

PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ
FACULTAD DE CIENCIAS E INGENIERÍA



PONTIFICIA
UNIVERSIDAD
CATÓLICA
DEL PERÚ

**Implementación de un algoritmo para la localización de la voz
de un orador en un plano tomando como punto de
referencia el nivel máximo de su patrón de radiación**

Tesis para optar el Título de Ingeniero Electrónico, que presenta el bachiller:

David Fernando Pinto Valdiviezo

ASESOR: Jorge Néstor Moreno Ruiz

Lima, Agosto del 2011

RESUMEN

Actualmente se vive en un mundo competitivo donde el conocimiento de las nuevas tecnologías y tendencias se ha convertido en la herramienta más poderosa. Las personas acuden más frecuentemente a todo tipo de conferencias y exposiciones para enterarse de los últimos avances e investigaciones en los campos que a cada uno le conciernen.

Este tipo de eventos tienen lugar dentro de salas de conferencias o auditorios los cuales han sido diseñados para brindar la mejor infraestructura y acústica posible para el público usuario. Sin embargo no se ha logrado controlar los altibajos de niveles de presión sonora del micrófono del expositor producido por los movimientos naturales de su cabeza, lo cual genera problemas de inteligibilidad y distracción en el público. El objetivo de la presente tesis es localizar la dirección de mayor intensidad de la voz de un orador respecto a un punto como parte de un proyecto que pretende ecualizar un micrófono para lograr en la medida de lo posible un mismo nivel y grado de inteligibilidad durante una exposición hablada que solucione el problema antes mencionado.

En el primer capítulo, se muestra información teórica sobre la voz humana y los micrófonos, puesto que es sumamente importante conocer la fuente y el receptor de señales y cómo es que interaccionan entre sí.

En el segundo capítulo, se describe el marco problemático y los posibles métodos de localización para resolver el problema, además se encuentra la investigación realizada de los últimos avances en localización sonora.

En el tercer capítulo, se define el método a emplear y el diseño del algoritmo que toma en cuenta los requerimientos del método y las características de la voz humana.

En el cuarto capítulo se muestran los resultados de la implementación del diseño basándose en las pruebas realizadas y el costo del proyecto.

Finalmente se presentan las conclusiones, las cuales muestran que se cumplieron con los objetivos planteados y se dan algunas recomendaciones para futuras investigaciones sobre el proyecto.





Agradecimiento

La culminación de la tesis es el fin de una etapa muy importante en la vida de cualquier persona, es el producto de todos los largos años de la vida universitaria que estuvo llena de buenos y malos momentos durante los cuales uno aprende a crecer profesional y personalmente.

Sin embargo finalizar esta etapa de mi vida no hubiese sido posible sin el apoyo incondicional de mis padres Luis y Nerisa, quienes fueron los promotores y animadores de esta etapa tan crucial, a ellos mi eterno agradecimiento por todo lo que hicieron y me inculcaron durante todo este tiempo. A mi hermano Alejandro, quien fue mi apoyo en esos momentos donde todo no parecía funcionar, le dedico mi gratitud.

Por otro lado a mi asesor Jorge Moreno y al profesor Richard Rivera quienes me apoyaron siempre con consejos y la disposición de su tiempo para poder sacar adelante este proyecto, se los agradezco.

Finalmente a Dios quien me dio los dones para poder seguir este camino.

Gracias

Índice

	Pág.
Carátula	1
Resumen	2
Tema de tesis	3
Agradecimientos	5
Índice	6
Título	8
Introducción	8
Capítulo 1: “Fuente y receptor de sonido: la voz y el micrófono”	
1.1 La voz humana	9
1.2 Micrófonos	12
1.2.1 Micrófono cardioide	12
1.2.2 Respuesta en frecuencia de los micrófonos	13
Capítulo 2: “Métodos de localización de fuentes sonoras”	
2.1 Marco Problemático	16
2.2 Técnicas de localización sonora	17
2.2.1 Intensidad Sonora	17
2.2.2 Beamforming	18
2.2.3 Holografía de campos acústicos cercanos	19
2.3 Estado del Arte	20
2.3.1 Presentación del asunto de estudio	20
2.3.2 Estado de la investigación	21
2.3.3 Síntesis sobre el asunto de estudio	22
Capítulo 3: “Planteamientos para la implementación de un algoritmo de localización para la mayor intensidad de la voz”	
3.1 Objetivos	23
3.1.1 Objetivo general	23
3.1.2 Objetivos específicos	23
3.2 Elección de la técnica de la localización sonora	23

3.3 Diseño de la sonda de intensidad sonora	25
3.3.1 Cálculo de la distancia de separación entre los micrófonos	26
3.3.2 Corrección de fase	27
3.3.3 Corrección de magnitud	34
3.4 Medición de intensidad sonora	40
3.5 Algoritmo para la localización de una fuente sonora tomando como referencia su mayor nivel de intensidad	41

Capítulo 4: “Pruebas y resultados del diseño”

4.1 Pruebas	45
4.2 Resultados	48
4.3 Presupuesto	50
Conclusiones	51
Recomendaciones	52
Bibliografía	53



Implementación de un algoritmo para la localización de la dirección de mayor intensidad de voz de un orador

Introducción

Un objeto puede ser localizado por medio de la visión, sin embargo existen otras formas de localización. El uso de ondas electromagnéticas permite a los radares determinar la posición de un objeto. Del mismo modo las ondas sonoras permiten a los sonares realizar una localización acústica de los objetos a su alrededor.

La localización de objetos por medio del sonido no sólo es bien aprovechado por los sonares. El “beamforming” técnica usada por los sonares es un método de localización de fuentes de sonido por medio de diferenciación de niveles de intensidad del sonido basado en la dirección de donde se originan.

Entre las aplicaciones de esta técnica se encuentran la localización de fuentes de ruido en objetos de mediano a gran tamaño como vehículos, componentes y equipos de construcción. Mediciones remotas en ambientes donde es difícil o peligroso medir cerca de la fuente.

Las aplicaciones de la localización de fuentes de sonido inclusive han sido utilizadas para mejorar la calidad de las videoconferencias, haciendo uso de la posición de la fuente sonora (la boca del expositor) para realizar el control de la cámara de video. Esto se realiza para compensar la incomodidad que sienten los participantes cuando el expositor no es enfocado por la cámara.

De la misma manera en una conferencia, mientras un expositor está hablando resulta incómodo para el público en general cuando se producen diferencias en el nivel de intensidad de la voz debido a los movimientos naturales del expositor frente al micrófono o porque voltea a leer su ponencia.

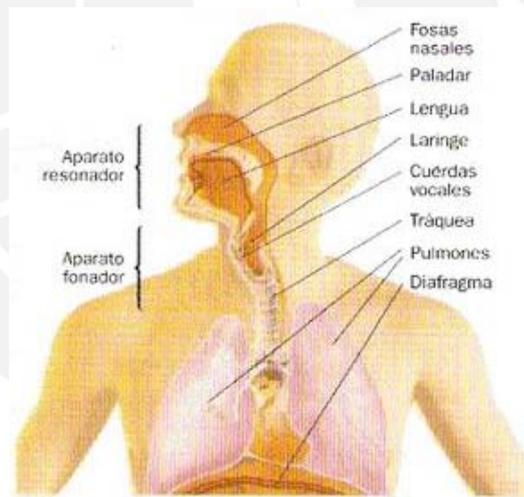
La presente tesis es parte de un proyecto que pretende contribuir con la ecualización automática de un micrófono durante la oratoria. Para ello plantea el uso de un sistema de micrófonos que permita localizar la dirección de la mayor intensidad de voz producida por un orador en un ángulo sólido con respecto al micrófono por el cual está hablando. La implementación de todo el sistema permitirá un mayor realismo y perfección de sistemas de audio en salas de oratoria.

Capítulo 1: “Fuente y receptor de sonido: la voz y el micrófono”

1.1 La voz humana

La presente tesis intenta en la medida de lo posible localizar la dirección de mayor intensidad de la voz, sin embargo para poder lograr esta localización es necesario conocer las características de la fuente de sonido.

La voz humana es producida en la laringe, donde el verdadero órgano de fonación humano es la glotis. La fonación se realiza cuando el aire procedente de los pulmones es forzado, por el diafragma y los intercostales, durante la respiración a través de la glotis y haciendo vibrar los dos pares de cuerdas vocales. Este sonido es muy débil por lo cual es amplificado por los resonadores nasal, bucal y faríngeo. Aquí también se produce el aumento de frecuencia y desvalorización del sonido. La figura 1 muestra el sistema vocal con los órganos que lo componen, mientras que la figura 2 muestra el modelamiento del mismo.



