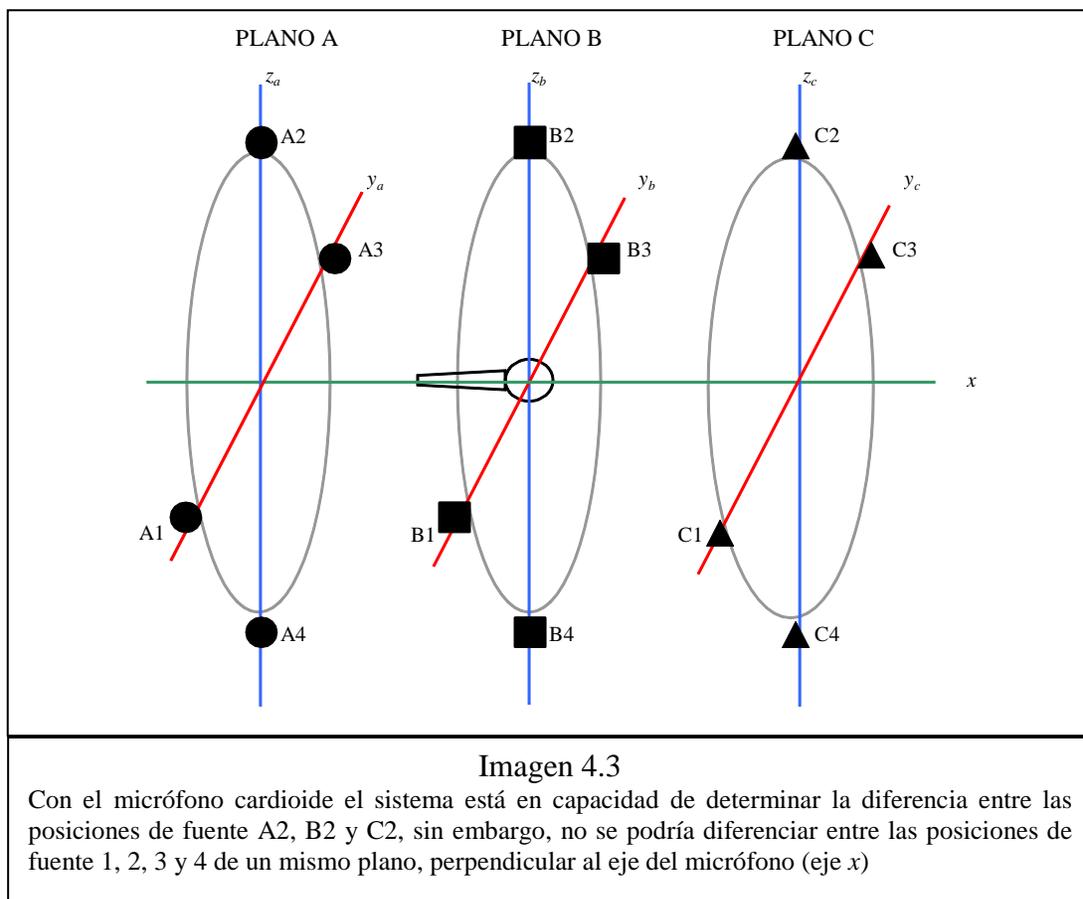
La limitación del sistema con el micrófono cardioide radica en la simetría que presenta el patrón polar, de modo que una fuente ubicada a 90 grados sobre el eje del micrófono (*eje x*), como en el caso del ejemplo anterior, podría situarse sobre un punto cualquiera de la circunferencia descrita sobre un plano perpendicular a dicho eje sin que el sistema detectara diferencia alguna.

En la imagen 4.3 se tiene un micrófono, que para este caso se supondrá de patrón polar cardioide, igualmente, se escogen tres planos perpendiculares al eje del micrófono (*eje x*), que se han denominado Plano A, Plano B y Plano C.

Sobre cada uno de estos planos se han señalado cuatro posiciones de fuente, indicadas como A1, A2, A3 y A4, y de manera similar para los otros dos planos. Con la imagen 4.3, se pretende explicar, que si se toman registros de audio para las cuatro posiciones de fuente de un mismo plano, sea el A, el B o el C, la variación de presión sonora sería nula, en virtud de la simetría del patrón de captación polar del micrófono. Por otra parte, si se toman registros de las posiciones A1, B1 y C1 considerando el micrófono estático en la posición del plano B, el sistema estaría en capacidad de discriminar entre estas tres posiciones, dependiendo de la distancia que exista entre dichos planos y la magnitud de la diferencia de presión entre los mismos.

También hay que considerar, que el patrón de captación, presenta un comportamiento diferente para cada frecuencia.

Dicho en otras palabras, el sistema provisto de un solo micrófono cardioide haciendo uso del criterio de función de transferencia, tendría capacidad de localización solamente en una dimensión o eje (*x*).



Habría que preguntarse entonces, ¿qué pasaría con el desempeño del sistema si se modificara la simetría del patrón polar de captación del micrófono?, pero, esto conllevaría a otra pregunta ¿Cómo podría cambiarse este patrón?

La respuesta a estos interrogantes es, la barrera acústica, de modo que, se adaptó una superficie al micrófono, separando la cara anterior y posterior de la cápsula del mismo, pero se debe tener en cuenta que la posición de la cápsula en la superficie no debe ser simétrica y la forma de la misma tiene que ser irregular.

Cumpliendo estas dos condiciones en la aplicación de la superficie, lo que se pretende es alterar las distancias que recorre el sonido proveniente de la cara posterior de la cápsula para ángulos y trayectorias de acercamiento diferentes.

Para las posiciones de fuente que estén del lado de la cara anterior de la cápsula se requeriría combinar otra superficie pero dispuesta de forma perpendicular a la primera. ¿Por qué?

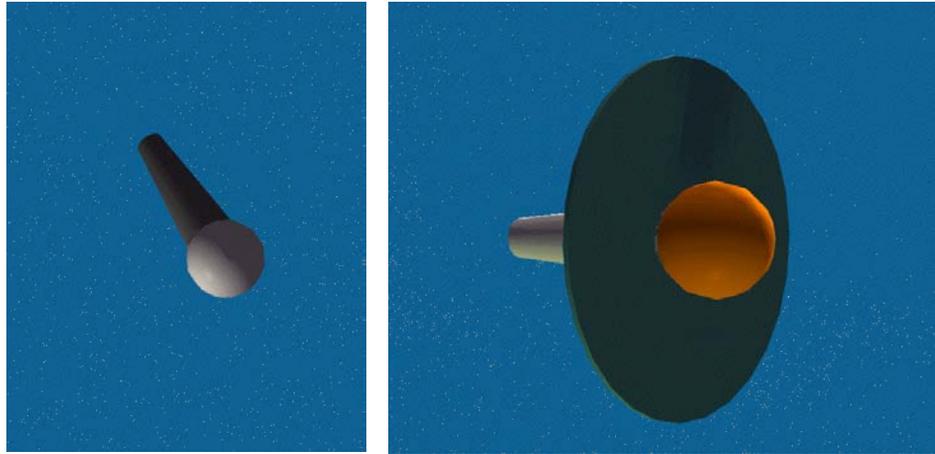


Imagen 4.4
Superficie simétrica

Retómese el caso del sistema provisto del micrófono omnidireccional que aunque inicialmente se afirmó que el micrófono tal como está diseñado no sería de mucha utilidad, en esta ocasión se le adaptará un elemento nuevo, una superficie dispuesta de modo tal que la cara anterior y posterior de la cápsula queden separadas, tal como se muestra en la imagen 4.4.

Con este recurso, lo que se busca es que para una distancia constante entre la fuente y la cápsula, los sonidos provenientes del lado de la cara posterior del micrófono deban recorrer una mayor distancia que los provenientes de la cara anterior, y en virtud de esta diferencia de distancias así como del efecto de sombra que proporciona la superficie, los sonidos de la parte posterior se registren con un nivel inferior a los provenientes de la región anterior. Esto obviamente, dependerá de dos condiciones, la primera, que se conserve la misma distancia fuente micrófono para cualquier posición de fuente y la segunda, que las dimensiones de la superficie sean comparablemente superiores a las de la longitud de onda del sonido incidente. Así, actuando dentro de un rango muy limitado de frecuencias, el sistema que inicialmente no servía para localizar fuentes, ahora cuenta con un accesorio que le aproxima al caso analizado con el micrófono cardioide y el nivel de presión sonora registrado por el sistema, será también una función dependiente del ángulo de incidencia del sonido, (*omítase por el momento lo referente al estudio de la fase para este caso*).

Hasta el momento solo se ha conseguido que el sistema tenga capacidad de localización limitada en frecuencia, en una sola dimensión, pero haciendo uso de un solo micrófono, sea omnidireccional o cardioide. El caso del micrófono omnidireccional con la superficie acoplada es particularmente importante, porque existe un elemento que hasta esta instancia no se había considerado, La Fase.

Cuando se separaron las caras anterior y posterior del micrófono, para generar diferencia de presión como consecuencia del enmascaramiento, también se estaba añadiendo el fenómeno de difracción al proceso, de modo que cuando el sonido proviene del lado opuesto de la cápsula, la señal envuelve la superficie por toda su periferia, haciendo que para posiciones específicas de fuente, se presenten tiempos de llegada diferentes ante una misma excitación acústica.

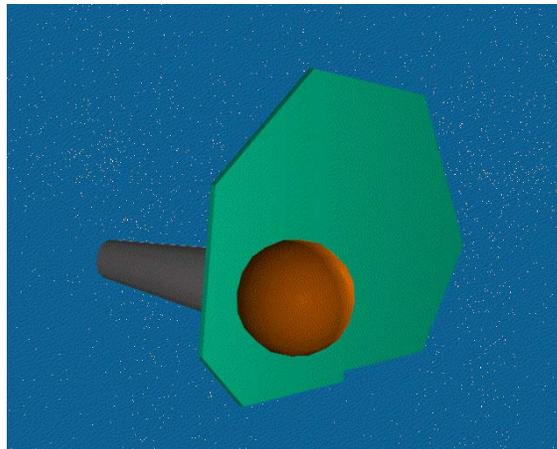


Imagen 4.5
Superficie asimétrica irregular

Como se muestra en la imagen 4.5, la onda sonora viaja por un número infinito de trayectorias generando retardos, algo que no es posible en el caso del sistema provisto del micrófono cardioide únicamente o con la superficie simétrica como en la imagen 4.6. De modo que, además de la diferencia de nivel de presión sonora que inicialmente se tomó como el criterio de comparación entre las señales A y B, ahora se cuenta con un nuevo elemento de estimación, que es bastante importante, si se tiene en cuenta que la mayoría de los autores que documentan el estudio de la localización sonora, lo hacen a partir de sistemas binaurales en los que se refieren a diferencias temporales entre los oídos izquierdo y derecho y en este caso, hasta el momento, solo se ha teniendo en cuenta solo un oído.

Considerando una situación con una unidad de captura como la ilustrada en la imagen 4.6, el sistema recae en las mismas limitaciones de simetría que en un principio se mencionó para el micrófono cardioide; si se sabe que ahora se está considerando la fase dentro del análisis, se podría sugerir una nueva modificación a la unidad de captura para mejorar la eficiencia del sistema.

De modo que se opta por una superficie con unas características especiales, la primera es que no sea una forma simétrica regular, como un círculo, cuadrado, o semejantes, y la segunda consideración es que la ubicación de la cápsula sobre esta superficie, debía ser en una zona No equidistante de los bordes de la misma.

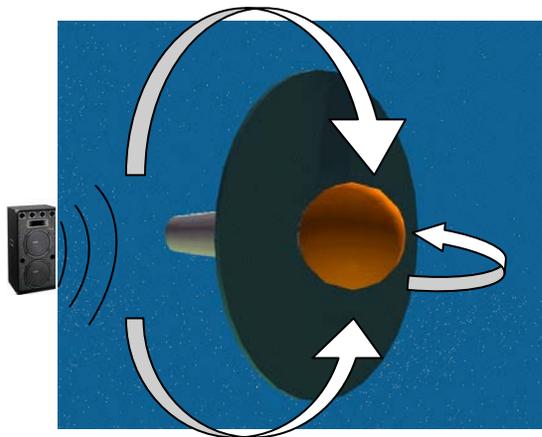


Imagen 4.6
Superficie simétrica regular

Todo este análisis se ha desarrollado sobre el caso del sistema con el micrófono cardioide y no se ha tenido en cuenta que ocurriría si se utilizara el micrófono de patrón omnidireccional. Inicialmente se afirmó que el micrófono tal como está diseñado no sería de mucha utilidad, sin embargo, no se planteó la posibilidad de adaptarle una superficie que, actuando como barrera acústica modifique el patrón de captación.

Si se acopla una superficie de modo tal que la cara anterior y posterior de la cápsula del micrófono queden separadas, todo sonido proveniente de la cara posterior tendrá que bordear necesariamente la superficie para ser registrada, cosa que no ocurrirá con los sonidos provenientes de la cara anterior. Igualmente, cuanto mayor sea el tamaño de la superficie con respecto a la longitud de onda del sonido incidente, mayor será el efecto de enmascaramiento, haciendo posible diferenciar por el momento, los sonidos provenientes del frente y de atrás.

Pero, ¿qué ventaja proporciona la nueva superficie añadida al sistema?

La geometría de la superficie empleada debe ser irregular, no simétrica, puesto que estas dos condiciones eliminan la posibilidad de que se presente una misma función de transferencia para más de una trayectoria de incidencia.

Gracias al fenómeno de difracción, el sonido proveniente de la cara posterior de la cápsula, bordea la superficie por un número infinito de trayectorias, de modo que si se la cápsula se ubica en una zona de la superficie cuya distancia de los bordes no sea equidistante para todos los ángulos de incidencia, se puede implementar un algoritmo que con base en la función de transferencia respectiva asigne un ángulo de trayectoria específico, para cada paquete de señales, cada una de las cuales evidencia un comportamiento de fase diferente, para un grupo de frecuencias específico. Dicho en palabras más cortas, cuanto más irregular sea la superficie, cada posición de fuente generará un paquete de información característico de cada punto. Si se emplea un algoritmo que esté en capacidad de integrar las n señales recibidas en un intervalo específico de tiempo, el criterio de localización será más sólido pues se tiene mayor cantidad de información y el margen de error disminuye, el ejemplo que sustenta esta afirmación es una situación cotidiana, si una persona escucha un sonido, de muy corta duración difícilmente podrá determinar su proveniencia, pero en la medida que la excitación acústica se hace más prolongada, el cerebro cuenta con un paquete de información más grande que le permite aumentar la certeza de la ubicación de la fuente.

La importancia de este nuevo planteamiento radica en que el sistema llega a tener capacidad de localización en dos dimensiones, aunque solo para los sonidos provenientes de la región posterior del micrófono, que son los afectados por la superficie.

Ahora, es pertinente aclarar las razones que impulsaron el anterior análisis. En primer lugar, se establece que el requerimiento de una cabeza binaural con forma craneana, provista de pabellones auditivos y un torso, para la unidad de captura del sistema tiene una justificación específica y es la siguiente:

El cerebro humano, está en capacidad de localizar fuentes sonoras en una dimensión, con ayuda de un solo oído, claro está, que en una medida muy inferior comparado con el caso de una persona provista de sus dos oídos, por ejemplo, la longitud de onda de la señal incidente debe ser comparable o menor que las dimensiones de la cabeza, luego estamos hablando de un rango de frecuencias superiores a los 1000 Herz, cuya longitud de onda es de aproximadamente 34 centímetros; el nivel de presión sonora del estímulo debe ser lo suficientemente elevado como para que el cerebro esté en capacidad de percibir con claridad la diferencia entre el sonido directo y el que ha bordeado la cabeza, pero a la vez cuidando que la presión no sea demasiado alta como

para que el sistema de protección auditivo restrinja la capacidad elástica del oído para recibir la información, además, la duración del estímulo debe ser lo suficientemente prologada, de modo que permita integrar la información suficiente para reducir el margen de incertidumbre.

El cerebro, aún estando provisto de un solo oído, al igual que el sistema con el micrófono de la imagen 4.5 integra las señales que lo atacan en diferentes momentos, producto de las n trayectorias diferentes que debe realizar la onda sonora proveniente de la región posterior al encontrar obstáculos de diferentes dimensiones en cada uno de sus n recorridos.

La razón que lleva a plantear una situación como la de la superficie de geometría irregular, obedece al hecho de querer ilustrar de una forma más simple lo que ocurre con la cabeza humana y los pabellones auditivos. La forma de los pabellones es irregular y la ubicación de la entrada al canal auditivo no es simétrica a los bordes del pabellón. También es preciso tener en cuenta que incluso, entre los oídos derecho e izquierdo de cualquier persona, no existe una simetría perfecta, por el contrario, en la mayoría de los casos, un pabellón está levemente más separado de la cabeza que otro, o, alguno es más grande que su opuesto, y lo interesante de este aspecto es que el cerebro diariamente se entrena y desarrolla funciones de transferencia específicas para cada necesidad, funciones de transferencia que se afinan mediante un sistema de retroalimentación y comprobación permanente en el que participan otros sentidos como la visión y el tacto.

Teniendo en cuenta lo anterior, se puede establecer que el sistema puede desarrollar funciones de transferencia para localización de fuentes sonoras en una dimensión, con ayuda de barreras acústicas, dispuestas de forma conveniente, tanto para micrófonos cardiodos como omnidireccionales.

Si esto se logra con solo un micrófono, se deduce que la capacidad de localización del sistema se incrementa cuando se incorpora el segundo micrófono, puesto que utilizando el mismo criterio de comparación espectral de potencia y de fase, la cantidad de datos aumenta, lo cual, es bueno para los propósitos del sistema, pero que a su vez exigirá una mayor capacidad de procesamiento al mismo.

Cuando la irregularidad de la superficie que enmascara la cápsula aumenta, la geometría del patrón de captación se aleja cada vez más de una forma simétrica, generando patrones de captación también irregulares

Con patrones de captación diferentes, se pueden implementar funciones de transferencia diferentes que se pueden asociar a ubicaciones específicas de fuente.

4.1.2 Evaluación del Sistema con dos micrófonos

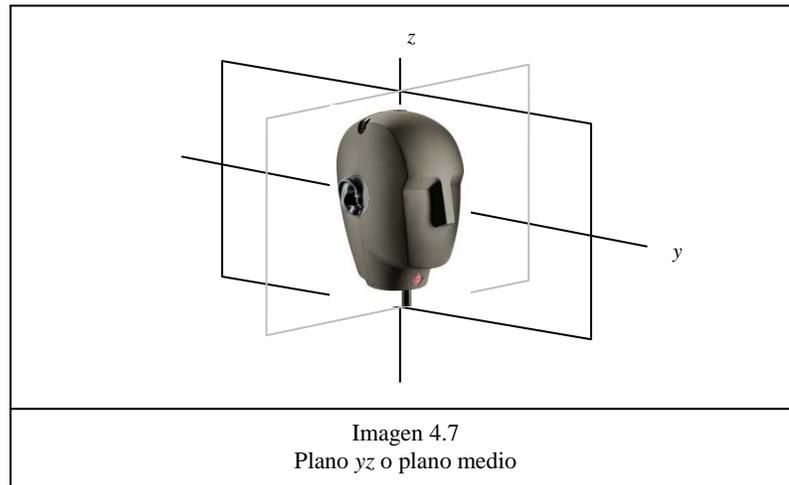
La capacidad de localización en el eje x es posible gracias a la diferencia en el tiempo de arribo de las señales a ambos oídos, sin embargo esto solo será válido para sonidos cuya longitud de onda sea comparable con las dimensiones de la cabeza porque para frecuencias muy bajas, la separación natural de los oídos no permitirá percibir con claridad una diferencia de tiempo, y para frecuencias muy altas, el criterio no será una diferencia temporal, sino, una diferencia de presión sonora entre los oídos. Esto es, en resumen, lo que ocurre con el sistema auditivo humano, y es lo que hace posible la localización de fuentes sonoras en el plano xy .