Fuente: <http://vozyaudicion.blogspot.com/>

Figura 1
Sistema vocal

Sin embargo a comparación del rango completo de la audición humana, la voz cubre un rango relativamente estrecho de frecuencias (entre 100Hz hasta 6kHz). Dentro de este rango, la distribución de la potencia se encuentra mayormente localizada por debajo de 1kHz, inclusive con el 80% de la energía concentrada por debajo de los 500Hz. Precisamente las bajas frecuencias corresponden a las vocales y su supresión resta potencia a la voz.

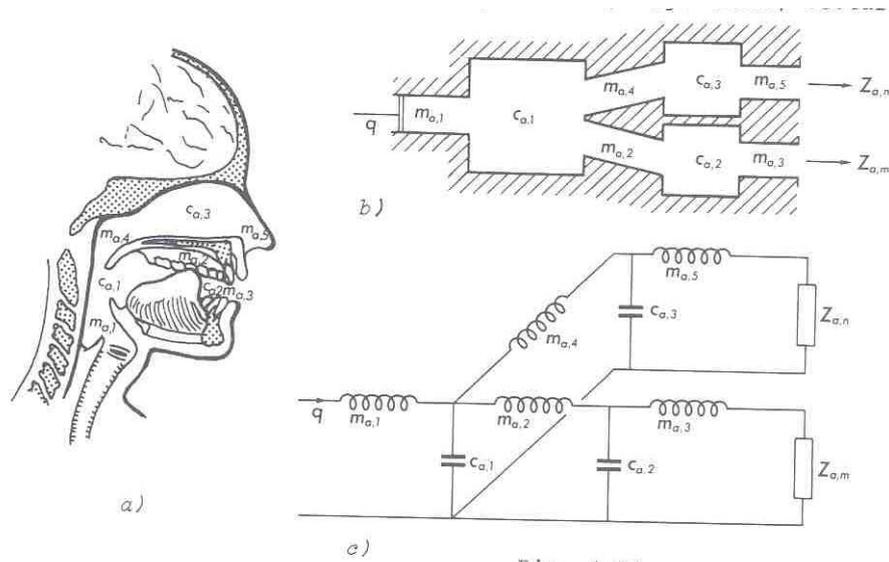


Fig. 4.31

Fuente: Analogier mellen MEKANISKE, AKUSTISKE OG ELECTRISKE SYSTEMER. K. Rasmussen

Figura 2

Sistema equivalente al sistema vocal

A pesar que la voz posee muy poco contenido de alta frecuencia, la inteligibilidad oral se debe a estas frecuencias, donde casi toda la energía de la voz por encima de 1kHz es producida por las consonantes. Para que el habla sea comprensible es indispensable la presencia de armónicos cuyo rango de frecuencias está entre los 500Hz a 3500Hz. Por lo tanto, la pérdida de altas frecuencias puede afectar drásticamente la inteligibilidad del habla.

Otra característica de la voz humana es su patrón de radiación donde se observa la distribución espacial del nivel de presión sonora, en función del ángulo θ , alrededor de la boca que es el centro del plano. La figura 3 muestra los patrones de radiación de un Simulador de cabeza y torso Tipo 4128C de Brüel & Kjær para diferentes frecuencias, los datos fueron obtenidos por medio de mediciones en el laboratorio de acústica de la Pontificia Universidad Católica del Perú.

Como se observa en las figuras 3a y 3b, los patrones de radiación en 250Hz y 500Hz son muy parecidos, en ellos se puede diferenciar 3 zonas en particular: la primera entre 0° y 30° donde la magnitud es casi idéntica, la segunda entre 30° a 60° donde existe una caída progresiva de hasta 2dB con respecto a la primera zona y la tercera entre 60° y 90° donde la caída va entre 2dB a 4dB.

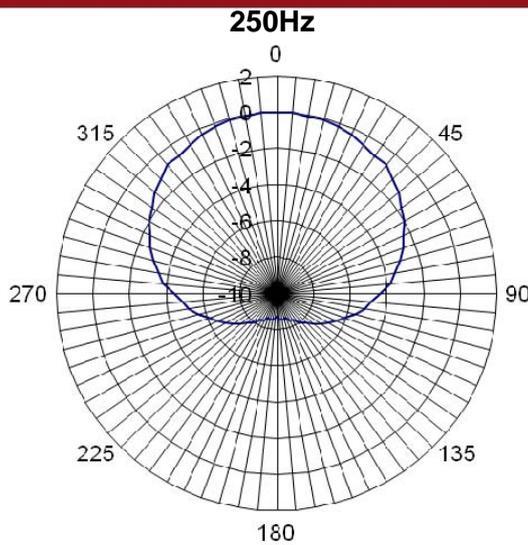


Figura 3a

Patrón de radiación a 250Hz

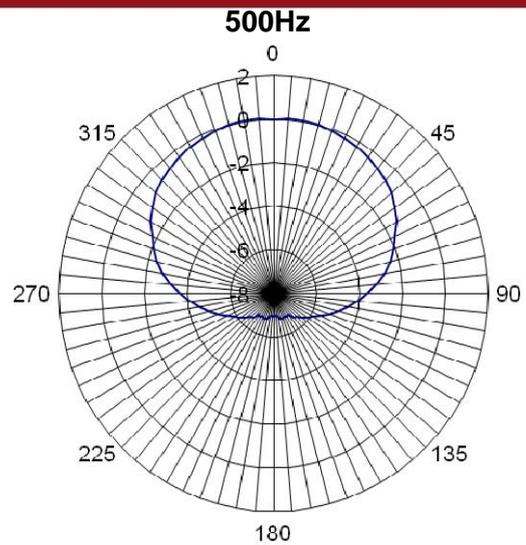


Figura 3b

Patrón de radiación a 500Hz

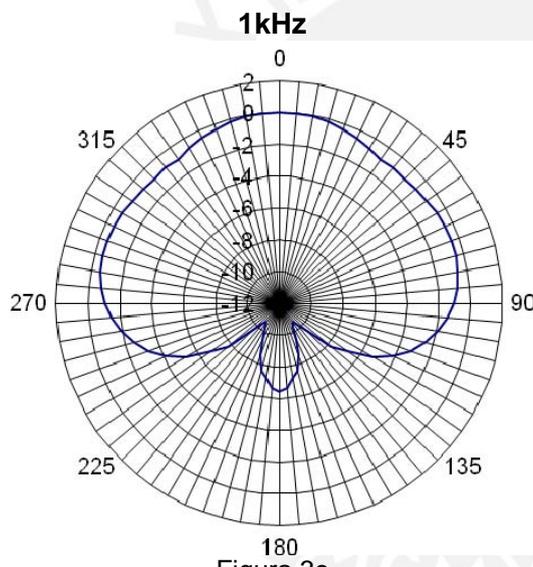


Figura 3c

Patrón de radiación a 1kHz

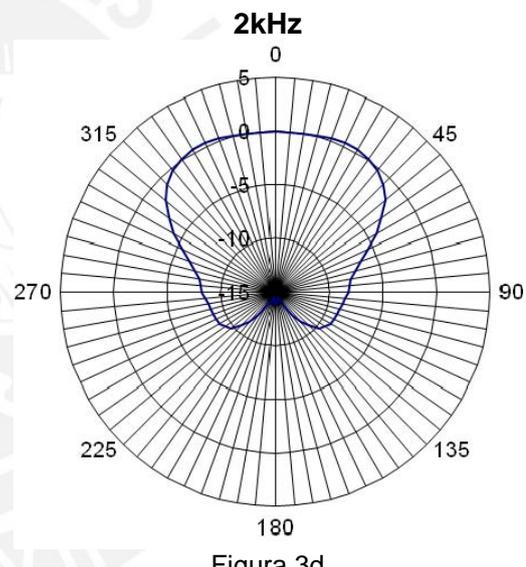


Figura 3d

Patrón de radiación a 2kHz

Fuente: creación propia

Por otro lado la figura 3c muestra el patrón de radiación en 1kHz el cual conserva su magnitud prácticamente en todo el rango de 0 a 90°, lo cual no es muy útil si se desea hallar el movimiento de la cabeza en esta frecuencia puesto que no se observa ninguna variación en la magnitud; mientras que la figura 3d muestra el patrón a 2kHz el cual es bastante direccional ya que las caídas de las magnitudes son bien pronunciadas a partir de los ángulos de 40° y 65° principalmente. Finalmente se deben resaltar dos características de los patrones de radiación: son simétricos respecto al eje de 0° y dentro de todo su rango de frecuencias las magnitudes siempre se conservan en los ángulos de 0° a 30°.

1.2 Micrófonos

Por otra parte, si bien es necesario conocer la fuente de sonido, también es necesario conocer los instrumentos con los cuales se mide la voz, y el instrumento por excelencia para ello es el micrófono.

El micrófono es el término genérico que se utiliza para denominar a los elementos que transforman energía acústica (sonido) en energía eléctrica. Un micrófono por ende es un tipo de transductor, dispositivos que traducen energía de una forma a otra forma de energía. Una clase particular de micrófonos puede incluir alguno de los siguientes transductores: carbón, cerámico, condensador, dinámico, inductancia, magnético, electrónico, semiconductor, entre otros.

Dentro de la gama de micrófonos existen varias formas para la clasificación de estos: micrófonos de presión, de gradiente de presión (velocidad), combinación de presión y gradiente de presión y de ondas de interferencia.

La respuesta eléctrica de los micrófonos de presión resulta de las variaciones de presión en el aire, estos tienen un patrón de recepción omnidireccional, puesto que la presión es una cantidad escalar que posee magnitud mas no dirección.

La respuesta eléctrica de los micrófonos de gradiente de presión resulta de la variación de la velocidad de partículas en el aire. El patrón de recepción es bidireccional (la figura de un ocho) esto se debe a que la velocidad de partículas es un vector, posee magnitud y dirección.

Por otra parte la respuesta eléctrica de los micrófonos que combinan presión con gradiente de presión también es proporcional a las variaciones de la velocidad de partículas. El patrón de recepción resulta ser cardioide, supercardioide o similar a una función con forma de limón, que puede ser ajustable o corregida.

1.2.1 Micrófono Cardioide

Debido a la popularidad y uso frecuente del micrófono tipo cardioide en todo tipo de eventos, como son los congresos y exposiciones, es necesario conocer más sobre ellos. La figura 4 muestra el patrón de recepción del micrófono el cual tiene forma de corazón, es de allí de donde proviene su nombre. Como se observa en esa

figura, este tipo de micrófonos son más sensibles a los sonidos provenientes del eje frontal y rechaza los sonidos de los costados y posteriores a él.



Fuente: <http://www.bateristas.net/lecciones/micros.htm>

Figura 4

Patrón de Recepción de un micrófono Cardioide

Las cualidades direccionales lo hacen una elección natural para refuerzo sonoro, debido a que ayudan a reducir la realimentación e incrementar la ganancia del sistema. Este efecto es muchas veces sobreestimado, y los micrófonos omnidireccionales son frecuentemente una mejor opción para trabajos cercanos que los cardioides.

Este tipo de micrófonos tienden a tener una mayor coloración cuando el sonido no proviene del eje debido a sus cualidades direccionales que varían con la frecuencia. Los micrófonos cardioides son bastante comunes en grabación, debido a que pueden ser usados para disminuir sonidos no deseados provenientes fuera del eje. A comparación de los micrófonos omnidireccionales su respuesta en frecuencia es usualmente mucho más tosca y son de alguna manera más sensibles al ruido del viento y de la respiración.

1.2.2 Respuesta en Frecuencia de los micrófonos

La respuesta en frecuencia es la relación entre la entrada del dispositivo y la salida tomando en consideración la amplitud y la frecuencia de la señal. Su uso más común es describir el rango utilizable de las frecuencias que pasan desde la entrada hasta la salida. Para un micrófono viene a ser la medición de la consistencia con la cual se traduce un nivel de presión sonora en un nivel de señal de audio a distintas frecuencias.

Idealmente un micrófono traduciría un nivel de presión determinado a un mismo nivel de señal sin importar la frecuencia de ésta. Este tipo de micrófonos ideales poseen lo que se denomina una respuesta plana en la frecuencia.

Mientras que muchos de los micrófonos de grabación e instrumentación se acercan a esta idea, la mayoría de las unidades usadas en audio profesional se desvían de esta respuesta plana. Sin embargo, las variaciones de la respuesta en frecuencia que muchas veces se dan en un micrófono no son necesariamente un defecto. Son muchos casos en donde estas variaciones son introducidas intencionalmente en busca de lograr ventajas específicas en la performance para aplicaciones específicas.

Una de las limitaciones frecuentes de los micrófonos es la caída de la respuesta en frecuencia por debajo de los 100Hz. Un ejemplo de esto son los micrófonos para oratoria, esto debido a que el espectro de frecuencia de la voz humana no contiene energía por debajo de los 100Hz.

Por otra parte en las altas frecuencias muchos micrófonos exhiben una ganancia, la cual se denomina pico presencial, y esta es una característica típica para los micrófonos de canto no lírico. Un pico presencial ayuda a incrementar la inteligibilidad de las palabras y la presencia de la voz en actuaciones en vivo, por lo que puede ser una característica deseable. Pero al mismo tiempo puede incrementar la posibilidad de una realimentación en el refuerzo sonoro.

Otro efecto que se presenta en la respuesta en frecuencia es el denominado efecto de proximidad. El cual es un incremento en la respuesta de la baja frecuencia cuando el micrófono se encuentra muy cerca de la fuente de sonido, esta es una característica inherente de los micrófonos direccionales (los omnidireccionales no exhiben este efecto).

Este efecto es incrementado drásticamente cuando el micrófono se encuentra a menos de 60cm de distancia de la fuente y puede producir un incremento de los bajos de 16dB o más.

Algunas veces esto puede producir una saturación en el pre-amplificador resultando en una distorsión de los graves. Tanto anunciadores y vocalistas frecuentemente emplean el efecto de proximidad para agregar consistencia al sonido de su voz, e

inclusive un artista experimentado lo incorpora como parte de su técnica de canto. El caso contrario ocurre con oradores inexpertos a quienes este efecto normalmente perjudica pues disminuye la inteligibilidad de la palabra.

Las variaciones en la respuesta en frecuencia son un factor primordial que gobierna las características sonoras de un micrófono y es necesario conocerlas al momento de la selección de uno para una aplicación específica.



Capítulo 2: “Métodos de localización de fuentes sonoras”

2.1 Marco Problemático

En la actualidad se vive cada vez más en un mundo competitivo donde el conocimiento se ha vuelto una herramienta de suma importancia para todas las personas. Es en búsqueda del conocimiento que las personas asisten a seminarios, conferencias y/o congresos de distintas índoles dependiendo de sus intereses.

Este tipo de eventos requiere de establecimientos diseñados para tal propósito, los cuales deberán cumplir tanto en la capacidad requerida como también en brindar la atmósfera adecuada para la atención de los participantes.

Una buena infraestructura propicia un ambiente que ayuda a la atención del público. Este ambiente es un factor decisivo en la elección de un establecimiento respecto de otro para lograr una mejor calidad del evento.

Bajo estas circunstancias los asistentes y expositores son los encargados de evaluar el desarrollo de los eventos. Uno de los problemas que se presentan es no lograr la atención deseada por cada uno de los participantes debido a que estos no escuchan con un grado de inteligibilidad las palabras del expositor. La infraestructura del local o el movimiento natural del expositor frente al micrófono podrían ser las causas de este problema.

Consecuente del movimiento del orador se producen variaciones en los niveles de presión sonora de las ondas que llegan al micrófono. El cual no cuenta con ningún tipo de compensación para este tipo de cambios, lo que produce altibajos en los niveles de sonido que posteriormente son amplificadas. Finalmente ello provoca incomodidad en el público y molestia en el expositor.

Se ha logrado controlar la infraestructura del local para las exposiciones mas no la otra causa del problema. Los movimientos naturales de un expositor frente al micrófono no son predecibles, lo que conlleva a que la ecualización establecida previamente no sea la más adecuada. Lo que se requiere es una ecualización constante y una compensación de ganancias del micrófono del expositor dependiendo de la posición en la cual se encuentre para lograr tener en la medida de lo posible un mismo nivel de intensidad sonora durante la exposición hablada.

Por lo que el problema se divide en dos partes: la primera es diseñar un sistema que permita determinar la dirección de mayor intensidad de la voz de un orador, mientras que la segunda es lograr en la medida de lo posible un mismo nivel y grado de inteligibilidad durante una exposición hablada con la ayuda de filtros adaptativos gracias a la información de la posición del orador, esta última etapa escapa a los alcances de la presente tesis.