En el eje vertical, se puede aplicar el mismo criterio para la localización, sin embargo, existe una dificultad, cuando la fuente se encuentra en la región del plano medio que separa los oídos, o plano yz , en esta región, para una distancia cualquiera, la diferencia de tiempo será nula. De igual forma, por la similitud que existe entre los lados izquierdo y derecho de la cabeza, el efecto de sombra que ejercen los pabellones será muy parecido para ambos oídos, entonces, ¿cómo es posible que el cerebro discrimine con tanta facilidad el sonido que proviene de la nuca y lo diferencie con claridad del que está al frente?

Como se dijo anteriormente, aunque la morfología humana tienda a ser simétrica, esta simetría no es perfecta; tanto la superficie como la forma de los pabellones izquierdo y derecho es diferente, además, la ubicación de estos es relativamente dispareja, alguno se encuentra levemente más arriba y/o más adelante que su opuesto, además, las condiciones de elasticidad y sensibilidad también son levemente disparejas, lo que comúnmente suele decirse, es que, se oye más por un oído que por el otro.

Todo esto, encierra en si una gran ventaja y es que el cerebro, no realiza un dictamen sobre una diferencia entre dos señales puntuales, sino que, realiza una comparación de paquetes robustos de información provenientes de cada oído.

Cuando un sonido viene del frente, justo en el medio, aunque el sonido directo aparentemente llegue al mismo tiempo a ambos oídos, el cerebro integra la información que se genera por los n rebotes que se produce en el cuerpo y la cabeza con sus pabellones; si se deja a una persona estando de pie y se comienza a desplazar una fuente sonora a lo largo de una circunferencia sobre su eje yz (imagen 4.7), las condiciones del campo acústico que se presentan son diferentes para ángulos distintos de ubicación de la fuente, ya que los rebotes que se producen son diferentes para ángulos diferentes.



Los rebotes más significativos, son los que se generan en las zonas más próximas a los oídos, es decir, los hombros y el torso, el efecto que producen las piernas o el abdomen bajo podría considerarse prácticamente despreciable en los oídos.

Habiendo expuesto las anteriores consideraciones, se establece la necesidad de diseñar una unidad de captación binaural provista de pabellones auditivos y un torso, que permitan simular un campo acústico similar en las inmediaciones de los oídos, como lo experimenta el sistema auditivo humano para efectos de localización de fuentes sonoras.

La cabeza binaural que se propone debe tener en cuenta la implementación del tipo de micrófono disponible para la simulación (AKG C480B), es decir, se debe considerar su tamaño y forma para adecuarlos al interior de la cabeza.

4.2.1 Descripción General del Sistema

Luego de un proceso de diseño, construcción, medición, evaluación y corrección, se logra desarrollar un sistema de localización binaural, que hace uso del método de función de transferencia, comparación espectral y de potencia para llevar a cabo de forma artificial el proceso de localización de fuentes sonoras con un nivel de eficacia por

El sistema consta de las siguientes partes:

Partes del Sistema de Localización Binaural.

4.2.1.1 Unidad de Captura.

Cabeza Binaural, construida en fibra de vidrio y masilla, con un relleno de lana de vidrio de media densidad. La cabeza que en su forma se aproxima a la forma humana, cuenta con compartimentos internos para alojar los micrófonos (AKG C480B), sobre una superficie de madera tipo mdf recubierto con espuma negra para brindar un soporte firme pero a la vez disminuir la transmisión por vía solida a los micrófonos desde el soporte.

En la imagen 4.8 se pueden observar los pabellones, así como la parte interna de la cabeza que cuenta con un cierre magnético, para evitar el uso de broches o tornillos.

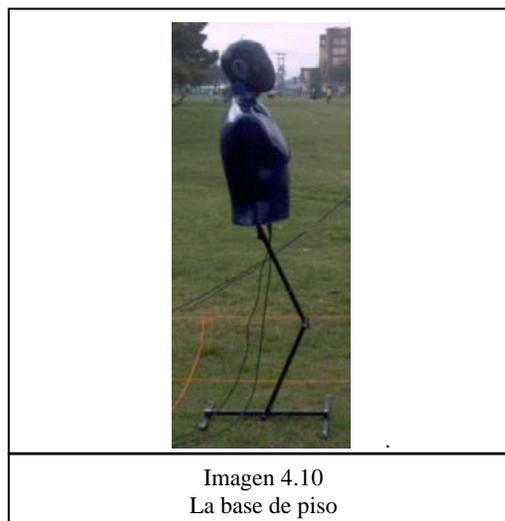


Imagen 4.8
Cabeza Binaural con pabellones

El torso, está también construido en fibra de vidrio y masilla, con un relleno de lana de vidrio de media densidad al igual que la cabeza, con el fin de reducir convenientemente las resonancias propias de esta estructura. En su interior, el torso cuenta con una estructura en madera tipo mdf de 15 mm, que tiene como función mantener unidas las caras anterior y posterior del torso, al mismo tiempo que brinda soporte a la cabeza y permite un acople seguro con la base de piso. En la imagen 4.9 se puede apreciar el torso sosteniendo la cabeza.



La base de piso, es una estructura articulada, construida en hierro común de (tubo cuadrado de 2cm de lado), que tiene por objeto soportar al torso, que a su vez sostiene a la cabeza. Ha sido diseñada teniendo en cuenta las diferentes posiciones que debe adoptar para simular la ubicación y postura de una persona promedio en diferentes condiciones y/o lugares de interés, como salas de grabación, auditorios, espacios abiertos, entre otros. Cuenta con 4 apoyos autorroscantes que permiten ajustar la base a las irregularidades del suelo y a su vez hace posible aislar considerablemente la estructura de las vibraciones del piso, tal como se puede observar en la imagen 4.10.



4.2.1.2 La interfaz Análoga Digital / Digital Análoga

Tarjeta de Sonido. En este caso, se cuenta con una tarjeta de sonido externa marca M-Audio, referencia MobilePre, que se conecta por puerto USB al computador. La tarjeta está diseñada para trabajar capturando un máximo de dos señales simultáneamente, balanceadas o desbalanceadas, a una tasa de muestreo entre 8 KHz y 48 KHz, con una respuesta en frecuencia de 20 Hz a 20KHz, rango dinámico de 88 dBA, relación señal a ruido S/N: -88 dBA, y una distorsión armónica THD: 0.013% (-78dB) -1dBFS a 1 KHz.



Imagen 4.11
Tarjeta de sonido USB (Interfaz AD y DA)

4.2.1.3 La Unidad de Procesamiento

El Computador, cuenta con una configuración específica de hardware y software, para adelantar los procedimientos lógicos que exige la aplicación de localización.

Requerimientos de Hardware:

- Procesador Pentium III de 800 Mhz, o superior. En su defecto, un procesador de similar rendimiento de cualquier otro fabricante.
- Memoria RAM de 800 MB como mínimo.
- Tarjeta madre (Main Board) con un bus de 400 MHz o superior, con disponibilidad de al menos un puerto USB (Universal Serial Bus) de 2.0.
- Espacio en disco duro de 10 GB. Menor capacidad podría llegar a generar conflictos de rendimiento de la máquina cuando los archivos de audio sean lo suficientemente numerosos y extensos.
- Unidad de CD ROM o DVD ROM
- Monitor de video estándar, de resolución 800 x 600, o superior.

Requerimientos de Software

- Windows XP SP2. No se recomienda trabajar con Windows Vista, puesto que se sabe que este sistema operativo destina una cantidad importante de recursos de hardware para efectos gráficos, en perjuicio del rendimiento del equipo.
- Matlab 7.0 y Simulink 6.0. o Matlab 2008. Versiones anteriores de Matlab y Simulink pueden funcionar, no obstante, la aplicación solo se ha probado en las versiones antes mencionadas.
- Software de instalación para la tarjeta Mobile Pre USB.

NOTA: En algunos equipos se requerirá actualizar el BLAS⁵ para poder ejecutar Matlab sin inconvenientes como ocurre con algunos procesadores como el AMD Turion 64x.

Otros requerimientos adicionales:

- Parlante con amplificador, según sea el caso del parlante empleado. Debe tener una respuesta en frecuencia estable entre las bandas de 250 Hz y 8 KHz.
- Se requerirá sonómetro para la calibración del sistema, en tanto no esté determinada una referencia previamente reconocida y configurada por el sistema de localización.
- Se puede considerar el uso de un juego de audífonos de alta definición con respuesta plana entre 20Hz y 20 KHz, para complementar el procedimiento de calibración.

4.2.2 La Aplicación de Software y consideraciones generales de su diseño

La aplicación para localización de fuentes sonoras USB-RADAR, recibe su nombre, en primer lugar, como reconocimiento a la Universidad de San Buenaventura Bogotá, que es la instancia académica que acogió, a través de su nodo de investigación el presente proyecto. El término Radar se atribuye por analogía y similitud entre la función que pretende cumplir la Aplicación de Software y el radar empleado para navegación en aeronáutica, cuyo fin ya es bien conocido.

La investigación realizada, ha permitido profundizar en los diferentes recursos que ofrece matlab para la elaboración de ventanas de diálogo, botones, menús desplegables, barras de desplazamiento, casillas de verificación y otros más, que

⁵ BLAS (Basic Linear Algebra Subroutines), es un conjunto de subrutinas en lenguaje de máquina que emplea el procesador para desarrollar procedimientos algebraicos específicos, como la multiplicación de vectores y matrices o la solución de sistemas de ecuaciones lineales.

hacen parte de la herramienta para diseño de entorno gráfico conocida como GUI⁶.

Igualmente, en desarrollo de las actividades de investigación, se detectó la necesidad de emplear el Simulink. Esta es una poderosa herramienta que brinda matlab para hacer montajes virtuales de entornos físicos reales que sean objeto de investigación.

Simulink permitió diseñar y modelar diferentes subsistemas para la reproducción y captura simultanea de señales de audio con características especiales. La versatilidad de esta herramienta no solo simplificó la carga operativa en lo que se refiere a la generación de código de programación necesario para procesar las señales, sino que además permitió hacer cambios y ajustes de importancia en tiempos muy cortos.

Inicialmente, se consideró la posibilidad de compilar la aplicación de localización como un archivo ejecutable (.EXE), para que fuera posible implementarla incluso en equipos que no tuviesen instalado el software sobre el cual se diseñó, es decir matlab. No obstante y a pesar de el beneficio que esta compilación representaría, se desistió de llegar a este punto, por lo menos en la presente investigación, al considerar que la generación de archivos ejecutables desde matlab para una aplicación de las dimensiones y complejidad que tiene USB-Radar, merece un tratamiento más profundo y detallado, extendiéndose más allá del objetivo de la presente investigación, además de considerar la limitación que existe al contar con una licencia de software académica y no profesional que es la requerida para tal fin.

4.2.3 Arquitectura de la Aplicación para Localización de Fuentes Sonoras

La gran mayoría de sistemas desarrollados en ingeniería y otros campos de investigación, parten de una etapa de observación y posterior imitación de otros sistemas ya existentes. USB-Radar no fue la excepción y haciendo diferentes análisis se determinó que la aplicación debía ensamblarse por módulos y funcionar por etapas. El sistema de ensamble por módulos permite trabajar paquetes que ejecuten tareas específicas sin perjuicio de alterar otros módulos en dado caso de presentarse errores de programación. Igualmente, es posible flexibilizar la aplicación alterando, según convenga, el orden en que se invoquen los módulos; dicho en pocas palabras, es un sistema de diseño que permite alterar el orden en que se ejecuten las instrucciones reduciendo el riesgo de dañar la aplicación global de forma accidental.

⁶ Graphic Unit Interface (GUI)

4.2.3.1 Etapas de la Aplicación para Localización de Fuentes Sonoras

Se diseñaron cuatro etapas a saber que son: **Calibración, Captura, Procesamiento y Análisis**. A continuación se exponen cada una de estas de manera más detallada.

La Calibración: En esta primera parte se ajusta el nivel de sensibilidad del sistema. Se requiere que los micrófonos instalados en la cabeza binaural presenten las condiciones de simetría más favorables posibles. Incluso cuando se cuenta con dos unidades de captación idénticas, como es el caso de este sistema, la sensibilidad de los micrófonos no era la misma para la banda de 2 y 4 KHz, por lo tanto, fue necesario implementar herramientas de ecualización.

Aunque matlab cuenta con recursos para hacer la ecualización de los micrófonos en una etapa previa al almacenamiento de las señales, se optó por utilizar ecualizadores externos para no recargar los recursos de memoria y procesador del equipo.

También es preciso asegurarse de tener un nivel de presión apropiado en la unidad de radiación o parlante. Un nivel de presión muy bajo dejaría las señales muy cerca del piso de ruido y un nivel muy alto puede saturar las señales de entrada de los micrófonos. Teniendo en cuenta que si se carece de una cámara anecoica para adelantar las mediciones, la presencia cercana de superficies reflectantes generan rebotes no deseados que alteran de forma importante la capacidad de localización; esta es otra razón para no emplear niveles muy elevados en la reproducción de los sonidos.

No está de más, incluir dentro del proceso de calibración, una rutina que verifique la respuesta en frecuencia de los micrófonos, ya que si estos elementos no son de uso exclusivo del investigador, es posible que alguno de sus componentes haya sufrido algún daño o alteración que deba ser tenida en cuenta antes de iniciar el ejercicio de localización.

La Captura: Esta etapa requiere que desde el simulink se ejecuten dos procesos: la reproducción de una señal de audio por un canal monofónico y la captura de dos señales (L y R) que son almacenadas de manera independiente en una matriz, todo se ejecuta simultáneamente.

Las señales que se pueden reproducir son: Ruido rosa y tonos puros de 250Hz, 500Hz, 1 KHz, 2 KHz, 4 KHz y 8 KHz., sin embargo, por razones prácticas que atañen al tiempo que puede llevar el análisis, en la mayoría de los casos el ejercicio para los tonos se limitó a la frecuencia de 1Khz. Las seis bandas escogidas están ubicadas dentro del espectro de mayor sensibilidad en el ser humano. La razón para no considerar otras frecuencias intermedias

obedece a que el sistema de audición humano reconoce el incremento de frecuencias en distancias tonales de octava con mayor facilidad y precisión que los incrementos por décadas. La tasa de muestreo de las señales almacenadas es de 44100 a una resolución de 16 bits.

El Procesamiento: En esta instancia, es donde se debe elegir el criterio con que se han de tratar las señales capturadas, es decir, se debe tener un Criterio de Localización (LC) Location Criteria, el cual debe estar apoyado en un fundamento teórico claro y comprobable. No implica que dicho criterio, tenga que ser en efecto, un método 100% efectivo e irrefutable, pero que por lo menos, obedezca a un planteamiento físico y acústico razonable.

Este criterio es la brújula que orienta la construcción de los algoritmos que procesan los archivos de audio almacenados. Por ejemplo, si se emplea el criterio de diferencia de nivel de presión inter auricular (ILD), no se justificaría hacer un procedimiento de correlación, porque bastaría con calcular la potencia promedio de las señales L y R en un mismo intervalo de tiempo para saber qué oído es el que recibió un mayor nivel de presión.

Criterios de Localización:

En el desarrollo de los algoritmos de localización se implementaron tres criterios a saber:

Criterio de Diferencia Inter auricular de Nivel (ILD Criteria): Basado el fenómeno ampliamente documentado y reconocido que consiste en la capacidad que tiene el cerebro humano de establecer que oído recibe una mayor cantidad de energía acústica en un mismo intervalo de tiempo.

Para este fin, se implementaron las siguientes instrucciones:

$$\begin{array}{l}
 A \quad \left\{ \begin{array}{l} Y1=yL.^2; \\ Y1=(\text{sum}(y1(n1:nn)))/(nn - n1); \end{array} \right\} \\
 B \quad \left\{ \begin{array}{l} Y2=yR.^2; \\ Y2=(\text{sum}(y2(n1:nn)))/(nn - n1); \end{array} \right\}
 \end{array}$$

Obsérvese que la sección A y la B corresponden al cálculo de potencia de los canales izquierdo (L) y derecho (R) respectivamente. Posteriormente se emplea alguna rutina de comparación para determinar el canal que ha registrado el nivel más alto. Para ejemplificar la situación se ha creado un vector de dos posiciones nombrado *lateralidad*.

```

Comparar(1,1)=Y1
Comparar(1,2)=Y2

```

Lateralidad= max(Comparar)

En el anterior ejemplo se empleó también un vector denominado “comparar” que simplemente es un recurso para mostrar como los valores calculados en las secciones A y B se almacenan en un nuevo arreglo, el cual es sometido a la función “max” que se encarga de encontrar el mayor valor, indicando con ello el hemisferio de la cabeza que fue expuesto a un mayor nivel de presión acústica.

Criterio de Diferencia Inter aural de Tiempo (ITD Criteria): Se basa en la habilidad que tiene el cerebro de percibir las diferencias en el tiempo de arribo de un sonido ante su variación de posición en el plano horizontal.

Este criterio establece que la percepción de diferencia temporal será más notoria cuanto menor sea la frecuencia a registrar y en la medida que la frecuencia de estudio aumenta, la posibilidad de percibir una diferencia temporal será cada vez menor. Por lo anterior es preciso establecer una frecuencia de corte o mejor aún, un rango de frecuencias a las que se les pudiese asignar un algoritmo para determinar diferencia en tiempo de arribo.

Los criterios ILD e ITD se complementan, para un rango de frecuencias mayores a una frecuencia de corte específica, el ILD indicará una ubicación lateral gracias a la dificultad que tienen las altas frecuencias para sobrepasar obstáculos a costa de una pérdida importante de su energía, por su parte, las frecuencias inferiores a la de corte, no presentarán una pérdida energética tan notoria como su diferencia temporal al llegar a los oídos en momentos diferentes.

De las 6 bandas escogidas para este modelo de simulación, se establece aplicar el criterio ITD a las tres primeras, es decir 250 Hz, 500 Hz y 1KHz, en tanto que el criterio ILD se aplica a las tres últimas que son 2 KHz, 4 KHz y 8 KHz.

Criterio de Convulsión / Correlación: Una vez considerados los criterios anteriores, surge una inquietud al considerar las situaciones en las que la fuente sonora se ubica en cualquier punto del plano medio que divide la cabeza, justo en medio de los oídos. Para estas posiciones de fuente las diferencias tanto en nivel de presión como en tiempo de arribo serán prácticamente nulas, por lo tanto, se presume que debe haber un tercer elemento que juega un papel importante en el proceso de localización.

Este criterio se presenta como la propuesta de la presente investigación, para tratar de responder a la pregunta que surge ante la presencia de

fuentes sonoras en el plano medio de la cabeza. Su planteamiento es el siguiente:

Si las diferencias de nivel y de tiempo no son perceptibles, el cerebro utiliza un recurso que bien puede estar presente incluso con un solo oído ya que el efecto simétrico es indiferente a su análisis. Pero, ¿Qué tipo de información puede sustraer el cerebro de un oído para determinar una posición en el plano medio?

Considerando la fisiología del oído humano, el cerebro procesa paquetes de información que integran relaciones de fase y energía entre diferentes frecuencias, producto no solo de los diferentes caminos que recorre el sonido al bordear la cabeza y el cuerpo, sino al generarse una cantidad indefinida de rebotes en las proximidades de los oídos, más concretamente por culpa de los hombros, la cara y desde luego los pabellones auditivos, proporcionando lo que en términos menos científicos se conoce como coloración del sonido, cualidad que puede asociarse con el timbre del mismo.

Dicho de otro modo, la coloración del sonido que el cerebro percibe cambia a lo largo del plano medio, debido a la variación de las superficies en las que rebota el sonido al aproximarse desde diferentes puntos.

La pregunta que surge ahora es: ¿Cómo percibir esta diferencia de coloración en una unidad de medida cuantificable? Para responder esta pregunta fue necesario introducir un nuevo concepto, también asociado con una función cerebral y casualmente mencionado en el análisis de señales aunque de forma diferente, se trata de la *memoria*.

Se hace referencia al término de memoria pero no como se entiende en el contexto del análisis de sistemas y señales, sino en el marco de las funciones cerebrales humanas.

El ser humano, desde su infancia está entrenando constantemente cada uno de sus sentidos, este es un proceso de reconocimiento permanente; es así como por ejemplo, asociamos radiaciones de cierta frecuencia con una determinada sensación y a su vez con un nombre específico, lo que después llamamos color. De igual modo, cada coloración específica de un sonido se asocia con una determinada sensación y una posición en el espacio, esto nos lleva a darle un nombre de procedencia, bien sea arriba, abajo, adelante o atrás, generando una especie de mapa espacial de coloraciones.

Para el presente criterio se establece que en la memoria del ser humano se almacena este hipotético mapa espacial y cada vez que se recibe un

estímulo acústico, la comparación entre lo que se escucha y lo que está memorizado permite dar un dictamen sobre la ubicación de la fuente en el plano medio.

Teniendo en cuenta el anterior razonamiento se puede ver que el propósito del algoritmo a utilizar, es hacer una comparación entre dos valores, el primero es el registro que se obtuvo en la etapa de calibración, que fue almacenado en una matriz y que está claramente identificado con una posición en el espacio, el segundo valor es el de la señal capturada a la cual se le quiere determinar la posición de fuente, de allí la necesidad de emplear la función de correlación.

Ahora ante la pregunta, ¿por qué emplear la Convolución?, hay que decir que:

Si se presume que al cambiar la dirección de aproximación del sonido, el contenido espectral de la señal cambia también, debido al filtrado que ocasionan las superficies irregulares de la cabeza; se puede inferir que para cada posición se debe presentar un tipo de filtrado diferente a los demás, luego si se conoce la señal original (señal de calibración) y se identifica el filtrado obtenido en la señal a localizar, se puede deducir por este medio la posición asociada al tipo de filtrado registrado.

Aunque en la teoría este planteamiento parezca sencillo, en la práctica es bastante complicado identificar diferencias importantes en el contenido espectral, más aún cuando el tamaño de los obstáculos (pabellones auditivos) es relativamente pequeño; además, también es difícil ver que estas diferencias se repitan de manera consistente. Por esta razón, se determinó el uso de una rutina que involucrará la combinación de los dos oídos para la localización en el plano medio.

El objetivo es que se pueda tomar cada par de señales LR, de cada posición y sean tan fácilmente identificables como si se tratara de una huella dactilar para cada posición del espacio. Para ello, se asumió que la cabeza binaural trabajaría como un sistema LTI con dos entradas independientes L y R, pero que al ser registradas se convolucionan como si una fuese la respuesta al impulso $h(t)$ de la otra. Considerando la propiedad de conmutatividad de la convolución, no hace falta especificar cual señal es $h(t)$. De modo que se entendería como un sistema en el cual para cada posición del espacio habría una única respuesta al impulso, o si se quiere, la función $h(t)$ sería una variable dependiente de la localización de la fuente, es decir $h(x,y,z)$. Este último planteamiento se ajusta de una mejor forma a la realidad del modelo, si se tiene en cuenta que para el caso de la localización sonora, el filtrado asociado a una posición no es

dependiente del momento en que se produzca el sonido sino del sitio de la fuente.

La pertinencia de este procedimiento se funda en el hecho de que la convolución es una función que integra el contenido de las señales en su totalidad, reduciendo la ambigüedad que se presenta al trabajar únicamente en el dominio de tiempo.

El Análisis

La posibilidad de que los espectros de dos posiciones adyacentes sean parecidos existe, pero la probabilidad de que sus señales convolucionadas sea igualmente parecida debe ser menor, además, en el análisis empleando la convolución debe generar una sola gráfica para una misma posición de fuente, sin importar la señal que se utilice, por ejemplo, para una fuente ubicada frente al oyente a la altura de la cara, la gráfica de la convolución debe ser igual cuando escucha un tono que cuando escucha una voz humana, la razón es que sin importar el tipo de señal registrada, las relaciones de fase y nivel bajo condiciones ideales se conservan intactas.

Retomando, lo que hacen los algoritmos es comparar las convoluciones entre los canales L y R de cada posición, que fueron asociadas en la etapa de calibración a una ubicación específica, con la convolución de las señales L y R del registro que se quiere localizar.

4.2.3.2 Los Módulos de La Aplicación

Módulos de Inicio: Son ventanas creadas para introducir al usuario en la aplicación y a su vez darle las instrucciones iniciales.

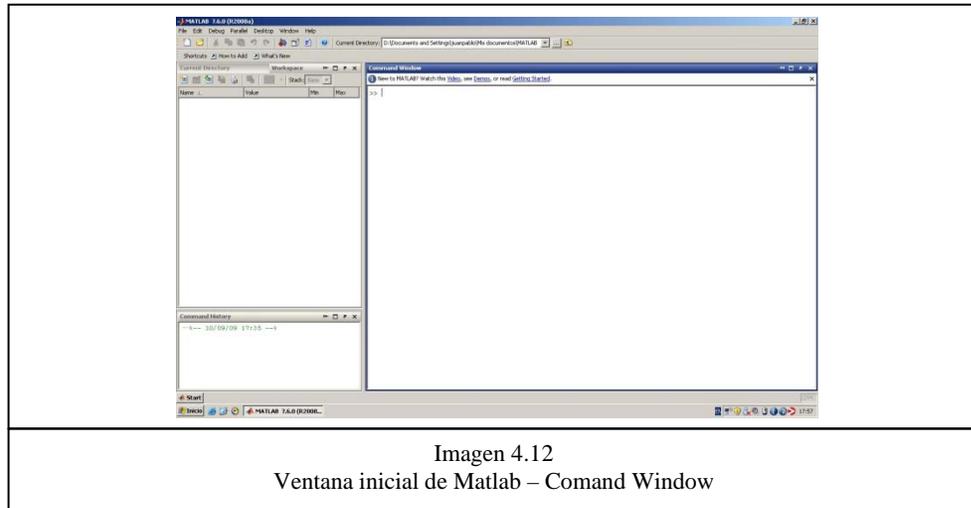
Módulos de Calibración: Son ventanas diseñadas para solicitarle al usuario la información sobre el tipo de calibración que desea realizar. Hay que mencionar que la aplicación le permite al usuario hacer una calibración para cualquiera de las seis bandas disponibles.

Módulos de Localización: Son ventanas diseñadas para realizar las operaciones matemáticas y los procesos que involucran todo el análisis de las señales, así como la gráfica de la ubicación de la fuente.

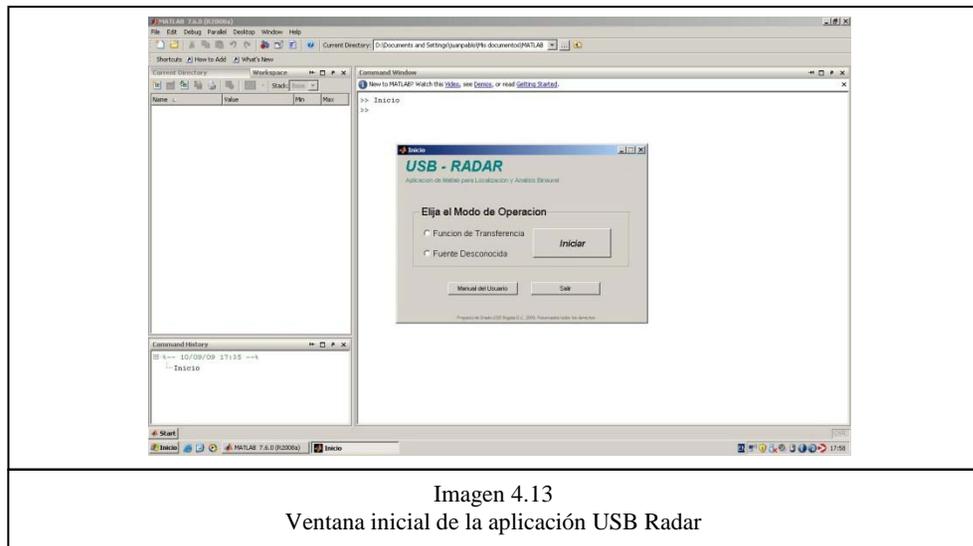
Módulos de Verificación: Son ventanas que advierten al usuario sobre alguna situación especial, como por ejemplo si desea cerrar la aplicación, o si quiere continuar con el siguiente punto de calibración. También en algunos casos cumplen funciones de diagnóstico que permiten detectar errores o comportamientos especiales.

4.2.4 Descripción de la Aplicación paso a paso.

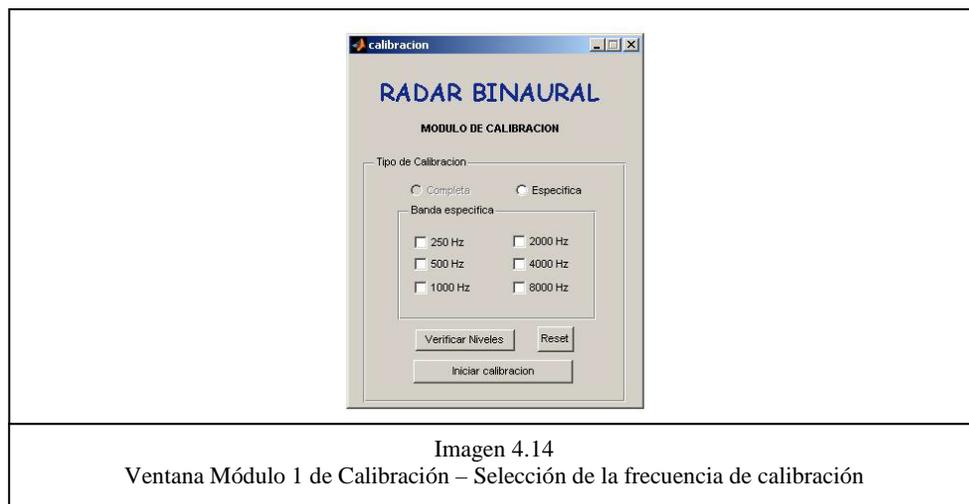
Como ya se explicó anteriormente, la aplicación no tiene en esta instancia del proyecto un archivo ejecutable, por lo tanto, para iniciarla es preciso ingresar a matlab y en la ventana del Comand Window digitar “Inicio” como se puede ver en las imágenes 4.12 y 4.13 respectivamente.



Matlab mostrará (Imagen 4.13) una ventana de inicio en la que se pregunta al usuario el modo bajo el cual desea trabajar, “Fuente Desconocida” o “Función de Transferencia”. El primero de estos se encuentra solamente indicado, en tanto que el segundo esta activo y se puede seleccionar y luego se da click en el botón “Iniciar”. También Están las opciones que ofrecen los botones “Manual del usuario” el cual abrirá el documento correspondiente. El botón “Salir” como lo indica, cerrará la aplicación.



Luego de dar inicio en la ventana de la imagen 4.13, se abrirá el módulo de calibración que se muestra en la imagen 4.14. En esta ventana se le pregunta al usuario el tipo de calibración que se desea hacer al sistema. La opción “Completa” es una calibración que tomará registro de todas las frecuencias para cada una de las 25 posiciones establecidas, en tanto que al elegir la calibración específica el sistema de solicitará al usuario escoger la banda que se quiere registrar. En la parte media de la ventana aparecen las opciones para cada una de las frecuencias posibles. En la parte inferior se encuentra un botón rotulado con “Verificar niveles” que al ser pulsado abrirá el control de niveles de la tarjeta de sonido instalada en el sistema operativo. El botón “Reset” está para eliminar los valores que previamente hayan podido quedar almacenados en una calibración anterior y por último, el botón “Iniciar calibración” conducirá al siguiente paso.

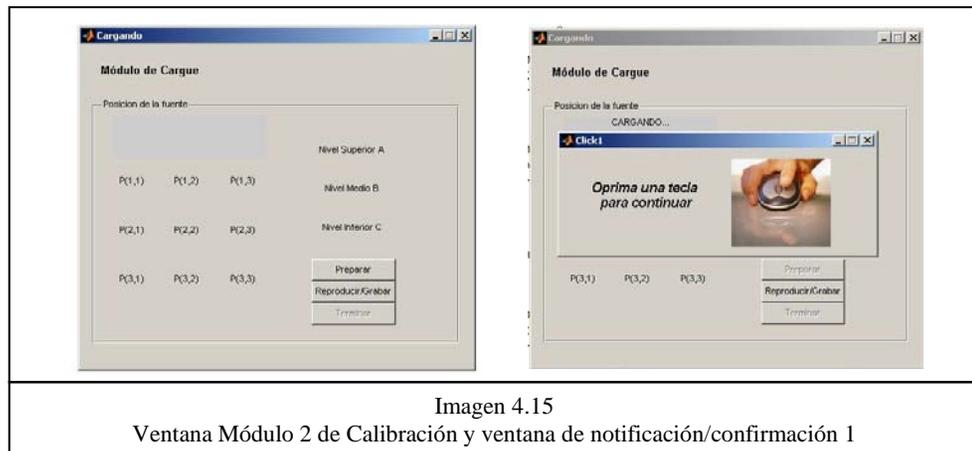


La etapa de Calibración se desarrolla en dos módulos que trabajan encadenadamente, es así como la selección hecha en la ventana de la imagen 4.14 se implementará en la ventana de la imagen 4.15, en donde se encuentran tres botones y una grilla que representa las posiciones de fuente vistas desde un punto elevado. La altura a la que se encuentra la fuente estará indicada por los avisos de texto ubicados al costado derecho de la ventana y cada vez que haya un cambio de nivel de los tres posibles, el texto correspondiente se resaltará visiblemente.

El paso a seguir es pulsar el botón rotulado como “Preparar”, en ese momento el programa pedirá que se verifique la posición correcta de la fuente y de manera no implícita se presume que el operario debe verificar que las condiciones de ruido sean favorables para evitar intromisiones en el momento de iniciar la simulación.

Una vez cumplido este requisito, se da click en el botón “Reproducir / Grabar”, con el cual se pondrán en marcha los módulos del simulink encargados de reproducir y grabar simultáneamente, a través de la tarjeta de sonido las señales de tono y ruido que se hayan seleccionado en la ventana de la imagen 4.14.

Inmediatamente se inicie este proceso, el texto que indica la primera posición de fuente “P(1,1)” cambiará de color, señalando con esto la ubicación que en ese momento debe tener el parlante. También estará resaltado el texto que dice “Nivel Superior A”.⁷



Por cada posición de fuente, el sistema reproducirá dos tipos de señales, primero el tono de la frecuencia seleccionada y posteriormente una señal de ruido rosa, cada uno tendrá una duración no superior a 5 segundos⁸. Para evitar que el programa caiga en errores de procesamiento y por alguna razón se llegara a reproducir una señal antes de tiempo, tan pronto como finalice la reproducción del tono aparecerá una ventana de aviso que solicitará la confirmación para reproducir el ruido rosa. Esta misma solicitud o petición para continuar aparecerá después de generarse el ruido rosa, dando la posibilidad de cambiar la fuente a la siguiente posición, sin perjuicio de activar la simulación antes de tiempo.

Cuando las nueve posiciones del primer nivel hayan sido registradas, una nueva ventana de aviso y petición de confirmación aparecerá, indicando que hay que modificar la altura de la fuente antes de reactivar el simulador. (Imagen 4.16)

Una vez que se han terminado de registrar las señales de la última posición del último nivel (Nivel C, P(3,3)), se activará el tercer botón de esta ventana de

⁷ Ver la sección correspondiente a la rutina de calibración

⁸ Un tiempo de muestreo superior hace considerablemente más demorado el procesamiento de las señales.

calibración, rotulado con “Terminar”, indicando que el registro está completo y se puede pasar a la siguiente etapa, denominada como Ventana de Localización o Finder (Figura 4.17).

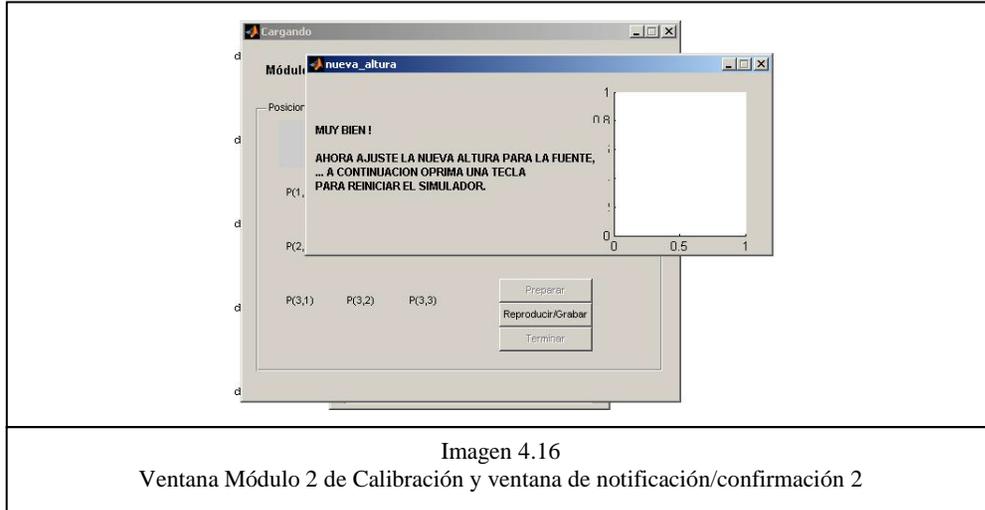


Imagen 4.16
Ventana Módulo 2 de Calibración y ventana de notificación/confirmación 2

El Módulo de Localización involucra todos los procesos realizados en las etapas anteriores. El propósito de esta ventana es recrear de una forma secuencial y clara, el proceso de localización que lleva a cabo el sistema auditivo humano, teniendo en cuenta las consideraciones y planteamientos teóricos del criterio de localización escogido.

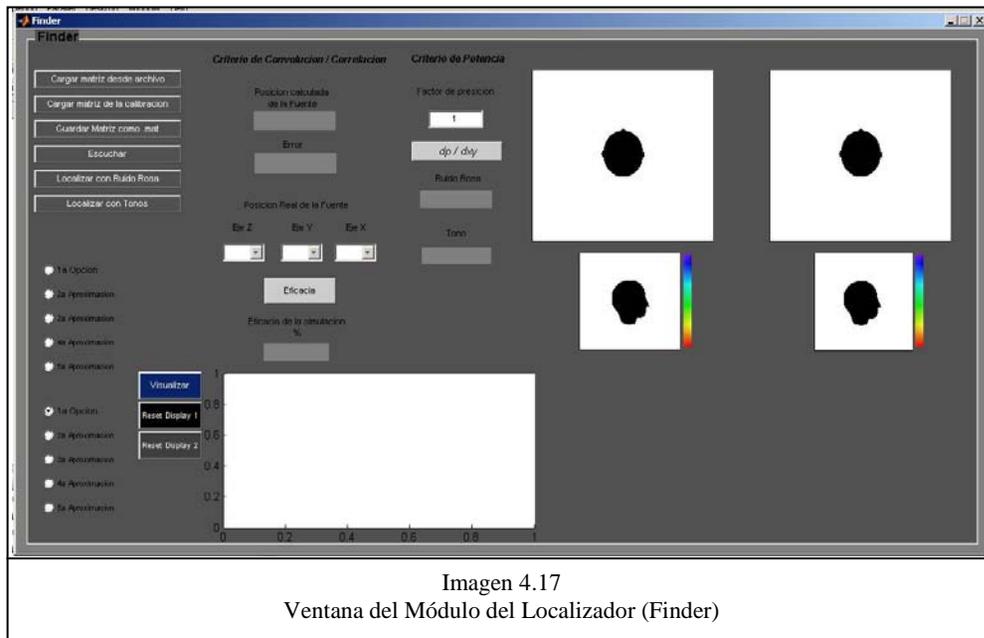


Imagen 4.17
Ventana del Módulo del Localizador (Finder)

Para comprender mejor la estructura y el funcionamiento del Localizador es preciso indicar las etapas que se propusieron en su diseño.