2.2 Técnicas de localización sonora

Dentro del mundo de la localización de objetos por medio del sonido, existen muchas técnicas que son aprovechadas por las industrias para realizar distintos ensayos, entre sus aplicaciones se encuentran: localización de fuentes de ruido, mapeo sonoro, reconstrucción de campos acústicos, etc.

Cada una de estas técnicas tienen sus ventajas y desventajas una respecto a las otras por lo cual son utilizadas según la aplicación. Estas técnicas son las siguientes:

- Intensidad Sonora
- Beamforming
- Holografía de campos acústicos cercanos (NAH)

2.2.1 Intensidad Sonora

La intensidad sonora es un vector que describe la cantidad y la dirección del flujo de energía acústica en una posición determinada, su unidad es W/m^2 . En un medio sin un flujo considerable, la intensidad sonora es igual al promedio en el tiempo del producto de la velocidad de partículas instantánea y la presión instantánea en la misma posición (ecuación (2.1))

$$I = \overline{p \cdot u} \quad (2.1)$$

donde:

I : Intensidad sonora

p : Presión sonora

u : Velocidad de partículas

Para poder calcular la intensidad sonora es necesario medir la presión instantánea y la velocidad de partículas, lo primero se puede realizar con cualquier tipo de

micrófono, sin embargo medir la velocidad de partículas directamente no es sencillo. Es por este motivo que se asocia la velocidad de partículas a un gradiente de presión con la ecuación linealizada de Euler (2.2), la cual es la segunda ley de Newton aplicada a fluidos. El apéndice A explica cómo calcular intensidad sonora a partir de las ecuaciones (2.1) y (2.2).

$$a = -\frac{1}{\rho} \text{grad } p \quad (2.2)$$

donde:

a : Aceleración

ρ : Densidad del fluido

p : Presión sonora

La medición de la intensidad sonora ha hecho un cambio radical en la industria. Anteriormente solamente se podía medir presión sonora la cual depende del campo sonoro. La potencia sonora puede ser relacionada con la presión sonora únicamente bajo condiciones cuidadosamente controladas donde se realizaban suposiciones acerca del campo sonoro (cámara anecoica y cámara reverberante).

Sin embargo, con la medición de intensidad sonora las mediciones pueden ser realizadas en cualquier campo sonoro siempre y cuando este no sea muy reactivo. Esta técnica permite realizar mediciones in situ de la intensidad y potencia acústica emitida por equipos aun cuando otras fuentes estén emitiendo ruido, inclusive a niveles más altos que la fuente que se desea medir. Esto es posible siempre y cuando el ruido sea estacionario, ya que el flujo neto a través de la superficie que se define para determinar la intensidad sonora es cero.

Una de las ventajas que proporciona la medición de la intensidad acústica es que provee una medición de la dirección así como de la magnitud, lo cual es muy útil cuando se trata de localizar fuentes de sonido.

2.2.2 Beamforming

El beamforming es una técnica de procesamiento de señales para la transmisión o recepción de señales direccionales. La selectividad espacial es lograda por medio de rayos patrones adaptativos. La mejora comparada con una recepción o

transmisión omnidireccional es conocida como la ganancia de recepción o transmisión respectivamente.

La técnica del beamforming es usada tanto en ondas electromagnéticas como acústicas. Posee variadas aplicaciones en radares, sonares, sismología, radio astronomía y biomedicina.

Esta técnica toma ventaja de la interferencia para cambiar la direccionabilidad del arreglo. Cuando se recibe información de distintos sensores es combinada de tal manera que el patrón esperado es preferentemente al observable. Por otra parte cuando se transmite información el beamformer controla la fase y la amplitud relativa de la señal de cada transmisor, para crear un patrón constructivo y destructivo de interferencia en el frente de onda.

El beamforming puede ser dividido ampliamente en dos categorías:

- Beamforming convencional
- Beamforming adaptativo.

Los beamformers convencionales utilizan un conjunto de coeficientes y retrasos de tiempo para combinar las señales provenientes del arreglo de sensores, primordialmente haciendo uso único de la información de la ubicación de los sensores en el espacio y de las direcciones de onda de interés. Por el contrario el beamforming adaptativo generalmente combina esta información con las señales que realmente son recibidas por los sensores, típicamente para mejorar el rechazo de señales no deseadas de otras direcciones.

2.2.3 Holografía de campos acústicos cercanos

La holografía de campos acústicos cercanos conocida por sus siglas en inglés NAH (Nearfield Acoustic Holography) es una poderosa técnica para la localización eficiente y precisa de fuentes de sonido, además analiza los campos sonoros cercanos a la fuente y deduce importante información acerca de la naturaleza de la fuente sonora. En la holografía de campos acústicos cercanos el campo sonoro es reconstruido a partir de las mediciones de posiciones discretas en una región de un plano finito. Generalmente la presión sonora es la que se mide, pero en principio

existe la misma información en la componente normal de la velocidad de las partículas acústicas.

El NAH debe contemplar dos aspectos importantes al momento de la medición de los datos. El primero es el espaciamiento entre los micrófonos de la grilla de medición, esta distancia determinará la mitad de la longitud de onda de la frecuencia máxima de medición, mientras que el segundo aspecto es el ancho del arreglo, éste depende del tamaño del objeto y de la mínima frecuencia de interés. Si el objeto es muy grande respecto al tamaño del arreglo, pueden realizarse mediciones por partes, hasta que toda la superficie del objeto sea cubierta, esto sólo se puede realizar si el objeto es estacionario. De no ser el caso, por lo menos debería de ser una condición transitoria repetitiva.

Una importante diferencia entre el beamforming y la NAH es que el beamforming puede ser utilizado con geometrías de arreglos irregulares, por ejemplo geometrías aleatorias de arreglos.

2.3 Estado del Arte

2.3.1. Presentación del asunto de estudio

En la actualidad la tecnología de localización de sonido tiene diversas aplicaciones, entre las cuales cabe mencionar: localización de fuentes de ruido, mapeo de fuentes sonoras, sistemas de adquisición de voz, entre otras. Existen diversos métodos y formas para la localización del sonido. El beamforming y la holografía acústica de campos cercanos son algunos de ellos.

El beamforming es un método que realiza mapeo de fuentes de sonido por medio de diferenciación de niveles de intensidad sonora, los cuales se basan en la dirección de donde se originan. Este método es bastante rápido puesto que permite calcular un mapa completo con una sola medición. [1]

Por otro lado la holografía acústica de campos cercanos (NAH) es el método que permite la reconstrucción de campos acústicos en 3D a partir de la medición de parámetros acústicos a través de un arreglo de transductores de presión y/o velocidad de partículas. [2]

2.3.2. Estado de la investigación

La compañía Brüel & Kjær ha diseñado un beamforming esférico que permite hacer un mapa en 360 grados del campo sonoro en ambientes confinados como automóviles o cabinas de aviones sin hacer suposiciones acerca del campo sonoro. Es posible el uso del beamforming esférico para hacer un mapeo rápido del campo sonoro y realizar análisis más profundos utilizando la aplicación de Mapeo que pueden ser ejecutados en el mismo sistema.

La adquisición de datos se realiza localizando la esfera en el centro del área de medición. La esfera tiene micrófonos y cámaras distribuidas en la superficie permitiendo realizar una medición sin conocer la geometría del objeto a medir. [1]

El beamforming provee una resolución muy pobre a bajas frecuencias pero una buena resolución a frecuencias altas con relativamente pocos puntos de medición. Por el contrario, la holografía acústica de campos cercanos provee una poderosa técnica de localización de fuentes de sonido en bajas y medianas frecuencias con una alta resolución, sin embargo con muchas posiciones de medición a altas frecuencias.

Por lo cual se ha desarrollado una técnica de medición que combina el NAH a bajas frecuencias y el beamforming a altas frecuencias por lo que provee lo mejor de ambas técnicas. Esto es posible debido a un sistema que permite mediciones de holografía acústica de campos cercanos y beamforming realizadas con un mismo arreglo irregular. [3]

Por otra parte esta misma compañía emplea los métodos de localización de sonido para la identificación de fuentes de ruido, conocido por sus siglas en inglés “NSI” (Noise Source Identification). Este método ayuda a optimizar las emisiones de ruido para una gama grande de productos desde automóviles, herramientas de potencia y maquinaria pesada hasta componentes como motores, cajas de cambios, electrodomésticos, etc.