El ser humano realiza el proceso de localización en tres pasos, que son: escuchar, procesar y reconocer o dictaminar la procedencia del sonido. De igual modo, el Localizador cuenta con un botón que activa la reproducción y grabación de una señal (“Escuchar”), dos botones de procesamiento (“Localizar con Ruido Rosa” y “Localizar con Tonos”) y un botón para mostrar la posición después de ser reconocida (“Visualizar”). Se sobreentiende que debido a su propósito, el orden en que se oprimen estos tres botones no admite cambios.

Como el criterio seleccionado involucra el concepto de memoria, se tiene la necesidad de invocar los datos cargados en la etapa de calibración, estos datos, son la memoria del sistema, un proceso que en un ser humano lleva toda la vida, en este caso se tuvo que implementar en aproximadamente 30 minutos. Es un paso necesario para poder tener funciones de comparación.

La forma de invocar estos datos es presionando el botón “Cargar Matriz de la calibración”, a lo cual el programa responde con una ventana de información que indica que se están procesando los datos (Imagen 4.18). El tiempo que tarde dependerá del tamaño de las muestras tomadas, el cual se advirtió anteriormente que no debía superar los 5 segundos.

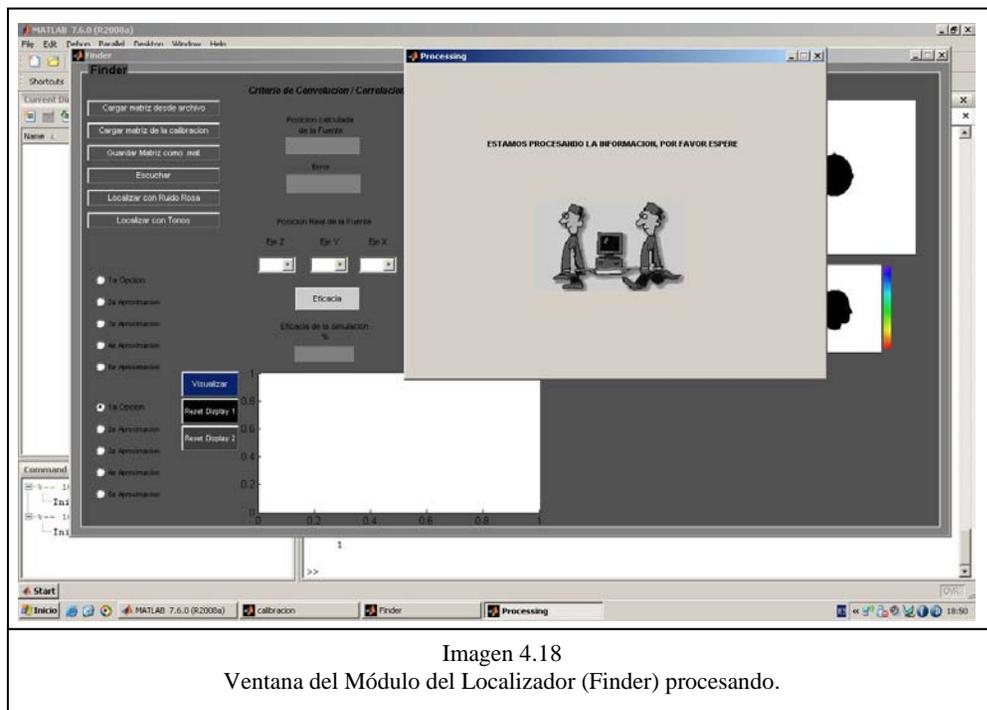


Imagen 4.18
Ventana del Módulo del Localizador (Finder) procesando.

Una vez desaparece la ventana de información, se puede iniciar la simulación

El proceso es bastante simple, solo hay que elegir una posición arbitraria para la fuente y ubicar allí el parlante; posteriormente se presiona el botón “Escuchar” con lo que se reproducirán las señales de tono y ruido ya escuchadas en la calibración. A continuación se presiona el botón “localizar con Ruido Rosa” y en este punto el programa ejecutará los algoritmos de comparación para hallar dentro de las 25 posiciones de fuente almacenadas en la matriz, a aquella cuya función de transferencia sea más próxima a la de la señal que se acabó de registrar.

El botón “Localizar con Tonos”, repite el mismo proceso de comparación pero empleando las señales del tono escogido, como su nombre lo indica.

El botón “Visualizar” funciona de manera combinada con el panel de selección única que se encuentra al costado izquierdo de la ventana del Localizador y con los cuatro indicadores bidimensionales de posición de la fuente que se pueden ver en la esquina superior derecha del mismo módulo.

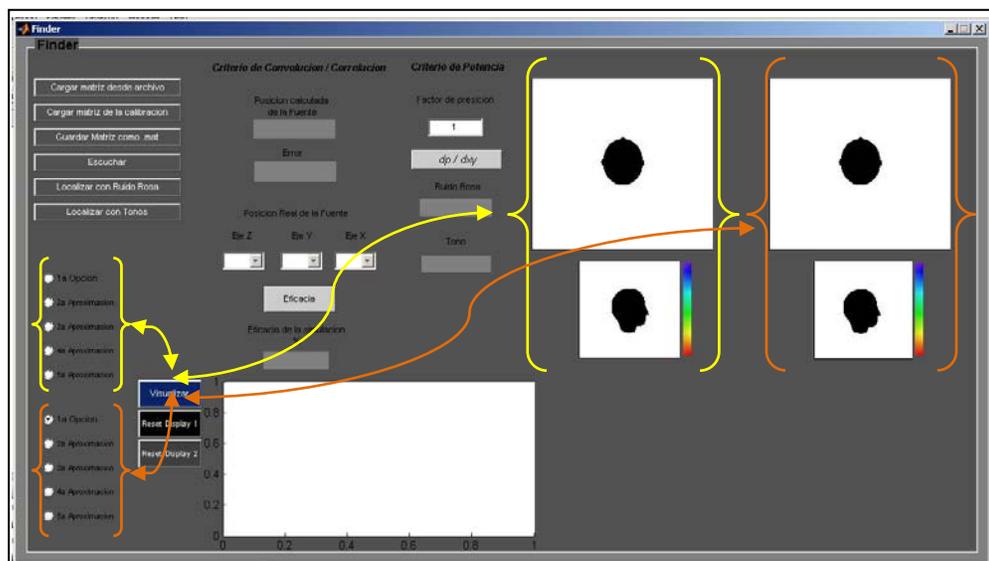


Imagen 4.19
Ventana del Módulo del Localizador (Finder), Sub módulo de visualización.

La función de visualización es aparentemente simple, puesto que le pide al programa mostrar en los indicadores bidimensionales una posición de fuente que ya ha sido calculada en la etapa anterior, no obstante, hay que saber qué es lo que se le está pidiendo al programa que muestre.

En el panel de selección única se encuentran dos grupos de botones rotulados de forma idéntica (Imagen 4.19). El grupo de la parte de arriba es para indicar entre las 25 posiciones posibles, aquellas 5 que se aproximaron más al valor esperado, luego de hacer los cálculos para las señales de ruido rosa. El grupo de abajo realiza la misma selección pero para las señales de tono. El módulo solo permite hacer una elección a la vez, pero permite visualizar el resultado de la localización con ruido rosa de forma separada al resultado obtenido con los tonos.

Es por esto que el módulo cuenta con dos pares de indicadores bidimensionales, las dos ventanas de la izquierda muestran la posición calculada para las señales de ruido rosa y los otros dos indicadores hacen lo mismo para las señales de tono.

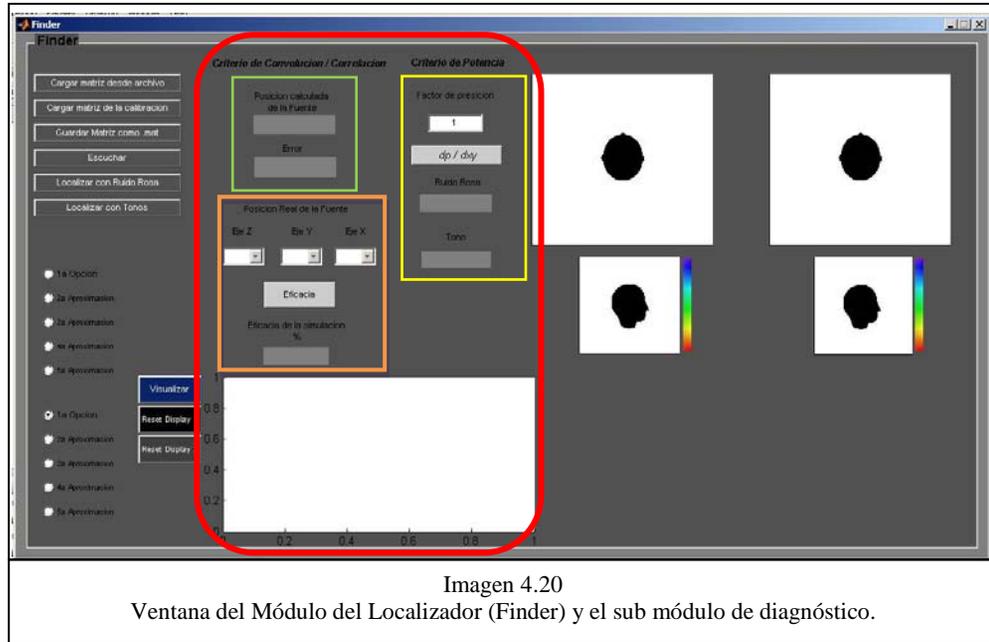
Los indicadores de la parte superior muestran la posición de la fuente en el plano xy, en tanto que el indicador de menor tamaño muestra la altura de la fuente, respondiendo con un cambio de color según la escala de colores que se encuentra a su derecha, de modo que, cuando la fuente sea calculada en una posición por encima de la cabeza el indicador se pondrá azul, si el sonido se detecta a la altura de la cabeza la coloración será verde y si el dictamen corresponde a un sonido proveniente del piso entonces el indicador se pondrá de color rojo.

Hasta este punto, solo se han expuesto las funciones básicas de La Aplicación y su módulo de Localización, son las mismas que se diseñaron en un principio cuando su implementación se limitaba a registrar un sonido y dictaminar después cual era su ubicación. No obstante, en este primer prototipo se estaba trabajando bajo la suposición de que el sistema en general sería altamente eficaz, lo cual constituyó un primer error de diseño. ¿Qué pasaría si luego de toda la rutina de preparación, calibración y ejecución de la simulación el sistema diera un dictamen errado?, ¿habría que repetir todo el proceso o solo la etapa posterior a la calibración?, ¿cuántas veces se toleraría el error en el dictamen antes de pensar en repetir todo el proceso?, estas fueron solo las preguntas iniciales.

El sistema difícilmente sería cien por ciento efectivo en su primera etapa de prueba, esto debió considerarse a la hora de diseñar el primer módulo de localización, sin embargo una serie de intentos fallidos hizo pensar en la necesidad de modificarlo.

Para el propósito de la investigación no era suficiente con que el programa arrojara un resultado que se limitara a dar una coordenada, sea que el dictamen haya sido acertado o no, era igualmente importante saber qué tan cerca se estuvo del resultado ideal y en qué porcentaje de intentos.

El localizador en su etapa inicial no contaba con el sub módulo de indicadores que se ubican en el centro de la ventana (Figura 4.20). Su propósito es brindar



al ingeniero una parte de la información que internamente se está procesando en la Matriz de Comparación, de manera que se puedan advertir los errores o comportamientos irregulares en el proceso.

En el sub módulo se destacan cuatro grupos de controles / indicadores. En la parte superior resaltado con un cuadrado verde, se encuentran dos casillas de texto en donde se visualizan la posición de la fuente y la magnitud del error calculadas. Por una convención establecida para este diseño, no se emplearon en un inicio coordenadas convencionales como las cilíndricas o las polares, esto obedecía a la idea de trabajar con posiciones fijas, establecidas con distancias angulares de 45 grados para todas las posiciones, en tres niveles de elevación para la fuente. Lo anterior no quiere decir que el módulo no se pueda ajustar al manejo de coordenadas convencionales, solo es cuestión de establecer las transformaciones.

El error calculado es una magnitud escalar, que resulta de un análisis numérico en el que se pondera la diferencia encontrada entre la señal cargada en la calibración y la registrada posteriormente.

El cuadrado de color naranja, encierra los controles e indicadores encargados de medir la eficacia de la simulación. Si bien el propósito del ejercicio es determinar la posición de la fuente en las tres dimensiones, es preciso

indicarle al programa la ubicación real del parlante, para esto se encuentran las tres casillas o menús desplegables rotulados “Eje Z”, “Eje Y” y “Eje X”. Podría pensarse que al ingresar la posición real de la fuente a localizar se estaría incurriendo en un error que invalidaría el análisis de comparación, es por eso que se advierte que este sub módulo trabaja de forma independiente y solo proporciona una interpretación porcentual de la eficacia del sistema sin alterar los algoritmos de localización.

Los valores posibles para la medición de la eficacia del sistema, bajo las condiciones de configuración actual, solo permiten cuatro valores que son: Eficacia del 99.9 % para un dictamen acertado en los tres ejes, Eficacia del 66.6 % para un dictamen acertado solo en dos ejes, Eficacia del 33.3 % para un dictamen acertado solo en un eje y finalmente un valor de 0% para una simulación fallida.

El cuadrado de color amarillo resalta los indicadores que evalúan las señales comparadas bajo el criterio de potencia o ILD. Se tiene una casilla para el resultado con ruido rosa y otra para el tono. El criterio de potencia establece la posibilidad de conocer la ubicación lateral de la fuente, de modo que su dictamen solo permite tres resultados que son: valores considerablemente negativos, que corresponden a posiciones de fuente a la izquierda de la cabeza, valores considerablemente positivos que son los generados por una fuente ubicada a la derecha del oyente y finalmente, valores ligeramente negativos o positivos que están asociados a las posiciones de fuente cercanas al plano medio que divide los oídos (plano yz).

La casilla rotulada con la expresión “Factor de Precisión” está allí para que el usuario digite el valor bajo el cual se determinarán los valores ligeramente negativos, es decir, marcará el umbral para los valores que se puedan considera centrales o laterales.

Finalmente, en la parte inferior del Localizador, se encuentra una ventana que grafica el margen de error de las diferentes posiciones, al ser comparadas con la señal a localizar, como se ilustra en la Imagen 4.21. Se trata de una gráfica de puntos, que se emplea en matlab bajo el comando “stem”.

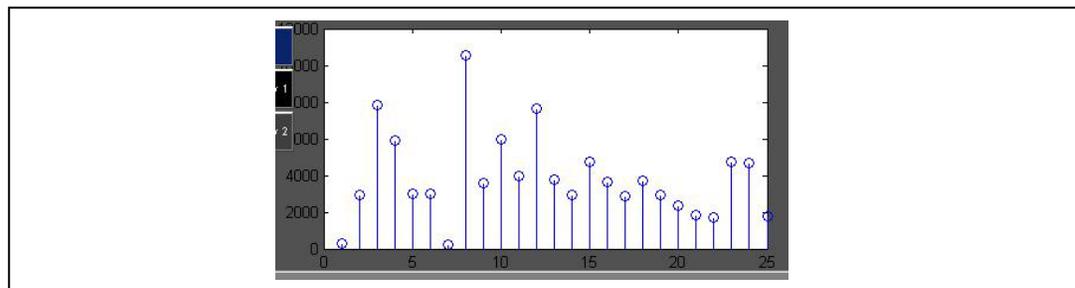


Imagen 4.21
Grafica del vector de Error en la ventana del Localizador.

¿Cómo interpretar la gráfica?

Hay que recordar el proceso que sufren las señales almacenadas en la matriz, en primer lugar, la convolución de las señales LR de calibración se somete a una correlación con la convolución de las señales LR del registro a localizar. Se presume que el mayor grado de correlación se presentará con aquella convolución de las señales LR de calibración que corresponda a la posición real de la fuente. Sin embargo, algunas variables como los rebotes generados por el recinto o el ruido presente en el mismo, pueden hacer que la función de transferencia de una misma posición de fuente cambie levemente de un instante a otro.

Teniendo en cuenta lo anterior, no se puede esperar que el algoritmo encuentre un caso de correlación perfecto, incluso para una misma señal registrada en dos instantes diferentes de tiempo, esto sería como recaer en el error que se cometió en el diseño de la ventana el Localizador. Lo que si se puede esperar es que la correlación se aproxime bastante para el caso de la posición real de la fuente y de igual forma, que se aproxime menos en los casos de las demás ubicaciones.

Una situación de correlación perfecta sería equivalente a calcular la auto correlación de cualquiera de las señales de calibración, es decir:

$$xcorr(y_{cal}, y_{loc}) = acorr(y_{cal})$$

Pero como ya se dijo antes, esto sería muy poco probable, no obstante se puede utilizar la anterior expresión para calcular el grado de aproximación de cada caso así:

$$Error = acorr(y_{cal}) - xcorr(y_{cal}, y_{loc})$$

Donde y_{cal} corresponde a la convolución de las señales LR de la calibración, y y_{loc} es la convolución de las señales LR del registro que se quiere localizar.

El cálculo del error es un procedimiento matemático, que se realiza con la posición que se quiere localizar y cada una de las 25 posiciones tomadas en la calibración, luego la expresión se modificaría así:

$$Error_n = acorr(y_{cal_n}) - xcorr(y_{cal_n}, y_{loc})$$

Para n que varía desde $n=1$ hasta $n=25$

De modo que, según este planteamiento, el valor de n para el cual se presente el menor valor de $Error$, correspondería a la posición real de la fuente.

La magnitud del vector error se grafica en la ventana de la parte inferior del módulo del localizador, teniendo en cuenta que cada uno de los 25 datos que componen el vector representan la diferencia para cada una de las posiciones de calibración, así, para el caso de la imagen 4.22, las posiciones más próximas a la ubicación real de la fuente son $n=1$ y $n=7$, en tanto que las más alejadas son $n=3$, $n=8$ y $n=12$, donde $n1$ y $n7$ corresponden a las posiciones $A(1,1)$ y $A(3,1)$, mientras que $n=3$, $n=8$ y $n=12$ son $A(1,3)$, $A(3,2)$ y $B(1,3)$ respectivamente.

En la imagen 4.22 se señalan las posiciones de mayor aproximación con flechas de color rojo y las de menor aproximación con flechas de color verde.

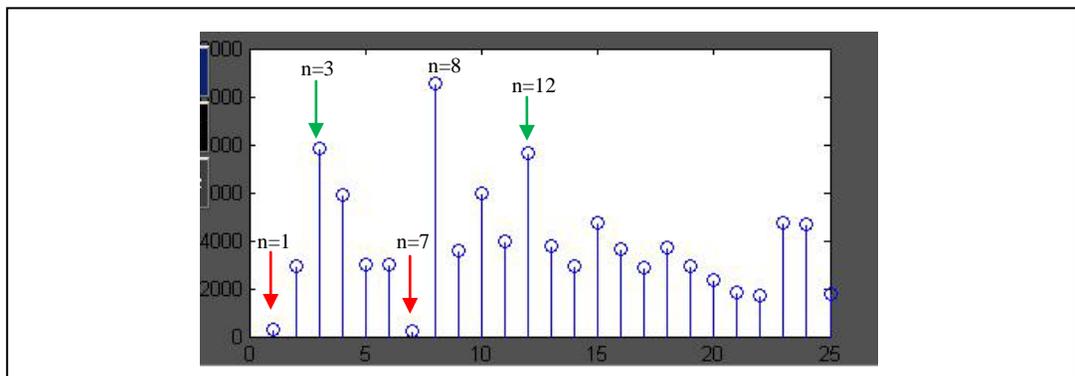


Imagen 4.22
Gráfica del vector de Error en la ventana del Localizador. Valores de mayor y menor aproximación.