El objetivo del NSI es identificar las sub-fuentes en un objeto en términos de posición, frecuencia y radiación sonora. La clasificación de estas sub-fuentes pueden ser usadas para identificar donde los cambios de diseño afectarían en mayor medida a los niveles de radiación de ruido. [4]

Por otro lado los arreglos de micrófonos poseen una variedad de usos en sistemas de adquisición de voz. Las aplicaciones incluyen teleconferencias, reconocimiento de voz y captura de sonido en ambientes adversos.

2.3.3. Síntesis sobre el asunto de estudio

De acuerdo a los métodos y aplicaciones de la localización de fuentes sonoras vistas anteriormente se deduce la siguiente síntesis:

- El uso de localización de fuentes sonoras es bastante utilizado por las industrias para diferentes aplicaciones.
- Existen diferentes métodos para la localización de fuentes de sonido. Entre ellos el Beamforming y el NAH.
- Entre las aplicaciones de localización de fuentes sonoras están localización de fuentes de ruido, mapeo de fuentes sonoras.
- Todas estas técnicas son mayormente utilizadas en la industria automotriz y aeronáutica.

Capítulo 3: “Planteamientos para la implementación de un algoritmo de localización para la mayor intensidad de la voz”

3.1 Objetivos

3.1.1 Objetivo general

Determinar la dirección de mayor intensidad de la voz de una persona haciendo uso de un sistema de micrófonos y el correspondiente procesamiento de señal.

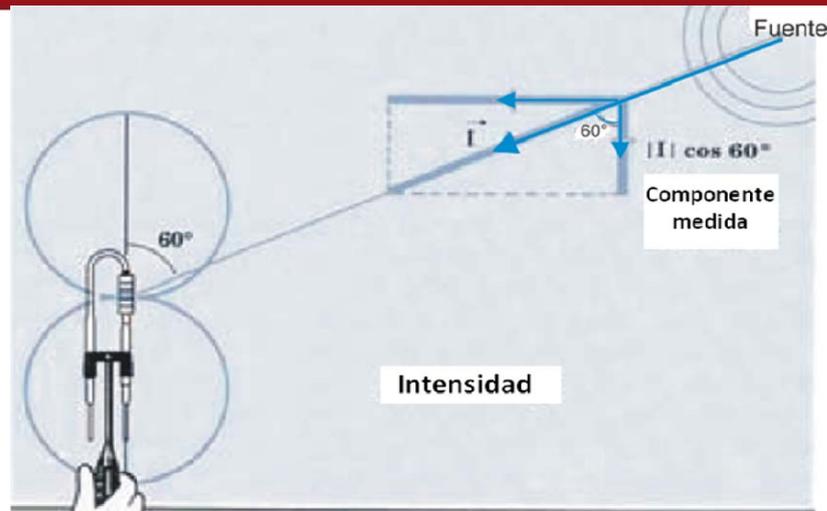
3.1.2 Objetivos específicos

- a) Determinar la cantidad y posición de los micrófonos a utilizar para obtener una precisión aceptable.
- b) Implementar el algoritmo adecuado basado en la información del arreglo de micrófonos para localizar la fuente de la voz respecto a un punto.

3.2 Elección de la técnica de localización sonora

Después de haber estudiado las tres técnicas de localización sonora (Intensidad sonora, beamforming y NAH) se ha optado por trabajar con la intensidad sonora para el desarrollo del algoritmo de localización. La principal razón por la cual se ha elegido este método para resolver la presente tesis es su simplicidad frente a los otros métodos. La intensidad sonora es una cantidad vectorial la cual es sensada por una sonda consistente de dos micrófonos de características dinámicas similares espaciados una distancia “d”, que mide la componente del vector intensidad paralelo al eje que contiene los micrófonos.

Por ejemplo, para un sonido incidente a 90° del eje de los micrófonos la diferencia de presiones sensada va a ser igual a cero, produciendo que el resultado de la intensidad sonora medida sea igual a cero. Por otro lado si el mismo punto es medido con un ángulo θ de incidencia respecto al eje de los micrófonos, el valor obtenido de intensidad sonora estará reducido por un factor $\cos\theta$ respecto al valor real, la figura 5 muestra esta característica.



Fuente: Primer: Sound Intensity de B&K

Figura 5

Característica de directividad de la intensidad sonora

Es esta reducción la que produce la característica de directividad de la intensidad sonora, la cual puede ser utilizada para hallar el ángulo de incidencia del sonido sobre la sonda siempre y cuando se conozca la magnitud real de la intensidad, esta magnitud se obtiene realizando una medición donde el ángulo de incidencia de la fuente sonora respecto a la sonda es 0° .

Finalmente el procedimiento para hallar el ángulo de incidencia sobre la sonda de intensidad es el siguiente:

1. Medición de la intensidad sonora con 0° de ángulo de incidencia entre la fuente sonora a analizar y la sonda de intensidad donde se obtiene la magnitud real de la intensidad (I).
2. Medición de la intensidad sonora con θ° de ángulo de incidencia entre la fuente sonora y la sonda de intensidad donde se obtiene una componente del vector de intensidad ($I \cdot \cos \theta$).
3. Tomar los valores de las mediciones 1 y 2 para realizar la ecuación (3.1) y hallar el ángulo de incidencia θ .

$$\theta = \text{ArcCos} \left(\frac{I \cdot \cos \theta}{I} \right) \quad (3.1)$$

Ecuación para determinar el ángulo de incidencia

Este método se ve limitado por tres factores. El primero implica que el primer paso del procedimiento que se denominará calibración del punto de referencia, debe ser realizado cada vez que la fuente sonora cambie, para el caso de la tesis implicaría realizarlo cada vez que un nuevo expositor utilice el sistema.

El segundo factor es el nivel de intensidad con el cual la persona habla, este es un factor relativo puesto que al momento de hablar uno tiende arbitrariamente a entonar o resaltar ciertas afirmaciones que pueden generar variaciones en la intensidad sonora, por lo cual se le pedirá al expositor que realice el proceso de calibración del punto de referencia con la misma entonación con la cual realizará su exposición, esto tiene por objetivo minimizar errores.

El último factor es el patrón de radiación de la persona, este es un factor muy importante y considerable puesto que va a delimitar el rango de ángulos que el sistema es capaz de diferenciar, por ejemplo en el capítulo 1 se observó que no existen diferencias en los niveles de presión sonora de una persona en el rango de 0° a 30° no importa la frecuencia que sea, por lo que diferenciar un ángulo exacto dentro de este rango basándose solamente en la voz es no es posible, a lo más que se puede llegar es indicar que se encuentra en una zona comprendida por estos ángulos.

Por otro lado, la medición de intensidad sonora se realiza por medio de una sonda de intensidad, la cual consiste básicamente en dos micrófonos colocados uno enfrente del otro o al costado del otro. Estos micrófonos deben cumplir con dos características muy importantes:

- La separación entre ellos debe estar dada por el rango de frecuencias que se desea medir.
- Deben ser exactamente iguales en su respuesta en frecuencia de magnitud y fase.

3.3 Diseño de la sonda de intensidad sonora

Para el diseño de la sonda se utilizaron dos micrófonos B&K 4189, pertenecientes al Laboratorio de Acústica de la Pontificia Universidad Católica del Perú, los cuales han sido adaptados en la medida de lo posible para que cumplan con las características mencionadas, lo cual es primordial si se desea medir intensidad sonora.

3.3.1 Cálculo de la distancia de separación entre los micrófonos

En primer lugar es necesario hacer los cálculos para la distancia de separación entre los dos micrófonos de tal manera que cumplan el rango de frecuencias deseado que es entre los 100Hz hasta 5kHz, se ha elegido este rango puesto que la mayor intensidad de la voz se encuentra en las bajas frecuencias correspondientes a las vocales donde prácticamente el 80% de la potencia de voz está por debajo de los 500Hz y es precisamente allí en donde se centra el estudio. Por otro lado el valor máximo de 5kHz se da porque es muy cercano al límite de la voz humana.