Luego de un ejercicio de simulación y localización, tanto para las señales de ruido rosa como el tono, el Localizador debe tener una apariencia como la de la imagen 4.23.

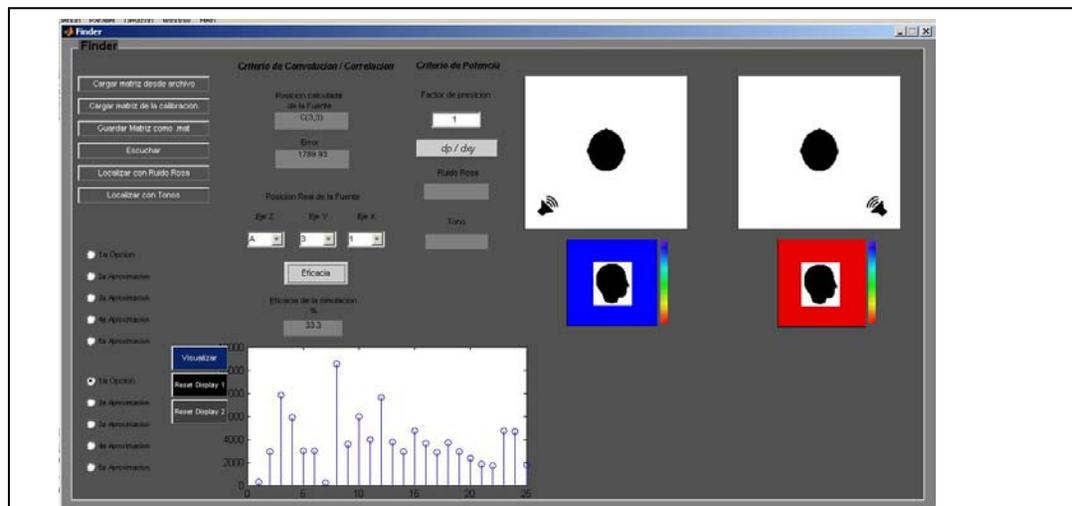


Imagen 4.23
Ventana del Módulo del Localizador (Finder) y visualización de la fuente con tonos con una eficiencia del 33%

Adicionalmente a las funciones y botones ya explicados, se agregaron dos rutinas más que se llamaron: “Cargar matriz desde archivo” y “Guardar matriz como .mat.”. La primera es una opción para obviar los datos de la calibración y llamar los datos de una calibración almacenada con anterioridad. La segunda opción es una rutina que permite al usuario almacenar la matriz de calibración que se halla realizado para utilizarla posteriormente, el formato en el que se guardan los datos lleva la extensión .mat que es el definido por matlab para este fin.

4.3 LA RUTINA DE CALIBRACIÓN

Se ha decidido dejar en un numeral aparte todos los detalles relacionados con el proceso de calibración, considerando que mezclar toda esta información en una misma sección, puede llegar a causar confusión y hacer más complicada la comprensión del funcionamiento de la aplicación.

Hay tres puntos a tener en cuenta con respecto a la calibración:

El primero es que el proceso consta de dos etapas: primero hay que hacer una verificación del correcto funcionamiento de los micrófonos, su respuesta en frecuencia y sensibilidad, asegurándose de que ante una señal de prueba el registro de los dos transductores sea lo más equilibrado posible, es decir que de ser necesario se debe implementar un ecualizador para igualar las respuestas de los micrófonos. La segunda etapa es el cargue de las señales de ruido rosa y tonos para cada una de las posiciones.

El segundo punto es que la calibración en su segunda etapa, requiere que se establezcan señales de prueba conocidas (ruido rosa y un tono de alguna de las 6 frecuencias posibles) así como posiciones de fuente claramente definidas.

El tercer punto es que sin importar el tipo de señal que se escoja, por limitaciones de tiempo de procesamiento, las señales no deben durar más de cinco segundos, pero como precaución y para evitar que la cantidad de muestras empleadas en el análisis sean suficientes, se sugiere no tomar registros con duraciones inferiores a los dos segundos.

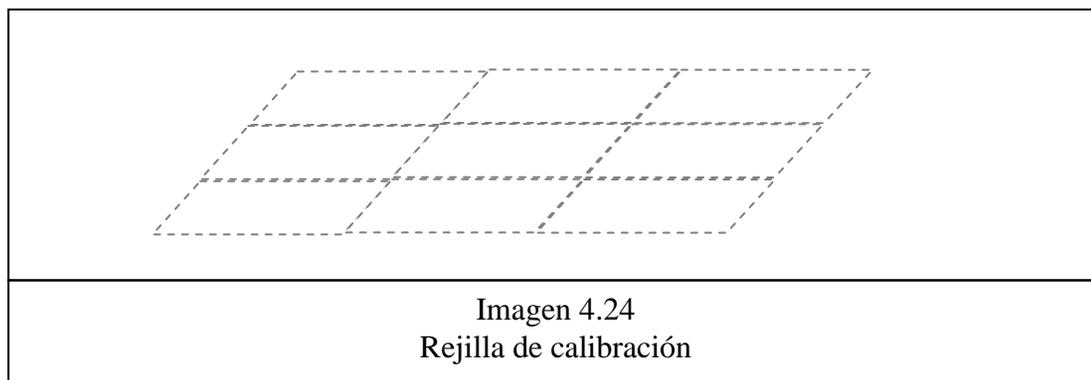
El proceso de calibración tiene principalmente dos objetivos que son:

- a. Verificar el correcto funcionamiento de las piezas que componen el sistema
- b. Generar la matriz de comparación o memoria del sistema.

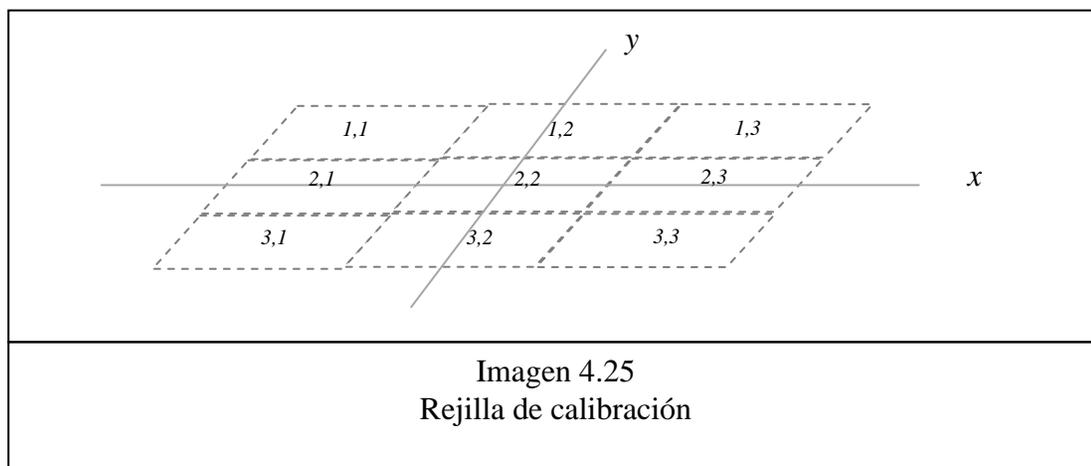
El primer objetivo no amerita mayor explicación y su cumplimiento radica en asegurar la integridad física de los elementos así como la correcta instalación y configuración de los dispositivos de software y hardware ya conocidos.

El segundo objetivo, por otra parte, implica recrear un modelo que a imitación del sistema auditivo humano, reciba una señal desde un punto en el espacio que sea conocido y posteriormente se almacene en una posición de una matriz a la cual se le asocia una coordenada específica.

Para iniciar el proceso, se debe marcar en el sitio de la calibración una rejilla cuadrada de 1.50 m de lado, con divisiones de 50 cm, obteniendo un figura con nueve cuadrados, cada uno de 50 cm de lado, como se muestra en la Imagen 4.24.



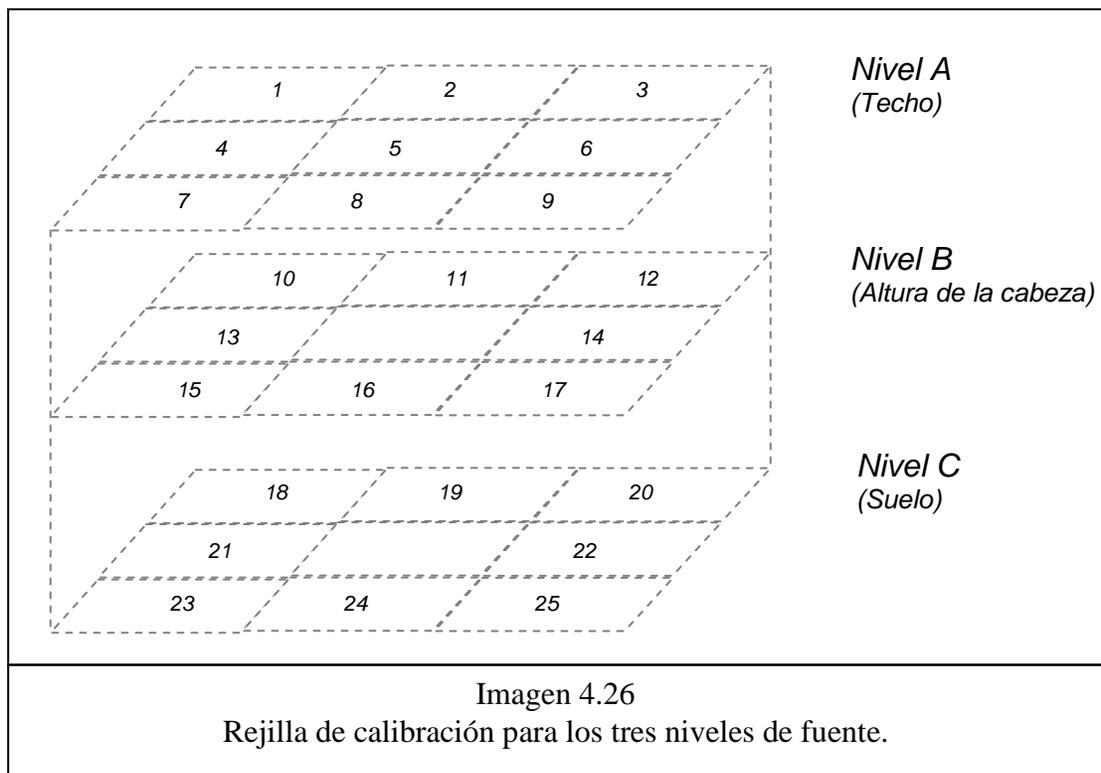
En el cuadrado del centro se ubica la cabeza binaural, cuidando que la altura de los micrófonos sea aproximadamente de 1.50 m. El propósito de la cuadrícula es tener claramente identificadas las posiciones que recorrerá la fuente, para asegurar que la variaciones sean mínimas entre un registro y otro.

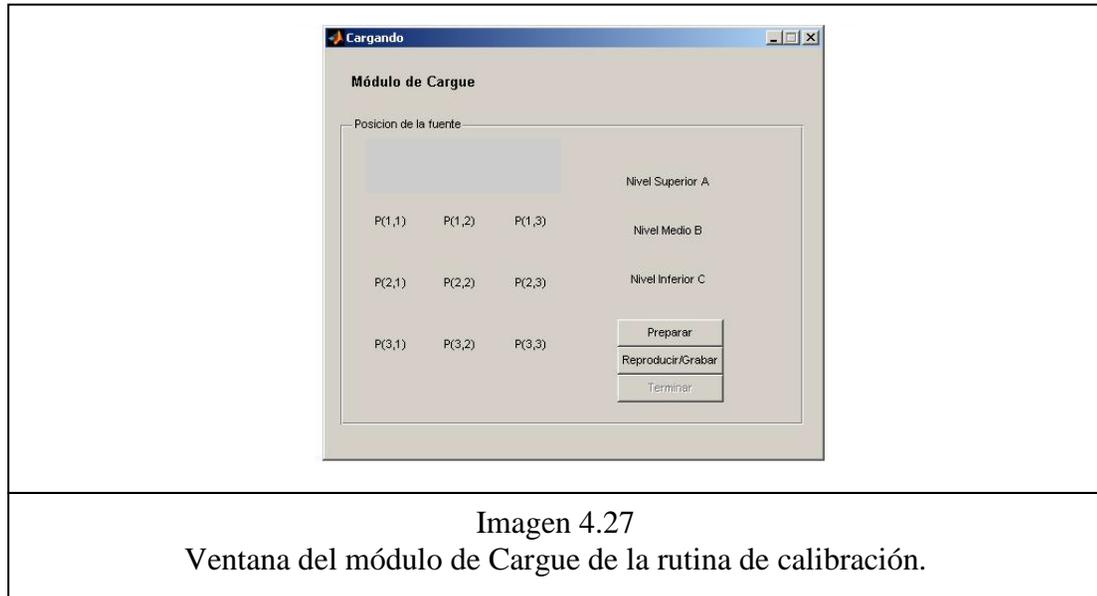


Las posiciones de fuente se establecieron siguiendo un esquema de organización matricial común, es decir, haciendo uso de dos números en cada posición de la cuadrícula, el primer número indica la fila y el segundo la columna. Esto implica una ventaja a la hora de trabajar directamente en la matriz.

Por la posición de los ejes como se muestran en la Imagen 4.25 se debe entender que la cabeza estará ubicada de manera que la nariz apunta hacia la posición (1,2), la oreja izquierda esta orientada del lado del punto (2,1) y la oreja derecha está del costado opuesto, es decir (2,3).

Estando la cabeza binaural en posición, se ubica la fuente en la primera casilla de la cuadrícula del nivel más elevado, es decir el nivel A, que por cuestiones de simetría, se sugiere ubicarla a una altura de 2.0 m de altura. En esta posición se ejecuta el comando de “Reproducir / Grabar” del respectivo módulo de calibración, para tomar registro del ruido rosa y luego del tono. Tan pronto como lo indique la ventana de notificación, se debe ubicar la fuente en la posición (1,2) del Nivel A y repetir el proceso de Reproducir / Grabar. Así sucesivamente hasta recorrer las 25 posiciones de los tres niveles, según el orden que se muestra en la imagen 4.26. Se puede observar allí mismo, que para las posiciones 2,2 tanto del nivel B como el C, no se toman registros, puesto que estas regiones están siendo ocupadas por la cabeza binaural y la base de piso.





En la imagen 4.27 se puede observar la relación que guarda la rejilla de calibración con la ventana del módulo de captura de registro de la rutina de calibración.

Por la estructura que actualmente tiene el módulo de captura para el proceso de calibración, en el eventual caso de que se presente un error durante la grabación de alguna de las señales, como por ejemplo un ruido externo, no será posible reemplazar únicamente el registro afectado sino que será necesario iniciar el proceso de Reproducción / Grabación desde la primera posición.

4.4 ARQUITECTURA GENERAL DEL SISTEMA DE LOCALIZACIÓN

Una vez visto en detalle el sistema y su funcionamiento, la estructura del mismo se puede sintetizar mediante el diagrama de la Imagen 4.28. Allí se puede ver de una forma sencilla la forma en que están interrelacionados los elementos tanto de software como de hardware involucrados en todo el proceso.

La flechas indican no solamente los vínculos existentes entre los diferentes elementos, también permiten deducir la dirección en que fluyen los datos para ser convertidos posteriormente en la información que el ingeniero requiera para el análisis.

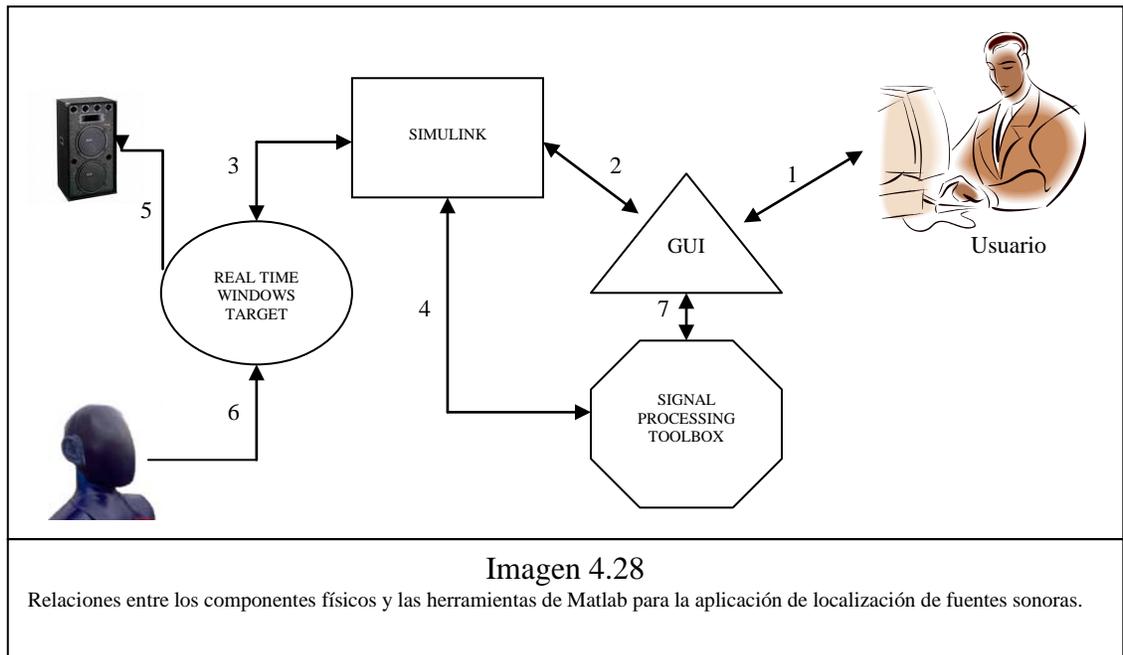


Imagen 4.28

Relaciones entre los componentes físicos y las herramientas de Matlab para la aplicación de localización de fuentes sonoras.

5 PRESENTACIÓN Y ANALISIS DE RESULTADOS

5.1 Datos Generales de la Simulación

Lugar: La simulación del Ejercicio de localización en su etapa final se desarrolló en el aula de sonido en vivo de la Universidad de San Buenaventura Bogotá. En etapas previas de diseño y ajuste de los algoritmos, se realizaron ejercicios de toma de registro en espacios abiertos, tales como parques y canchas de fútbol, con el fin de corroborar el funcionamiento y correcto encadenamiento de todas las piezas del sistema, de igual modo, para poder identificar los principales inconvenientes por presencia de ruido.

Horario: Considerando la disponibilidad de los equipos durante la jornada académica de la Facultad, las simulaciones de localización dentro de la Universidad para la etapa final del proyecto, se llevaron a cabo en jornadas que iban desde las 5:00 p.m. hasta las 8:30 o 9:00 p.m., en un periodo de tiempo de tres meses del primer semestre de 2009.

En las etapas iniciales de prueba, el horario de las simulaciones estuvo en las horas de la mañana y el medio día.

Ruido: La principal razón para trabajar en un recinto cerrado es reducir el nivel de ruido durante el ejercicio. El nivel de ruido registrado en la sala de sonido en vivo, para las simulaciones fue de 59 dB en promedio, para las jornadas de medición antes señaladas.

Señales empleadas para la simulación: Ruido rosa y tonos de 250 Hz, 500 Hz 1KHz, 2KHz, 4KHz y 8KHz.

Número de intentos realizados por ejercicio: 10 registros para cada banda.

5.2 Resultados Obtenidos de las simulaciones

A continuación se presentan los datos obtenidos en los ejercicios realizados para la simulación de localización con la señal de ruido rosa y con los tonos ya mencionados.

En las tablas se han resaltado algunas de sus celdas con tres colores diferentes, con el fin de exaltar tres comportamientos especiales que se presentaron durante las simulaciones.

En primer lugar, el color verde indica la posición para la cual se hizo un cálculo correcto de la localización de la fuente, es decir, se obtuvo una localización exitosa. El color azul, indica todos aquellos registros para los cuales se identificó de forma errónea una misma función de transferencia, es decir, una misma posición de fuente, a pesar de tratarse de posiciones diferentes reales del parlante.

El color naranja por su parte indica al igual que en el caso anterior, registros reales de posiciones diferentes, que según el programa corresponden a funciones de transferencia iguales o por lo menos de una misma posición de fuente (también se trata de dictámenes equivocados).

El color verde indica que la posición calculada coincide con la posición real de la fuente pero para la opción correspondiente. Si el sistema tuviese una eficacia del 100%, se obtendría una tabla con todas las celdas resaltadas con color verde en las columnas correspondientes a la “Opción 1”, sin embargo, como se puede apreciar en la tabla 5.1, el sistema está bastante apartado de este nivel de desempeño.

Intento	Posición real de la fuente	Ruido Rosa					Tono (250 Hz)				
		Opción 1	Opción 2	Opción 3	Opción 4	Opción 5	Opción 1	Opción 2	Opción 3	Opción 4	Opción 5
1	C(3,3)	C(1,2)	B(3,3)	C(2,3)	C(1,1)	A(2,3)	C(2,1)	B(3,3)	C(2,3)	C(1,3)	B(1,1)
2	C(1,2)	C(3,3)	C(2,3)	C(1,1)	A(2,3)	B(3,2)	C(2,1)	A(3,2)	B(3,3)	C(2,3)	A(3,3)
3	B(1,1)	C(2,3)	C(1,1)	A(2,3)	B(3,2)	C(1,2)	C(2,1)	B(3,3)	C(1,3)	C(2,3)	B(1,1)
4	B(2,1)	B(2,1)	B(1,1)	C(2,3)	C(1,1)	A(2,3)	C(2,1)	B(3,3)	C(2,3)	C(1,1)	C(1,2)
5	B(3,1)	C(2,3)	C(3,3)	C(1,1)	A(2,3)	B(3,2)	C(2,1)	B(3,3)	C(2,3)	C(1,3)	B(1,1)
6	B(2,3)	B(3,2)	C(2,3)	A(2,3)	C(1,1)	C(1,2)	C(2,1)	B(3,3)	C(2,3)	A(3,2)	C(1,3)
7	B(3,3)	C(2,3)	C(1,1)	A(2,3)	B(3,2)	C(1,2)	C(2,1)	B(3,3)	C(2,3)	C(1,3)	B(1,1)
8	B(1,3)	A(2,3)	C(1,1)	C(2,3)	B(3,2)	C(1,2)	C(2,1)	B(3,3)	C(2,3)	C(1,3)	B(1,1)
9	A(1,2)	C(2,3)	C(3,3)	C(1,1)	A(2,3)	B(3,2)	C(2,1)	B(3,3)	C(2,3)	C(1,3)	B(1,1)
10	A(1,3)	C(3,3)	C(2,3)	C(1,1)	A(2,3)	B(3,2)	C(2,1)	B(3,3)	C(2,3)	C(1,3)	A(3,2)

Tabla 5.1 Simulación con ruido rosa y tono de 250 Hz (Cabeza con pabellones)

El color azul, señala como se dijo anteriormente, las opciones para las cuales se detectó una misma posición de fuente a pesar de que en la realidad se trataba de ubicaciones diferentes del parlante. Lo particular de este comportamiento, radica en que es la evidencia de que el sistema lo está presentando de manera más pronunciada con las señales tonales que con las señales de ruido. La razón es que los micrófonos de la cabeza no pueden percibir una variación importante para posiciones diferentes de fuente a una frecuencia de 250 Hz, principalmente porque la longitud de onda de la señal generada supera enormemente la longitud de la cabeza, que es el obstáculo que separa los oídos, esto ya se había previsto. No obstante,

también hay que considerar el hecho de que esta frecuencia se presta para generar ciertos estados estacionarios dentro del recinto, haciendo más difícil producir un efecto notorio de la variación de la posición de la fuente sobre la cabeza binaural. Cuando se presentan estados estacionarios dentro del sitio de medición, las variaciones que se registran en los micrófonos nada tienen que ver con el propósito del proyecto, pues las variaciones de presión detectadas en la cabeza no obedecen a cambio de ubicación de la fuente, sino a cambios en la posición de la cabeza dentro del mismo recinto.

Las celdas resaltadas con tonos de color naranja, señalan la repetición de una misma posición de fuente para ubicaciones reales que están en posiciones diferentes, pero a diferencia del caso anterior, este comportamiento del sistema no obedece a la longitud de onda de las señales registradas.

Intento	Posición real de la fuente	Ruido Rosa					Tono (500 Hz)				
		Opción 1	Opción 2	Opción 3	Opción 4	Opción 5	Opción 1	Opción 2	Opción 3	Opción 4	Opción 5
1	C(3,3)	C(1,1)	A(3,2)	B(1,3)	A(3,3)	A(2,2)	B(1,2)	A(2,2)	C(2,1)	B(3,1)	C(3,2)
2	C(1,2)	C(1,1)	A(3,2)	B(1,3)	A(3,3)	A(2,2)	B(1,2)	A(2,2)	B(3,2)	C(3,2)	C(2,1)
3	B(1,1)	C(1,1)	A(3,2)	B(1,3)	C(3,2)	A(2,2)	B(1,2)	A(2,2)	C(2,3)	C(2,1)	C(3,2)
4	B(2,1)	C(1,1)	A(3,2)	B(1,3)	A(2,2)	A(3,1)	B(1,2)	A(2,2)	C(2,1)	B(3,1)	A(2,3)
5	B(3,1)	C(1,1)	A(3,2)	A(1,3)	B(1,3)	A(3,3)	B(1,2)	A(2,1)	A(2,2)	C(2,1)	C(3,2)
6	B(2,3)	C(1,1)	A(3,2)	B(3,2)	B(1,3)	A(2,2)	B(1,2)	A(2,2)	C(2,1)	C(3,2)	B(3,2)
7	B(3,3)	C(1,1)	A(3,2)	B(1,3)	B(3,2)	A(2,2)	B(1,2)	C(3,3)	A(1,2)	A(2,2)	C(1,3)
8	B(1,3)	C(1,1)	A(3,2)	B(1,3)	A(2,2)	C(3,2)	B(1,2)	A(2,2)	C(2,1)	A(2,1)	C(3,2)
9	A(1,2)	C(1,1)	A(3,2)	B(1,3)	A(3,3)	B(3,3)	A(3,2)	C(1,2)	B(1,2)	C(3,1)	C(1,1)
10	A(1,3)	C(1,1)	A(3,2)	A(1,3)	B(3,3)	B(1,3)	B(1,2)	A(2,2)	B(3,2)	C(2,1)	C(3,2)

Tabla 5.2 Simulación con ruido rosa y tono de 500 Hz (Cabeza con pabellones)

Si se observa detenidamente, en la tabla 5.2 las celdas resaltadas con naranja son en su totalidad registros obtenidos con las señales de ruido rosa. Lo anterior permite interpretar que el sistema está procesando la información, en algunos casos, bien sea con un nivel de ruido más elevado de lo conveniente, o también, se puede estar presentando en algún modo retroalimentación del sistema (feed back), lo cual hace que se obtengan dos registros muy similares para dos o más posiciones diferentes de fuente. La presencia de retroalimentación en el sistema durante las simulaciones es un comportamiento que se sabía que podía presentarse y por ello se sabía que debía controlarse cuidadosamente el nivel de radiación del parlante, así como el nivel de ganancia de los micrófonos, sin embargo existe la posibilidad de que ciertas frecuencias tiendan a retroalimentarse con mayor facilidad que otras, una vez que los micrófonos están montados en la cabeza, como resultado de la alteración ocasionada por los mismos pabellones, por lo tanto, no convenía realizar filtrados,

que con la intención de eliminar el feed back, suprimieran información que pudiera llegar a ser importante.

Vale la pena observar la tabla 5.3, en la que tomando registros a una frecuencia de 1KHz se aprecia que las celdas resaltadas con color azul han desaparecido, pero las de color naranja han aumentado. Bajo las condiciones de esta banda, ya no se obtienen los problemas de una longitud de onda grande y la respuesta modal del recinto; pero a 1KHz se está trabajando con una longitud de onda cercana a los 0.34m y la mitad de esta distancia es bastante próxima a la distancia promedio de separación de los oídos. Lo anterior puede entenderse, como una condición de cancelación de fase. El problema de este fenómeno es un poco más profundo de lo que parece, porque no se trata de una cancelación que simplemente reduce el nivel de presión en un oído cuando el sonido proviene de una dirección específica, desde la cual se puede generar un retardo equivalente a media longitud de onda. El problema radica en que este fenómeno equivale a reducir intencionalmente la ganancia de cualquiera de los micrófonos, lo cual el sistema interpreta como un cambio de posición y no como un cambio de frecuencia de registro. Por otra parte, al estar obligado a tener que controlar tan cuidadosamente los niveles de ganancia de los micrófonos, una reducción en la sensibilidad de algunos de estos, deja al sistema muy cerca del piso de ruido, generando de este modo un margen de incertidumbre mayor.

Intento	Posición real de la fuente	Ruido Rosa					Tono (1 KHz)				
		Opción 1	Opción 2	Opción 3	Opción 4	Opción 5	Opción 1	Opción 2	Opción 3	Opción 4	Opción 5
1	C(3,3)	A(2,2)	B(2,3)	A(3,2)	C(1,2)	A(2,1)	A(3,3)	A(2,2)	A(1,3)	B(3,2)	C(1,3)
2	C(1,2)	A(2,2)	B(2,3)	A(3,2)	C(1,2)	A(2,1)	C(1,2)	A(2,1)	B(3,2)	C(2,3)	C(2,1)
3	B(1,1)	B(2,3)	A(3,2)	A(2,2)	C(1,2)	A(2,1)	B(3,2)	A(1,2)	A(2,1)	C(1,2)	C(2,3)
4	B(2,1)	B(2,3)	A(3,2)	A(2,1)	C(1,2)	A(2,2)	A(3,2)	B(2,1)	B(3,2)	A(2,3)	A(3,1)
5	B(3,1)	A(2,1)	C(1,2)	A(3,2)	B(2,3)	A(2,2)	B(2,1)	A(2,3)	B(3,2)	A(3,2)	A(3,1)
6	B(2,3)	A(2,2)	B(2,3)	A(3,2)	A(2,3)	C(1,2)	B(1,1)	B(3,1)	C(3,2)	A(1,1)	C(3,1)
7	B(3,3)	A(2,2)	B(2,3)	A(3,2)	C(1,2)	A(2,1)	B(3,2)	A(2,1)	C(1,2)	C(2,3)	A(1,2)
8	B(1,3)	C(1,3)	A(2,2)	B(3,3)	A(1,3)	B(2,3)	A(3,2)	B(3,2)	B(2,1)	A(2,1)	A(2,3)
9	A(1,2)	C(1,2)	A(3,2)	B(2,3)	A(2,1)	A(2,2)	B(3,3)	C(3,1)	A(1,1)	C(3,2)	B(1,1)
10	A(1,3)	B(2,3)	A(2,2)	A(3,2)	C(1,2)	A(2,1)	B(2,1)	A(3,2)	B(3,2)	A(2,3)	A(3,1)

Tabla 5.3 Simulación con ruido rosa y tono de 1KHz (Cabeza con pabellones)

Las ventanas de visualización que se encuentran en la parte inferior de la ventana del Localizador (Imagen 5.1), están para mostrar de forma gráfica, para cada una de las 25 posiciones posibles de fuente, el grado de diferencia que hay entre la función de transferencia de la posición real de la fuente y la posición calculada de la misma. Dicho de otro modo, en esta gráfica (Imagen 5.2) observamos 25 magnitudes que representan el margen de error o diferencia de cada registro, de modo que, de las 25 posiciones, aquella que tenga una menor magnitud corresponderá a la posición de

fuente que según los cálculos del programa tiene una función de transferencia más parecida a la que se registró en el proceso de calibración. También se puede entender, como la gráfica de la función opuesta a la correlación, en la que se compara cada uno de los 25 registros almacenados en la memoria con el registro que se quiere localizar.

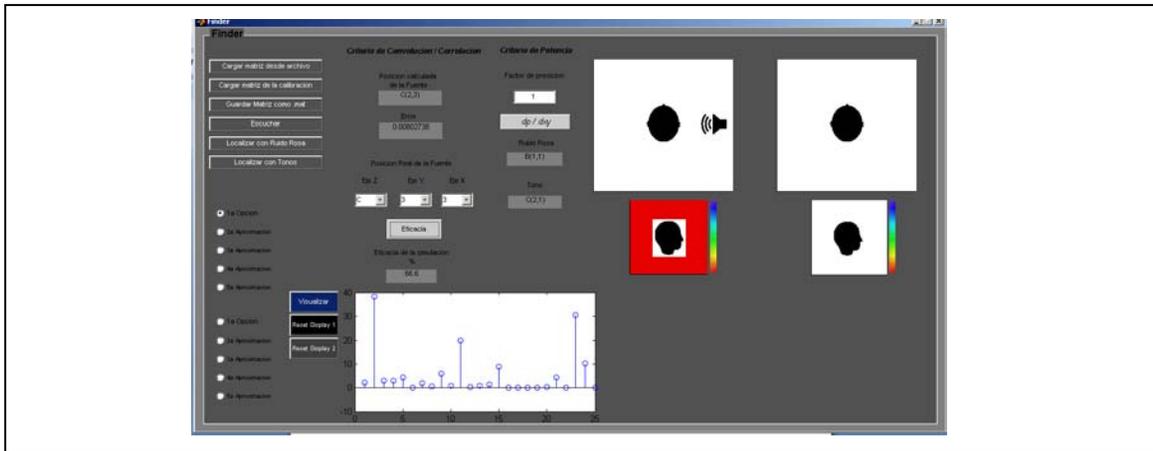


Imagen 5.1
Ventana de localizador y ventana de graficación del vector de error o aproximación

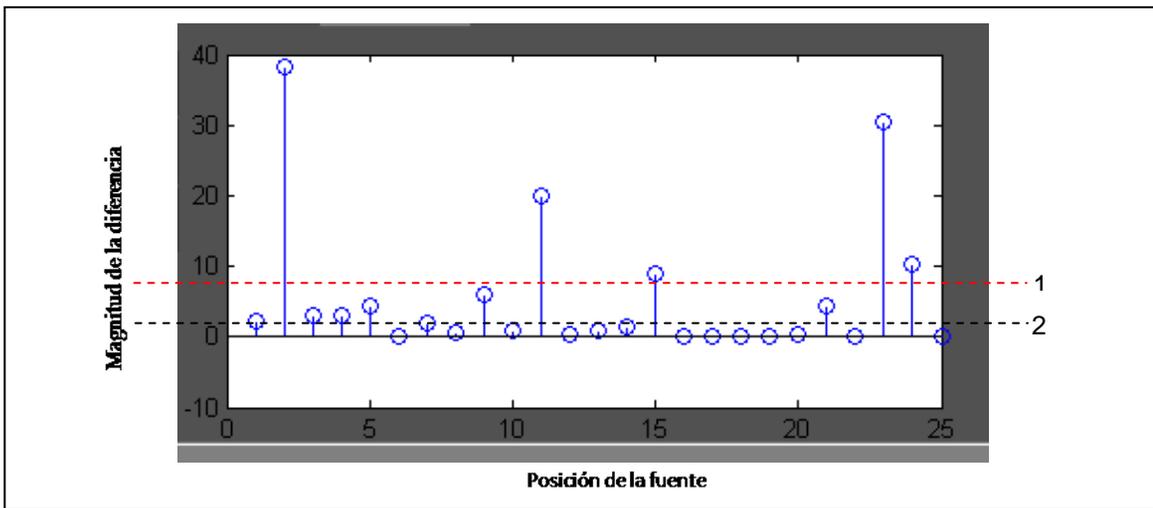
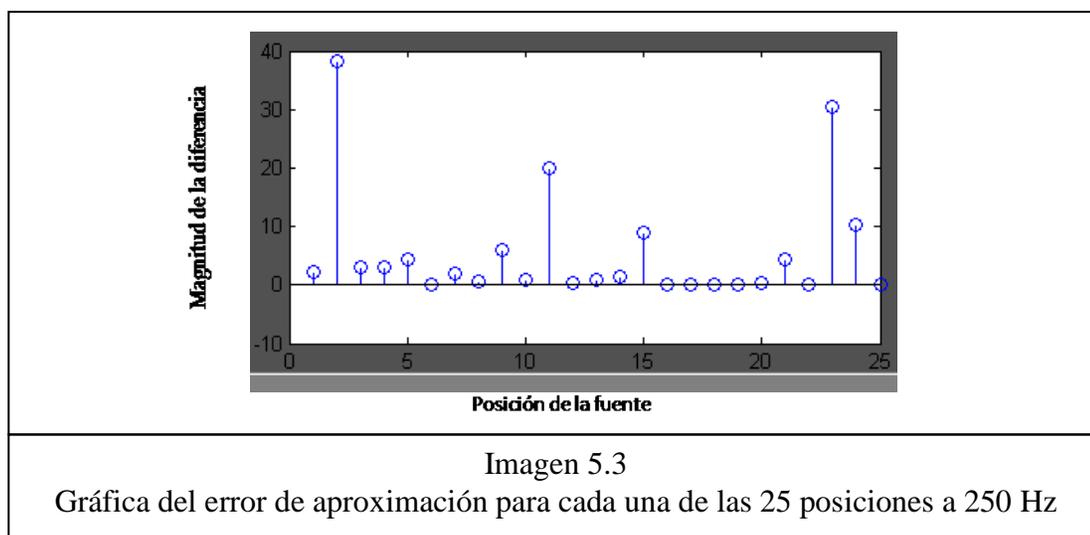


Imagen 5.2
Ventana para graficar el vector de error o aproximación

De la Imagen 5.2 se pueden extraer tres grupos de datos. El primero corresponde a los valores que están por encima de la línea puntada de color rojo y señalada con el número (1), estos son los valores de las funciones de transferencia más alejados de la posición real de la fuente. El segundo grupo es el de los valores encerrados entre

las líneas puntadas roja y negra, marcadas con los números (1) y (2) respectivamente, que corresponden a funciones de transferencia que son más similares a la de la posición real del parlante. El último grupo es el de todas aquellas posiciones que se registraron con valores o funciones muy similares entre si y además bastante próximas al valor de la posición real de la fuente, todas ellas ubicadas por debajo de la línea negra (2).