La distancia mínima de separación entre los dos micrófonos está limitada por la frecuencia máxima de la señal a analizar, en este caso por 5Khz. Los sensores deben cumplir las ecuaciones (3.2) y (3.3).

$$c = f_{\max} \cdot \lambda_{\min} \quad (3.2)$$

$$d_{\max} < \frac{\lambda_{\min}}{2} \quad (3.3)$$

Donde:

c = velocidad de la onda en el medio

f_{\max} = frecuencia máxima de la onda

λ_{\min} = mínima longitud de onda

d_{\max} = mínima separación de los sensores

La velocidad de la onda en el medio es la velocidad del sonido (340 m/s), mientras que la frecuencia máxima de la voz hablada es de 5kHz, reemplazando estos datos en la fórmula (3.2) se tiene:

$$c = f_{\max} \cdot \lambda_{\min} \quad (3.2)$$

$$340 = 5 \cdot 10^3 \cdot \lambda_{\min}$$

$$\lambda_{\min} = 0.068\text{m}$$

Luego, reemplazando en la fórmula (3.3) se tiene:

$$d_{\max} < \frac{\lambda_{\min}}{2} \quad (3.3)$$

$$d_{\max} < 3.4\text{cm}$$

Por los cálculos realizados la separación máxima de los sensores debe ser menor que 3.4cm.

3.3.2 Corrección de Fase

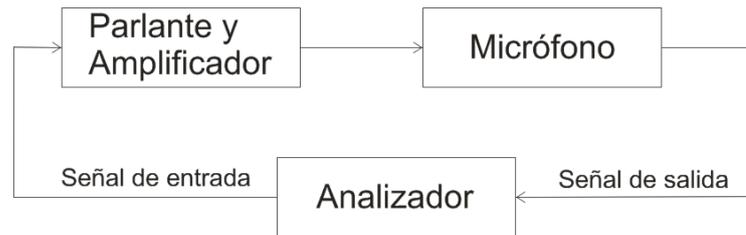
El segundo requerimiento al diseñar la sonda de intensidad sonora es que la diferencia de fases entre los dos micrófonos sea igual a cero.

La corrección de fase se realiza en dos etapas: la primera es encontrar la respuesta en frecuencia de los dos micrófonos para la sonda y la segunda es el post-procesamiento para corregir la diferencia de fase entre ambos micrófonos.

Para la primera etapa se realiza un ensayo en el laboratorio el cual consiste en comparar la respuesta en frecuencia de un parlante medido con los dos micrófonos que componen la sonda. Este ensayo tiene por objetivo encontrar la respuesta en frecuencia de los micrófonos que serán utilizados en la sonda. El ensayo fue realizado en el laboratorio de acústica de la Pontificia Universidad Católica del Perú donde los materiales utilizados fueron:

- 2 micrófonos Bruel & Kjaer 4189.
- 1 parlante con su amplificador
- Analizador de Sonido Tipo 2012 de Bruel & Kjaer.

La figura 6 muestra el diagrama de bloques de la experiencia donde se puede observar que la señal de entrada del parlante proviene del generador de señales del propio analizador y por otro lado la señal de salida que ingresa al analizador viene a ser el sonido del parlante captado por el micrófono. Mientras que la figura 7 muestra la disposición de los elementos durante la misma. La prueba consistió de cuatro mediciones, dos mediciones por micrófono, donde el objetivo era encontrar la respuesta en frecuencia del sistema comprendido por el parlante y el micrófono con ayuda del analizador.



Fuente: Creación propia

Figura 6
Diagrama de bloques del experimento



Fuente: Creación propia

Figura 7
Medición del parlante y micrófono en campo cercano

Se realizaron dos ensayos por cada micrófono debido al rango de frecuencias de 50Hz a 20kHz, una en campo cercano para bajas frecuencias y otra en campo lejano para altas frecuencias.

La primera medición en el campo cercano se realiza con el objetivo de disminuir el efecto de las ondas estacionarias de baja frecuencia causadas por las reflexiones del cuarto. A estas bajas frecuencias el diafragma del parlante se comporta como un pistón rígido y la medición del campo cercano es directamente proporcional a la del campo lejano, en condición de empotramiento, por lo que no es afectada por el ambiente en el cual el parlante se encuentra. Sin embargo, esta medición tiene un límite máximo de frecuencia que está condicionada por el diámetro efectivo del cono del parlante. A partir de la ecuación (3.4) encontramos la frecuencia máxima válida de medición para el campo cercano.

$$F_{\max} = \frac{c}{2\pi r} \quad (3.4)$$

Donde:

F_{\max} : Frecuencia máxima

c : Velocidad del sonido (340 m/s)

r : Radio del cono (en m)

El parlante del ensayo en el laboratorio tiene un radio efectivo del cono de 6.5cm por lo cual reemplazando en la ecuación (3.4) encontramos que la frecuencia máxima válida de medición en el campo cercano es de 832.5Hz.

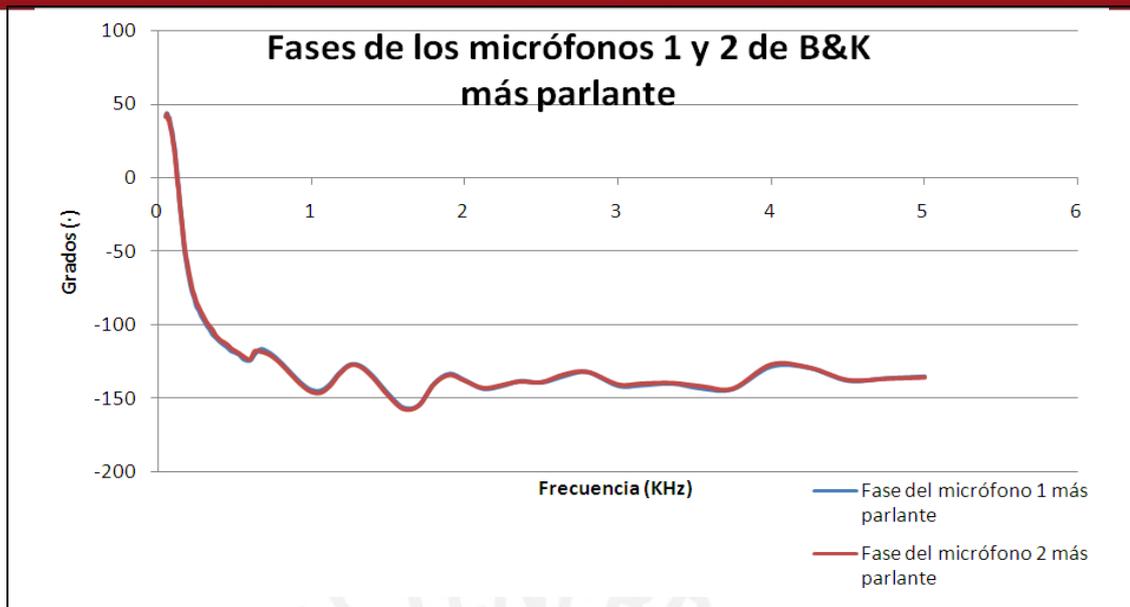
$$F_{\max} = \frac{340}{2\pi(0.065)}$$
$$F_{\max} = 832.5\text{Hz}$$

Por otro lado la segunda medición se realiza en el campo lejano para las altas frecuencias, en este caso el micrófono se encuentre aproximadamente a un metro del parlante y la frecuencia máxima que se desea medir es de 20Khz.

Posteriormente, el analizador obtiene la respuesta en frecuencia del parlante más el micrófono de cada medición por medio de la Espectrometría de Retardo en el Tiempo, conocido como TDS por sus siglas en inglés (Time Delay Spectrometry). Al unir los datos de cada medición se tiene la respuesta en frecuencia en el rango de 50Hz a 20kHz.

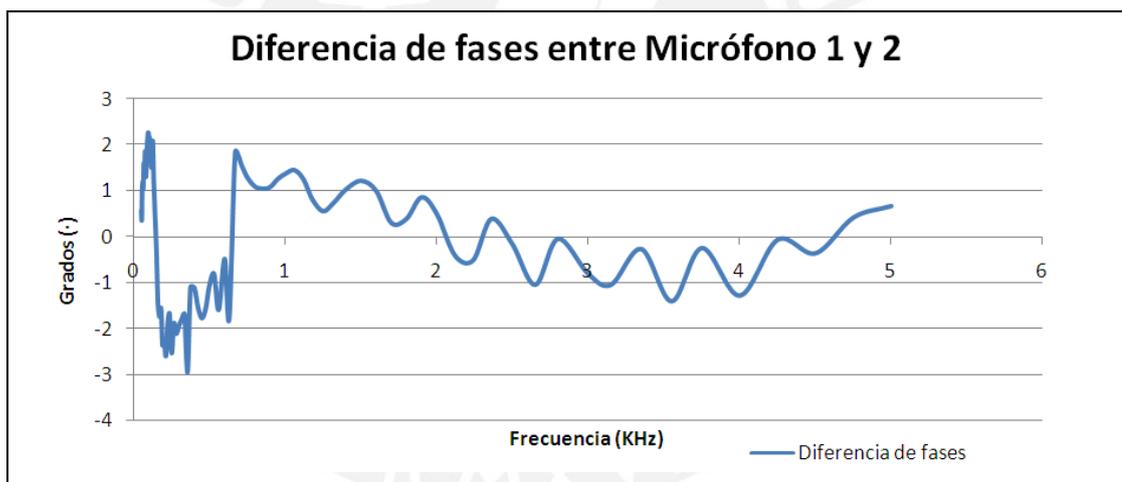
A continuación se repite el mismo procedimiento utilizando el segundo micrófono y con los datos de ambas pruebas se genera la gráfica de la Figura 8, la cual representa la fase de cada uno de los sistemas. Por otro lado, la figura 9 muestra la resta de los datos obtenidos en ambas experiencias, la cual representa la diferencia de fases entre el micrófono 1 y el micrófono 2.