EL sistema emplea una rutina de selección, en la que busca dentro del vector “error”, aquel valor cuya magnitud sea menor, luego se puede ver de la imagen 5.2 que el valor para este caso será alguno de los que se encuentran por debajo de la línea punteada negra (2). La dificultad y lo que reduce el porcentaje de eficacia del sistema, es que el **criterio de Convolución – Correlación**, no permite aislar a una parte importante de las muestras; como se puede observar en la imagen 5.2 de este ejemplo, solo se tienen 5 posiciones de fuente en la zona superior y 7 posiciones más en la zona intermedia entre las líneas roja (1) y negra (2), lo que suman 12 posiciones de fuente que numéricamente han sido descartadas, dejando 13 posiciones más, que guardan entre si semejanzas muy grandes. Lo anterior quiere decir, que para el ejemplo de la imagen 5.2, el **criterio de localización de Convolución - Correlación** solo ha podido reducir el margen de incertidumbre sobre la posición de la fuente, en un 48%. La posibilidad de que el sistema encuentre dentro de las 13 posiciones restante, aquella que corresponda en realidad a la ubicación de la fuente, es mínima, teniendo en cuenta que la variaciones en este rango de datos puede obedecer a alteraciones involuntarias y no controladas como puede ser el ruido del lugar o señales de estática registradas durante el ejercicio. La imagen 5.2 del ejemplo anterior, corresponde a la gráfica del vector de error para las posiciones de fuente del caso en el que se tomó registro de la señal tonal de 250Hz. Este comportamiento se observó en todas las posiciones de fuente de esta frecuencia, dejando ver la dificultad para determinar una función de transferencia específica en cada posición de fuente por el criterio de Correlación – Convolución.



Comportamientos como el observado en la Imagen 5.3, se presentaron para ruido rosa con mayor frecuencia que para señales de tono.

La Imagen 5.4 muestra la gráfica del vector de error para la banda de 500Hz, a la izquierda para el tono y a la derecha para el ruido rosa. Como se puede observar, con la señal del tono, el simulador solo descartó 1 de 25 posiciones de manera consistente y con la señal de ruido rosa, las opciones se redujeron, pero en resumen no hubo un avance importante, pues se cuentan aún, cerca de 10 posibles posiciones de fuente.

Es evidente, que la presencia de señales tonales produce registros que al ser llevados a un proceso de convolución y posteriormente correlación, generan lecturas bastante diferentes y no son suficientemente útiles en el proceso de localización.

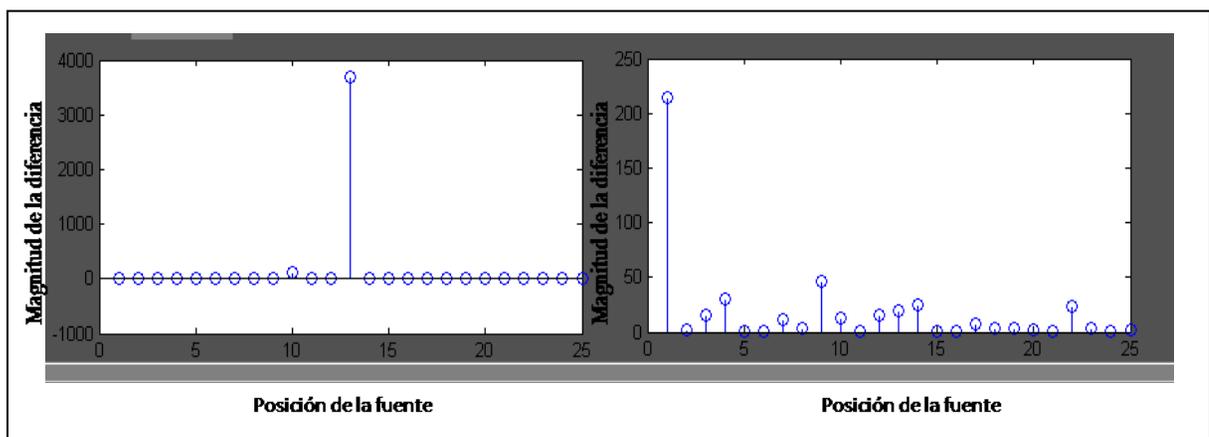


Imagen 5.4

Gráfica del error de aproximación para cada una de las 25 posiciones para la señal de 500Hz (a la izquierda) y el ruido rosa (a la derecha).

En la tabla 5.3 se presentan los resultados obtenidos en el sistema, para la simulación en la banda de 1KHz, en las mismas posiciones de los ejercicios realizados para las frecuencias de 250Hz y 500Hz y las graficas del vector de error de la misma simulación se presentan en la Imagen 5.5.

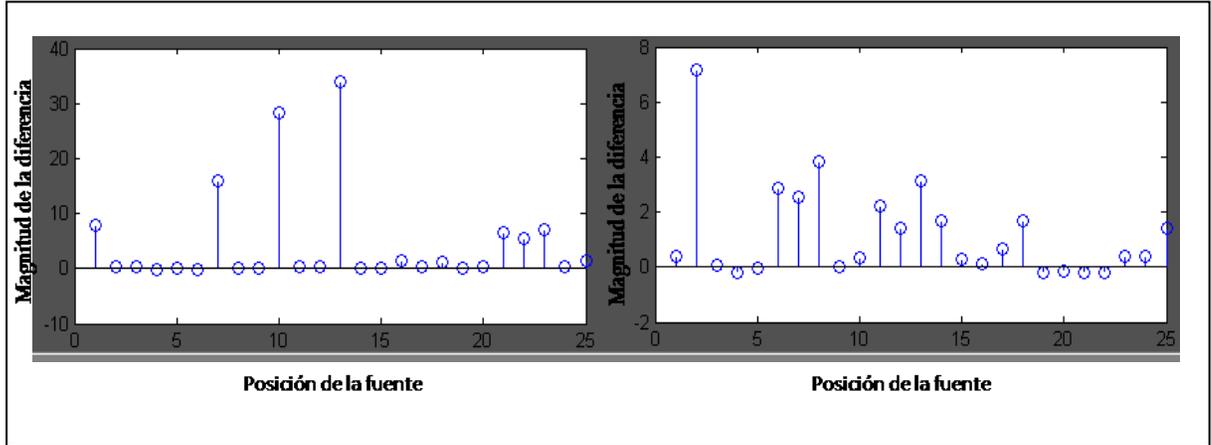


Imagen 5.5

Gráfica del error de aproximación para cada una de las 25 posiciones para la señal de 1KHz (a la izquierda) y el ruido rosa (a la derecha).

En las gráficas de la imagen 5.5, se puede observar que el grado de discriminación mejora tanto en los registros tonales como en el caso del ruido rosa comparando con las frecuencias de 250Hz y 500Hz, no obstante, continúan presentándose diferencias muy pequeñas entre una cantidad importante de posiciones.

Intento	Posición real de la fuente	Ruido Rosa					Tono (2 KHz)				
		Opción 1	Opción 2	Opción 3	Opción 4	Opción 5	Opción 1	Opción 2	Opción 3	Opción 4	Opción 5
1	C(3,3)	C(1,2)	C(2,3)	C((2,1)	A(1,3)	A(2,1)	B(1,2)	B(3,3)	B(1,3)	C(1,1)	C(3,2)
2	C(1,2)	C(2,3)	C(1,2)	C(2,1)	A(1,3)	A(2,1)	B(1,3)	C(1,1)	B(1,2)	B(1,1)	B(2,3)
3	B(1,1)	C(1,2)	C(1,1)	C(2,3)	C(2,1)	A(1,3)	B(1,2)	B(1,3)	C(1,1)	B(1,1)	B(2,3)
4	B(2,1)	C(1,2)	C(2,3)	C(2,1)	A(1,3)	C(3,3)	C(1,1)	B(1,3)	B(1,2)	B(1,1)	B(2,3)
5	B(3,1)	C(1,2)	C(2,3)	C(2,1)	A(1,3)	A(2,1)	B(1,2)	B(1,3)	C(1,1)	B(1,1)	B(3,3)
6	B(2,3)	C(1,2)	C(2,3)	B(3,2)	C(2,1)	A(1,3)	B(1,2)	A(3,1)	B(1,3)	C(1,1)	B(1,1)
7	B(3,3)	C(1,2)	C(2,3)	C(2,1)	C(1,1)	A(1,3)	B(1,2)	B(1,3)	C(1,1)	B(1,1)	C(3,2)
8	B(1,3)	C(1,2)	C(3,2)	C(2,1)	A(1,3)	A(2,1)	C(1,1)	B(1,3)	B(1,2)	B(1,1)	B(3,3)
9	A(1,2)	C(1,2)	C(2,3)	C(1,1)	C(2,1)	A(1,3)	B(1,2)	B(1,3)	C(1,1)	B(1,1)	B(2,3)
10	A(1,3)	C(1,2)	C(2,3)	B(1,1)	B(3,3)	C(2,1)	B(1,3)	C(1,1)	B(1,2)	B(1,1)	B(2,3)

Tabla 5.4 Simulación con ruido rosa y tono de 2KHz (Cabeza con pabellones)

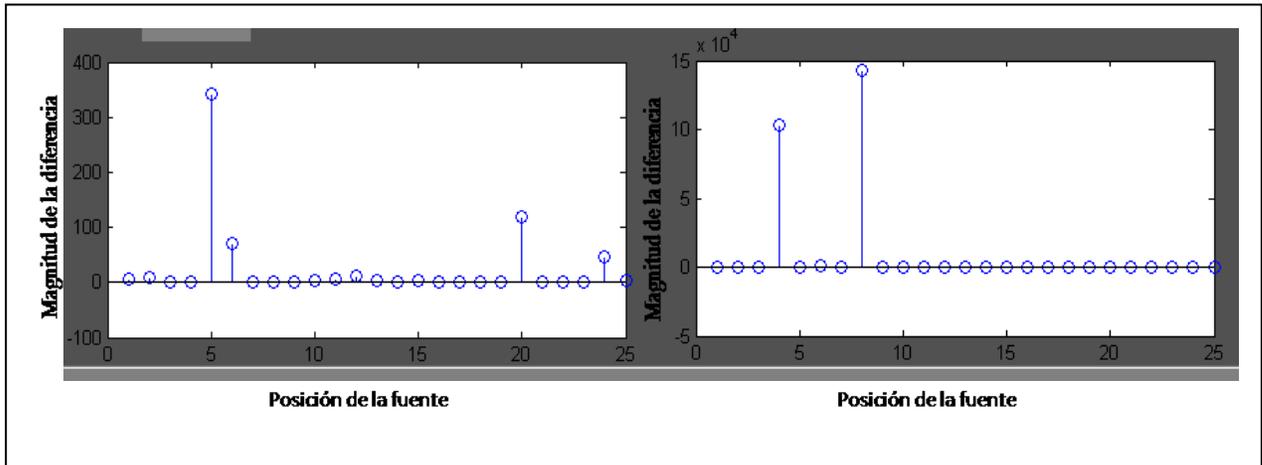


Imagen 5.6

Gráfica del error de aproximación para cada una de las 25 posiciones para la señal de 2KHz (a la izquierda) y el ruido rosa (a la derecha).

Al observar los resultados obtenidos en las simulaciones a 2KHz, se puede notar la aparición nuevamente de posiciones que se repiten en diferentes ubicaciones de la fuente, pero gracias a las gráficas del vector error, es posible ver que los registros tomados tanto para la señal tonal como para el ruido rosa no fueron de utilidad para este ejercicio, puesto que en el caso del tono, solo se tienen cuatro posiciones claramente descartadas y en el caso del ruido rosa no se llegaron a descartar de forma contundente más de dos posiciones de fuente.

Intento	Posición real de la fuente	Ruido Rosa					Tono (4 KHz)				
		Opción 1	Opción 2	Opción 3	Opción 4	Opción 5	Opción 1	Opción 2	Opción 3	Opción 4	Opción 5
1	C(3,3)	C(3,3)	B(3,2)	C(3,1)	B(1,2)	B(2,3)	C(1,2)	C(1,3)	C(2,3)	C(3,2)	A(1,2)
2	C(1,2)	C(1,1)	B(2,3)	A(3,2)	A(2,1)	A(1,2)	C(1,3)	C(1,2)	A(1,2)	C(2,3)	C(3,2)
3	B(1,1)	C(3,1)	B(3,2)	B(1,2)	B(2,3)	C(1,3)	C(1,3)	A(1,2)	C(1,2)	C(2,3)	C(3,2)
4	B(2,1)	C(1,3)	B(2,3)	B(3,3)	A(1,3)	B(1,2)	C(1,3)	A(1,2)	C(1,2)	C(2,3)	C(3,2)
5	B(3,1)	A(1,3)	B(3,3)	B(2,3)	C(1,3)	B(1,2)	C(1,3)	A(1,2)	C(1,2)	C(2,3)	C(3,2)
6	B(2,3)	A(1,3)	B(2,3)	B(3,3)	C(1,3)	B(1,2)	C(1,3)	A(1,2)	C(1,2)	C(2,3)	C(3,2)
7	B(3,3)	B(3,3)	B(2,3)	A(1,3)	C(1,3)	B(1,2)	C(1,3)	C(1,2)	A(1,2)	C(2,3)	C(3,2)
8	B(1,3)	B(3,3)	A(1,3)	B(2,3)	C(1,3)	B(1,2)	C(1,3)	A(1,2)	C(1,2)	C(2,3)	B(3,1)
9	A(1,2)	B(2,3)	A(1,3)	B(3,3)	C(1,3)	B(1,2)	C(1,3)	A(1,2)	C(1,2)	C(2,3)	C(3,2)
10	A(1,3)	B(2,3)	A(1,3)	B(3,3)	C(1,3)	B(1,2)	C(1,3)	A(1,2)	C(1,2)	C(2,3)	C(3,2)

Tabla 5.5 Simulación con ruido rosa y tono de 4 KHz (Cabeza con pabellones)

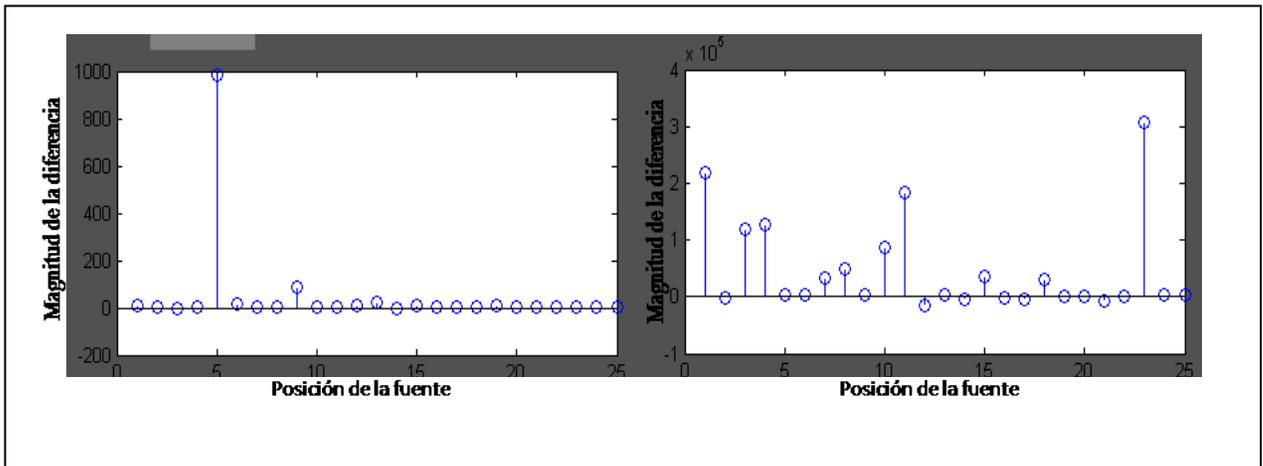


Imagen 5.7

Gráfica del error de aproximación para cada una de las 25 posiciones para la señal de 4KHz (a la izquierda) y el ruido rosa (a la derecha).

Para las gráficas de la Imagen 5.7, correspondientes al grupo de datos de la tabla 5.5, se puede determinar que se obtuvieron mejores resultados con la señal de ruido rosa que con la señal tonal de 4KHz. Vale la pena aclarar que este tipo de diferencias entre los comportamientos de las dos clases de señales, eran una de las razones que motivó a modificar la ventana que inicialmente tenía el Localizador, pues en su versión de prueba solo se podía saber si la posición calculada de la fuente era o no correcta, más no se podía apreciar el margen de error en el cálculo de aproximación en cada intento y con cada una de las señales.

La simulación para la banda de 8KHz no se tuvo en cuenta, considerando que las lecturas y los registros que arrojó la gráfica de error para esta frecuencia se asemejaba bastante al caso de los 2KHz, por lo tanto no existen posiciones de fuente calculadas exitosamente.

5.3 Análisis del Ejercicio de Simulación y de los resultados obtenidos.

5.3.1 Los Resultados

Lo que en la presente investigación se consideran como resultados, se han clasificado en dos categorías, la primera es la que se denominó *Funcionamiento del Sistema* y la segunda es la *Eficacia del Sistema*.

5.3.1.1 *Funcionamiento del Sistema:*

El funcionamiento del sistema evalúa que todas las piezas que lo constituyen, tanto físicas como lógicas, estén encadenadas y configuradas adecuadamente, de manera que se permita una captura, almacenamiento y procesamiento de la información, conforme los requerimientos de diseño que se quieran implementar.

Bajo estas consideraciones, hay que decir que se consiguió desarrollar un sistema de análisis binaural, con una aplicación de software en Matlab para localización de fuentes sonoras, haciendo uso de herramientas gráficas y ventanas de diálogo que permiten al usuario manipular los recursos del sistema, desde el momento de la captura de los datos hasta la graficación de los mismos.

Para el desarrollo del sistema se elaboró una rutina de calibración basada en la observación y análisis del funcionamiento del proceso de localización en los seres humanos, introduciendo el concepto de memoria. Se implementaron algoritmos que asocian las señales almacenadas en la memoria con la señal que se quiere localizar, todo esto, a través del entorno gráfico del GUI.

La Eficiencia de los elementos físicos del sistema es la esperada desde su etapa de diseño, tanto la cabeza, como el torso y la base de piso se ajustaron a las necesidades de los diferentes entornos de simulación. Los recursos de hardware representados en la tarjeta de sonido y el computador no presentaron ningún inconveniente importante, salvo la susceptibilidad a la estática en el aula de sonido en vivo⁹.

La Eficiencia de la Aplicación está dentro de los límites esperados, no se debe confundir este término con el de *Eficacia* de la Aplicación, pues esta última está asociada con la asertividad del Criterio de Localización escogido. La Eficiencia se refiere a la manera como se comportan todos los algoritmos tanto independientemente como al trabajar encadenados, entre las diferentes etapas del proceso, generando resultados, solicitándole información al usuario, procesando los datos dentro de límites de tiempo razonables, etc. En este aspecto, también se obtuvieron resultados positivos, no obstante, es preciso resaltar, que el proceso de cálculo de las convoluciones y posteriores correlaciones, en el momento de cargar la matriz, tarda cerca de 15 minutos en un equipo de configuración promedio.

⁹ Pequeñas descargas pueden ser registradas en el sistema al trabajar en un piso alfombrado, y aunque una conexión a tierra puede reducir el riesgo de las descargas, dependiendo de las características y conexiones internas de los elementos electrónicos involucrados, no aterrizar el sistema en algunos casos ayudó a reducir el ruido eléctrico en el parlante y/o amplificador, cuando este fenómeno se presentaba, particularmente con cabinas auto amplificadas.

5.3.1.2 Eficacia del Sistema:

Debe asociarse este término al grado de asertividad que tiene el Criterio de Localización escogido. En este aspecto hay que decir y apoyándose en la información consignada en las tablas de este mismo capítulo, que la eficacia del sistema es deficiente, en las cinco frecuencias estudiadas, de las seis que se habían propuestos al comienzo, ningún ejercicio logró demostrar que el *Criterio de Localización* de Convolución – Correlación fuera efectivo en más de un 50% de los intentos realizados. El bajo porcentaje de aciertos obtenidos durante las simulaciones (Tabla 5.6) no permite tener un nivel de confianza siquiera aceptable, al trabajar bajo el criterio de Convolución – Correlación.