Como se observa en la figura 9 la diferencia de fases entre ambos micrófonos B&K es muy pequeña, por lo cual no es necesaria una corrección de ello. Sin embargo en muchos casos la diferencia de fases entre diferentes micrófonos es considerable y es necesaria una corrección para cumplir los requerimientos de la sonda.



Fuente: Creación propia

Figura 8
Fase de los sistemas

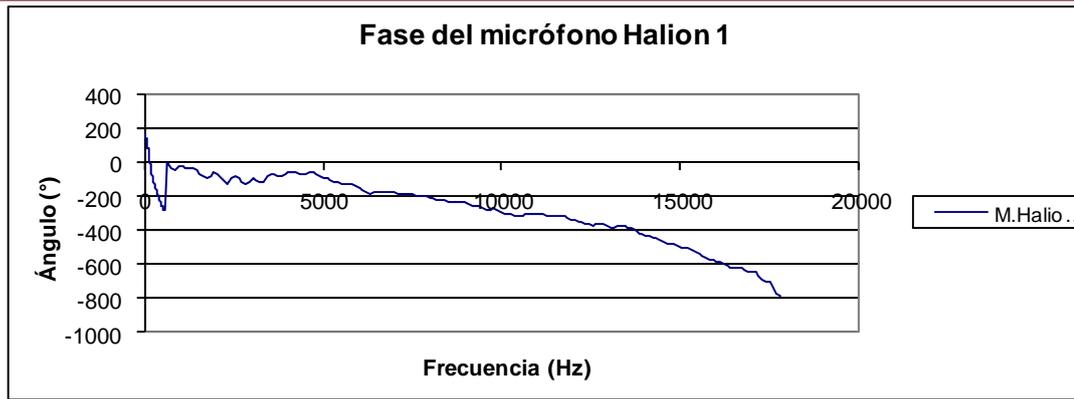


Fuente: Creación propia

Figura 9
Diferencia de Fases del micrófono 1 y 2

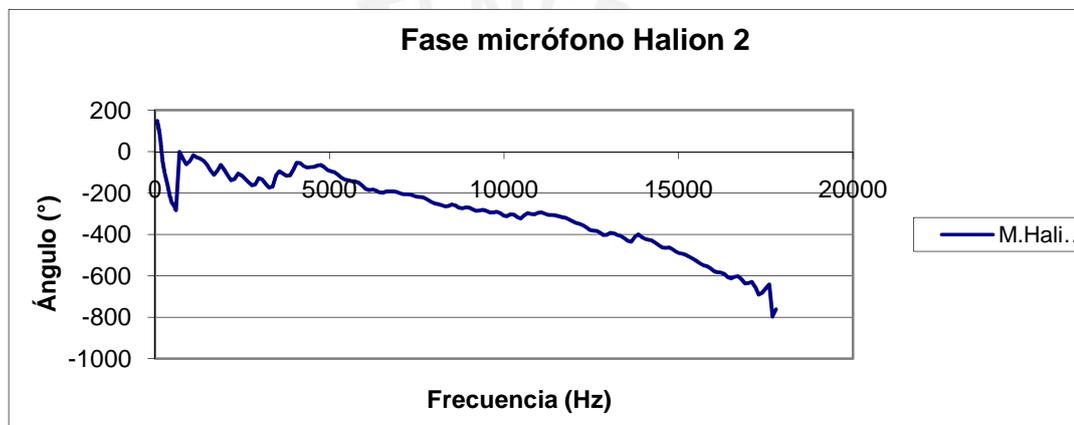
Para demostrar dicho caso se realizó una prueba en el laboratorio idéntica a la ya mencionada anteriormente pero con otros micrófonos con el objetivo de mostrar como micrófonos incluso de la misma serie pueden tener una diferencia de fases considerable y cómo se puede corregir dicha diferencia. Los materiales utilizados para la prueba fueron:

- 2 micrófonos Halion OV380A.
- 1 parlante con su amplificador
- Analizador de Sonido Tipo 2012 de Bruel & Kjaer.



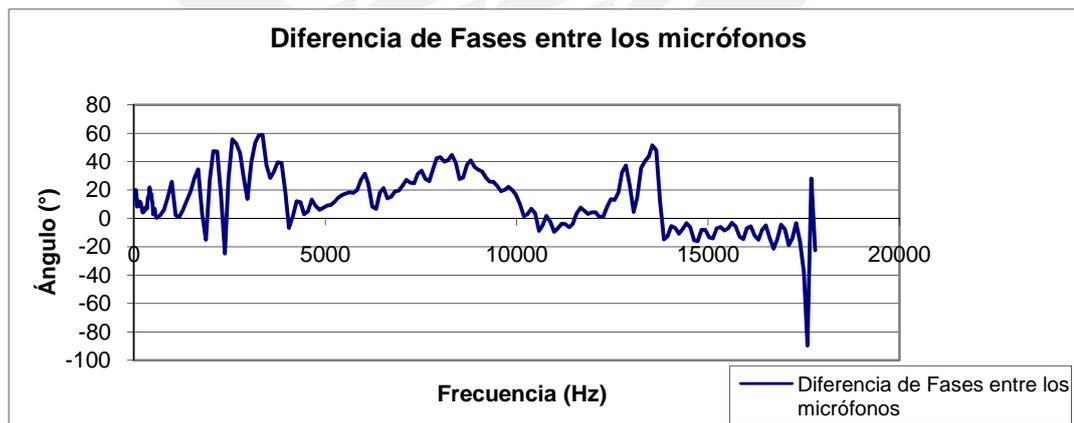
Fuente: Creación propia

Figura 10
Fase del parlante y micrófono Halion 1



Fuente: Creación propia

Figura 11
Fase del parlante y micrófono 2



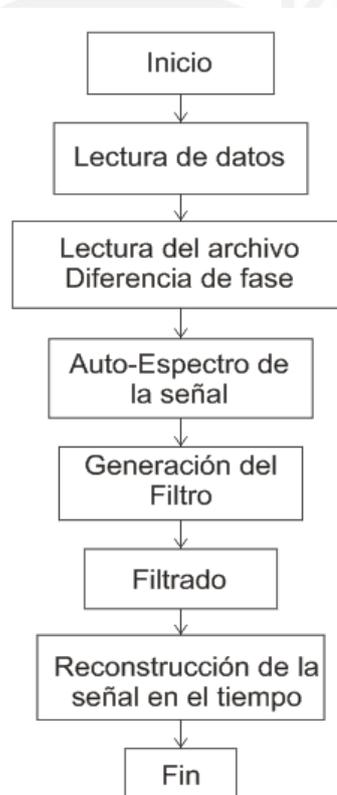
Fuente: Creación propia

Figura 12
Diferencia de fase entre los micrófonos

De dicho ensayo se obtienen las figuras 10 y 11, las cuales representan la fase de cada uno de los micrófonos Halion OV380A sumada a la fase del parlante, mientras

que la figura 12 representa la diferencia de fases de ambos micrófonos, la cual fue obtenida al restar las figuras 10 y 11, al realizar esto se elimina la fase del parlante quedando únicamente la diferencia de fases de los micrófonos.

Basado en los datos de la última gráfica se pasa a la etapa del post-procesamiento donde es posible realizar una corrección de fase a uno de los micrófonos de tal manera que luego del procesamiento la diferencia de fase entre ambos sea cercana a cero. Para esto se generó un código en Matlab (Apéndice B) que filtra la señal de uno de los micrófonos por medio de un filtro pasatodo que tiene como características magnitud uno en todo el rango de frecuencias y como fase la inversa de la función mostrada en la Figura 12, este código sigue el diagrama de flujo de la Figura 13.