Frecuencia (Hz)	No de aciertos	Porcentaje
250	3 de 20	15%
500	2 de 20	10%
1000	4 de 20	20%
2000	3 de 20	15%
4000	6 de 20	30%

Tabla 5.6

A pesar de que el sistema trabaja conforme a los parámetros de diseño, el margen de incertidumbre que hay trabajando bajo el Criterio de Localización actual es bastante amplio y por lo mismo, es muy difícil para el sistema determinar la posición real de una fuente sonora, bien sea con una señal tonal, o con un sonido cargado de mayor cantidad de información, como el ruido rosa.

No obstante, y gracias a las herramientas de visualización de la ventana del Localizador, se puede establecer que es muy pequeña la posibilidad de obtener resultados más eficaces. Se podría pensar que si las condiciones de simetría de los micrófonos se mejoraran y/o las condiciones anecoicas del lugar de la simulación se optimizan en un recinto más apropiado se podrían reevaluar los resultados actuales, pero siendo consciente de que estas dos condiciones no garantizan una eficacia mayor, repetir el ejercicio de simulación sería un gasto injustificado de tiempo sin un nuevo criterio de localización para evaluar.

En este punto y habiendo mencionado al inicio del proyecto los criterios de localización denominados como ITD e ILD, cabe hacer la pregunta ¿Por qué no se emplearon estos otros dos criterios para realizar las simulaciones?

La verdad es que si se emplearon, pero en la etapa inicial de desarrollo del primer prototipo de la aplicación de Localizador, no obstante se detectaron las siguientes dificultades:

En primer lugar, el criterio de *Diferencia Interaural de Tiempo (ITD)*, que implica un análisis de fase, estaba seriamente comprometido al desarrollar el ejercicio en un recinto de condiciones no anecoicas. La cantidad indefinida de rebotes solo se prestaba para que el sistema registrase como si hubiera no una sino varias fuentes.

En segundo lugar, el criterio de *Diferencia Interaural de Nivel (ILD)*, que es el principal responsable de la localización en el plano horizontal, comprometía seriamente la eficacia del sistema, considerando que la respuesta en frecuencia de las dos cápsulas (oído izquierdo y derecho de la cabeza) son diferentes.

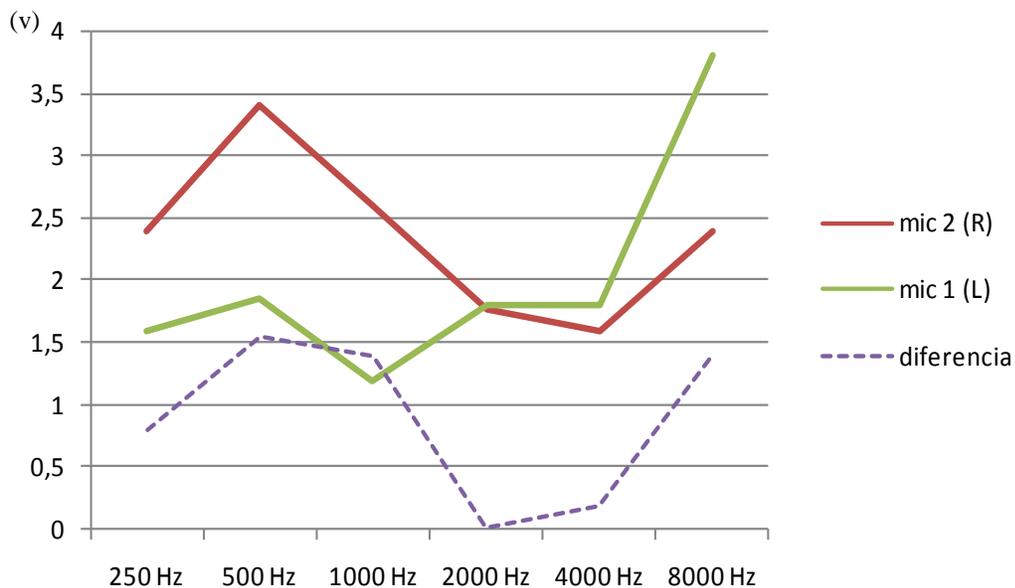
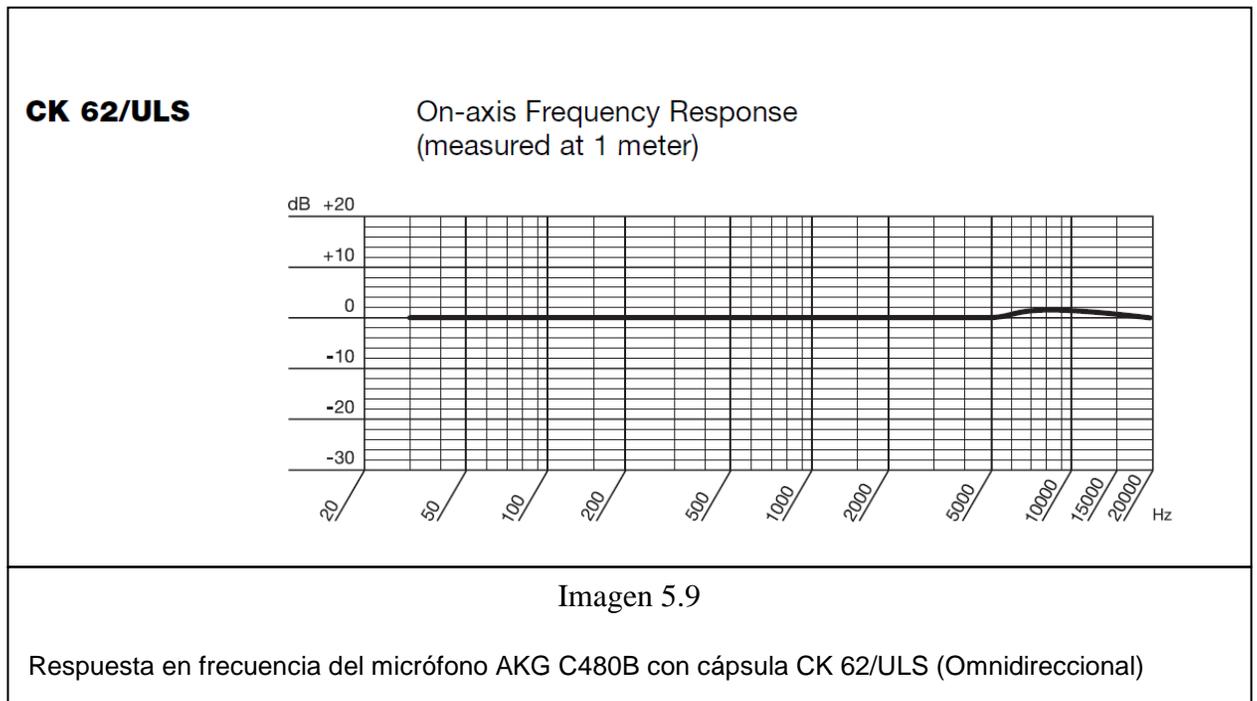


Imagen 5.8

En la gráfica de la imagen 5.8, se ilustra la respuesta de los micrófonos AKG C480B, luego de realizar una prueba de sensibilidad a los mismos, haciendo uso de las señales tonales respectivas a 94 dB, a una distancia de 1m. Los valores fueron registrados en unidades de voltio con ayuda de osciloscopio.

Considerando que el comportamiento de los micrófonos, de acuerdo con la carta de especificaciones del fabricante deben tener unas características como las que se pueden ver en la imagen 5.9, que a pesar de ser una gráfica en escala logarítmica y unidades de decibel, corresponde a una respuesta completamente plana para el rango de frecuencias de interés, gracias a esto, se puede entender, en parte, la dificultad para ubicar al sistema en un punto de trabajo que esté por encima del piso de ruido y a la vez por debajo del punto de retroalimentación, con una ecualización que no sea tan drástica que elimine información que pueda ser importante para el procesamiento de la

señal ni tan ligera que permita el paso de todo el ruido ambiente, teniendo en cuenta también, que cada unidad externa de ecualización hace una contribución de ruido y un proceso de ecualización desde matlab hace más lento el procesamiento de datos y el funcionamiento de la aplicación en general.



Así pues, se ideó un nuevo criterio, del cual no se tenía antecedentes en el campo de la localización binaural y que fue implementado en el sistema desarrollado en el presente proyecto, el Criterio de Localización por Convolución-Correlación Interaural.

Sin embargo, los resultados dejan en claro, que no hay un margen de confianza aceptable, que permita llevar al sistema a etapas de simulación donde se requiera una eficiencia de por lo menos el 75%. Si se hubiese evidenciado que de las cinco opciones posibles dentro de la ventana del Localizador, la ubicación real de la fuente siempre se encontraba dentro de estas cinco, se podría afirmar que el sistema bajo el criterio de localización escogido, estaba en capacidad de ubicar una fuente con un margen de error del 20%, lo cual sería equivalente a decir que el sistema tiene la capacidad de reducir el margen de incertidumbre en un 80%.

5.3.2 Resultados del sistema USB Radar y los antecedentes documentados

En una investigación adelantada por el Departamento de Acústica de la Universidad de Aalborg¹⁰, se realizaron pruebas de localización, empleando tres tipos de unidades de captura. En primer lugar, emplearon personas comunes que simplemente indicaban el lugar de la fuente sonora tras escuchar un sonido. En otro ejercicio, utilizaban a personas a las cuales les instalaban pequeños micrófonos al interior de los oídos y luego tomaban registro de las señales a localizar. En un último ejercicio, emplearon cabezas artificiales o dummies como se les conoce y tomaron nuevamente registro de las señales a localizar.

En estos experimentos se emplearon señales de audio con registro de voz y ruido blanco, grabado a una tasa de muestreo de 48000 Hz, a una distancia de 1m para 19 posiciones de fuente diferentes. Inicialmente se realizaron ejercicios en cámara anecoica y posteriormente en recintos comunes.

Se emplearon al menos ocho cabezas binaurales diferentes entre las que se encuentran: KU100, 4100, HMSII, MK1, ITA, VALDEMAR, AVH, y DOL, correspondientes a casas o fabricantes diferentes.

Las pruebas arrojaron resultados que se resumen en los siguientes enunciados:

- Se emplea el criterio de Correlación Interaular Cruzada (HRTF)
- Resultados altamente diferentes a favor de las pruebas con cabezas humanas.
- Se evidenció comportamiento menos eficiente en el modelo sin hombros ni torso.
- Dificultad de discriminación entre las regiones delantera y trasera.
- Dificultad de localización en el plano medio.
- El 60 % de las cabezas naturales tenían un desempeño mejor que la mejor de las cabezas artificiales.
- Se hace necesario mejorar los modelos de cabezas artificiales.

¹⁰ PAULI MINNAAR, *AES Member*, SOREN KRARUP OLESEN, FLEMMING CHRISTENSEN, *AES Member*, AND HENRIK M-ILLER, *AES Member*. Localization with Binaural Recordings from Artificial and Human Heads. Department of Acoustics, Aalborg University, DK-9220 Aalborg, Denmark. J Audio Eng. Soc, Vol 49, No 5, 2001 May

El anterior, es en resumen, un trabajo publicado por AES, que tiene por objeto en la presente investigación, exponer las grandes dificultades que se tienen para igualar el nivel de desempeño de una cabeza natural, incluso en condiciones de laboratorio. No obstante, en esta publicación, no se observa una manipulación de datos semejante a la que se propuso con la aplicación USB Radar

6 CONCLUSIONES

Tomando como base la información obtenida en las lecturas arrojadas por la aplicación de software, la cual esta consignada en las tablas, se puede afirmar que si el fenómeno de localización en humanos emplea el recurso de memoria para hacer un reconocimiento acústico de la dirección de proveniencia del sonido, no es por medio de un proceso que involucre de manera exclusiva un mecanismo de correlación, tal como se plantea en la hipótesis, de manera que, aunque el instrumento elaborado para el estudio del presente proyecto cumple con las condiciones de diseño planteadas, se reconsidera la efectividad del criterio de localización.

El Sistema de Localización Binaural Artificial desarrollado en la presente investigación, cumple con los parámetros de diseño establecidos para la misma, atendiendo los principales requerimientos del Ingeniero en cuanto a las herramientas de software y hardware disponibles. Igualmente, el sistema responde de manera adecuada a los cambios y modificaciones que el diseñador considere conveniente, teniendo en cuenta dificultades que se presenten en el proceso. La aplicación está propuesta de forma que se pueda ajustar a nuevos parámetros de diseño y condiciones de localización, que puedan extenderse más allá de los objetivos del presente proyecto.

Basándose en el funcionamiento del sistema de localización humano, es preciso decir que un sistema de simulación no puede tener un criterio de localización tan simple que lo limite a localizar en solo una o dos dimensiones, pues no cumpliría con su cometido y no sería de gran utilidad en tareas que requieran una simulación real, pero tampoco puede ser un criterio tan complejo que no se pueda ajustar a las herramientas tecnológicas con que pueda contar el ingeniero.

La identificación de una posición de fuente debe ser un procedimiento que no represente un número gigantesco de algoritmos matemáticos para una sola ubicación; si lo que se quiere es imitar al modelo humano, se debe tender a la simplicidad y no a la complejidad. En el desarrollo de los algoritmos para estas simulaciones se emplearon secuencias que pusieron a prueba los recursos del equipo de procesamiento, lo cual hace pensar que la falla en el sistema de localización radica esencialmente en la escogencia del criterio de localización, más que en las condiciones de funcionamiento de los elementos físicos del sistema, o del recinto como tal.

A pesar de que la eficacia del sistema no estuvo dentro de los límites esperados, la eficiencia del sistema en general, permite al investigador proponer en un futuro

nuevos criterios de localización que se pueden implementar en este mismo sistema. Hay que recordar que, apartándose de las dificultades que representan las condiciones de los micrófonos y el recinto, la eficacia del sistema obedece al Criterio de Localización escogido y que este a su vez, proviene de un planteamiento teórico específico, por lo tanto, en una posible extensión de la presente investigación se podrían sugerir nuevos planteamientos teóricos para el procesamiento de las señales, dando origen a nuevos criterios de localización.

Es preciso replantear el fundamento teórico del criterio de localización, considerando que, en un ambiente comúnmente acompañado de ruido, el sistema de localización humano está en capacidad de identificar ciertas fuentes, incluso, le es posible hacerlo de manera simultánea para diferentes clases de estímulos, como ocurre en el momento de escuchar una pieza musical mezclada en estéreo en un sistema de audio para tal fin. En una situación así, el oyente puede ubicar espacialmente algunos o la totalidad de los instrumentos sin tener que escuchar la pieza instrumento por instrumento.

El fenómeno de localización, aunque en un planteamiento inicial se ajusta a los criterios de ILD e ITD, en el fondo reviste una mayor complejidad de la que se pudo observar por parte de la presente investigación. Es preciso, como se hizo en el caso de la investigación publicada por AES, tener otras referencias y modelos de comparación, ya que trabajar con una única unidad de captura, provista de un único juego de micrófonos, reduce la posibilidad de identificar variables importantes dentro del proceso.

7 RECOMENDACIONES

La principal recomendación es estudiar la posibilidad de ajustar la estructura de la cabeza binaural, de manera que puedan acoplarse en su interior otro tipo de cápsulas, haciendo posible desarrollar ejercicios que arrojen nuevos resultados. De este modo se reduciría de manera importante el grado de dependencia que en la actualidad existe de los micrófonos AKG 480 B.

Dentro de esta misma propuesta, se plantea la posibilidad de ajustar el diseño interno de la cabeza para que además se pueda elaborar, de alguna forma, el conducto auditivo, el cual se ha comprobado que aporta un filtrado particular en el caso de la audición humana.

Adicionalmente, se recomienda adelantar los procesos concernientes a la generación de la Aplicación en su extensión ejecutable .exe, para que pueda llevarse a equipos que no tengan instalado Matlab para que pueda ser objeto de observación y futuras mejoras.

Se recomienda la implementación de otros criterios de localización, que permitan considerar nuevas rutas y alternativas para llegar a un sistema que pueda acercarse de una forma más efectiva al desempeño del sistema de localización humano. Las posibilidades que ofrecería un sistema de localización con estas cualidades son muchas en diferentes áreas de investigación.

Se pueden implementar cientos de ejercicios y técnicas para estudiar el fenómeno de localización, pero ello implica una inversión de tiempo y dinero que se puede reducir apoyándose en las investigaciones existentes, no obstante, estas publicaciones en muchos de los casos están restringidas como es el caso de las publicaciones de AES.

BIBLIOGRAFIA

OPPENHEIM Alan S., WILLSKY Alan S Señales y Sistemas,. Editorial Pearson Educación, Segunda Edición. 1997

PROAKIS John G., MANOLAKIS Dimitris G Tratamiento Digital De Señales, Principios, algoritmos y aplicaciones.. Editorial Prentice Hall, Tercera edición. 1998

ROBERTS Michael.J. Señales y Sistemas, Análisis mediante método de transformada y matlab. Editorial Mac Graw Hill, primera edición. 2005.

McCARTHY Bob. Sound Systems: Design and Optimization: Modern Techniques and Tools for Sound System Design and Alignment.

Referencias en Internet

GARCIA Alejandro, Sistema de Sonido 3D, Trabajo de electroacústica teórico práctico. Universidad del país vasco y Escuela Superior de Ingenieros, Bilbao, Junio de 2005

ORDOYO ANSORENA, Fabiola; Estudio comparativo de las técnicas estéreo y la técnica binaural en la grabación de ruido interior de vehículos. Ingeniería La Salle, Universidad Ramon Llull. REFERENCIA PACS: 43.38.Md. Barcelona. España

PAULI MINNAAR, *AES Member*, SOREN KRARUP OLESEN, FLEMMING CHRISTENSEN, *AES Member*, AND HENRIK M-ILLER, *AES Member*. Localization with Binaural Recordings from Artificial and Human Heads. Department of Acoustics, Aalborg University, DK-9220 Aalborg, Denmark. J Audio Eng. Soc, Vol 49, No 5, 2001 May.

Localización espacial de sonido

<http://www.sacom.org.ar/secciones/segunda/Sesiones%20Tem%C3%A1ticas/Di%20Liscia.htm>

Audio virtual 3D para auriculares

<http://www.sensaura.com/whitepapers/pdfs/devpc007.pdf>

GLOSARIO

Convolución: Procedimiento matemático mediante el cual se calcula la respuesta de un sistema ante una señal de entrada arbitraria, teniendo como referencia la respuesta del mismo sistema ante un impulso $h(t)$.

Correlación: Procedimiento matemático que permite determinar el grado de similitud entre dos señales. Se calcula de manera semejante a la correlación, pero omitiendo la inversión en el eje temporal de las señales a comparar.

Correlación interaural cruzada: Parámetro que mide la diferencia en los tiempos de llegada a los oídos izquierdo y derecho ante una excitación acústica.

Fracción de energía lateral: Parámetro que relaciona la cantidad de energía acústica que llega a la cabeza del receptor desde las paredes laterales con la energía total que proviene de todas las demás direcciones.

Ventana inicial de retardo (ITDG): Tiempo que transcurre entre la llegada del sonido directo y de la primera reflexión significativa al receptor.

Diferencias interaurales de tiempo (ITD): Diferencia del tiempo de arribo de una señal a cada uno de los oídos.

Diferencias interaurales de nivel (ILD): Diferencia en el nivel de presión sonora percibido por los oídos ante una misma señal acústica.

Cabeza Binaural: Dispositivo que en su forma se asemeja a la cabeza humana y que cuenta con dos micrófonos de iguales características, dispuestos en las regiones laterales (como los oídos) para simular de manera artificial el tipo de captación binaural humano.

GUI: (Graphic Unit Interface) Unidad de generación de herramientas gráficas para programación en matlab. Es un recurso que permite diseñar ventanas de dialogo para controlar recursos de matlab según las necesidades del programador.