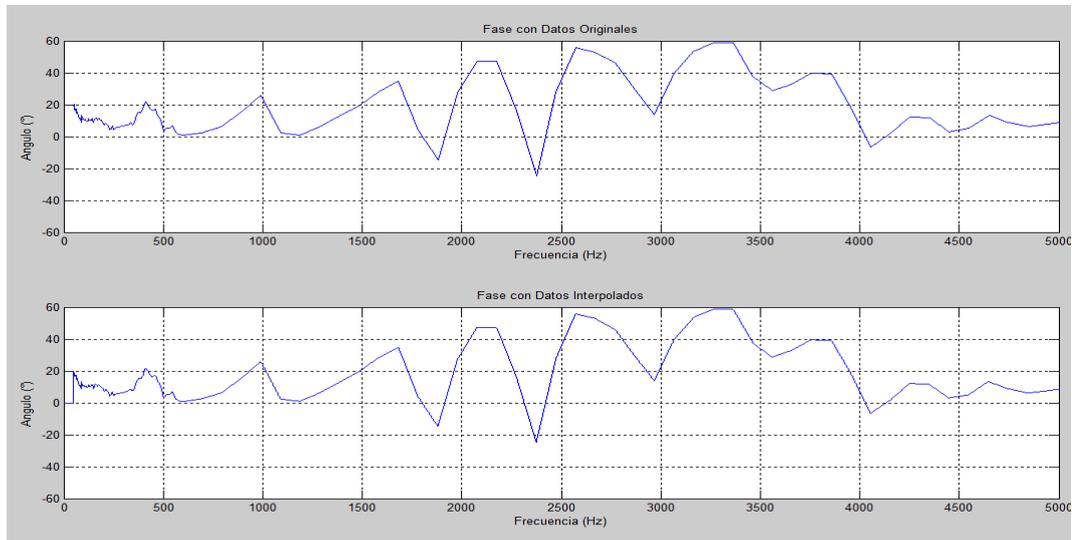
Fuente: Creación propia

Figura 13

Diagrama de flujo: Corrección de fase

En primer lugar se lee la señal a ser filtrada y el archivo en Excel que contiene los datos de la diferencia de fase entre los micrófonos, después se obtiene el auto-espectro de la señal por medio de la FFT, luego se obtiene los valores de frecuencia en donde ha sido analizada la señal por la FFT, en base a estos valores

se realiza una interpolación simple con los datos leídos de la parte anterior con el objetivo de reconstruir la gráfica en esos puntos. La figura 14 en su parte inferior muestra este resultado.



Fuente: Creación propia

Figura 14

Comparación entre gráficas de diferencia de fase

Posteriormente para la generación del filtro se parte de la premisa que un filtro analizado en frecuencia es un vector de números complejos que está comprendido por magnitud y fase en un rango de frecuencias. Para el caso de un filtro pasa-todo la magnitud es uno en todas las frecuencias y la fase es la que varía de acuerdo a los requerimientos. Para este caso en particular la fase será la inversa de los datos obtenidos por medio de la interpolación. Esto se puede representar como un vector de números complejos (Ec. (3.5)) donde cada componente tiene la forma:

$$Z = re^{i\phi} \quad (3.5)$$

donde:

Z : Número complejo

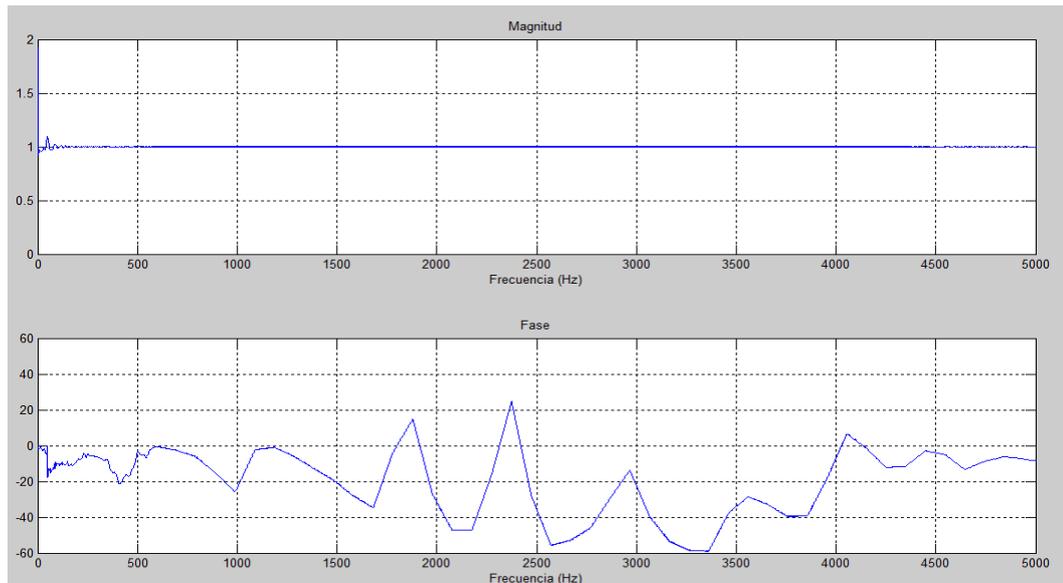
r : magnitud (respuesta unitaria en todo el rango de frecuencias)

ϕ : fase (según requerimiento)

Luego se procede al filtrado de la señal el cual se realiza por medio de la división de su auto-espectro y el vector que contiene al filtro, el resultado es la conservación de la magnitud y la variación de la fase. Después a los datos se les aplica

Transformada Inversa de Fourier (IFFT) y se almacenan en un archivo de formato .wav que reconstruye la señal en el tiempo.

Por último la respuesta del filtro se puede observar en la Figura 15 donde se demuestra que el filtro tiene una respuesta de magnitud prácticamente unitaria en el rango de frecuencias de 0Hz a 5kHz, mientras que la fase es la inversa de la diferencia de fases entre los dos micrófonos.



Fuente: Creación propia

Figura 15
Respuesta del filtro pasatodo

3.3.3 Corrección de la magnitud

El último requerimiento para la creación de la sonda es la corrección de magnitud, la cual en nuestro caso en particular no es necesaria puesto que se está utilizando dos micrófonos B&K 4189 los cuales tiene respuesta plana en magnitud en el rango de frecuencias de la medición a realizar.

Sin embargo es necesario calibrar ambos micrófonos de tal manera que emitan un mismo nivel de voltaje a un mismo nivel de presión, para ello fue necesario un calibrador de presión sonora y un multímetro, para la tesis se utilizó un multímetro de banco Fluke 8840A (Figura 16a). La calibración consistió en colocar un micrófono en el calibrador el cual emitía 93.8dB de presión sonora y se observó que el valor de voltaje del primer micrófono fue de 47.9mV a ese nivel de presión. A continuación se colocó el segundo micrófono en el calibrador y se observó el valor

de voltaje que emitía, el cual era mucho mayor que el primero, esto se debe a que uno de los micrófonos estaba conectado a un preamplificador, mientras que el otro estaba conectado a un sonómetro, el cual le brindaba una amplificación mayor al del preamplificador. Se realizó el ensayo de esta manera ya que el laboratorio en ese momento no contaba con dos preamplificadores para los micrófonos. Por lo cual fue necesario utilizar un atenuador de décadas (Figura 16b) para reducir los niveles de voltaje e igualar el valor del primer micrófono, el resultado obtenido fue de 48.0mV a 93.8dB de presión sonora.



Fuente: Creación propia

Figura 16a

Multímetro Fluke 8840A



Fuente: Creación propia

Figura 16b

Atenuador por décadas

Una vez realizada la calibración ya se ha cumplido con todos los requerimientos para la creación de la sonda de intensidad sonora y con esto se ha cumplido con el primer objetivo específico de la tesis.

A pesar de ello, si en algún caso se deseara utilizar micrófonos que no tienen respuesta plana en la magnitud es necesario realizar una corrección de ellos. Esto se logra en dos etapas: la primera es encontrar la respuesta en frecuencia de los micrófonos y la segunda es el post-procesamiento para corregir la magnitud.

Para demostrar que es posible realizar esta corrección, a continuación se muestra un ejemplo de cómo realizarlo. Para la primera etapa se realiza un ensayo en el laboratorio el cual consiste en comparar la respuesta en frecuencia de un parlante medido con dos micrófonos diferentes: uno con magnitud plana y el otro el que se desea corregir. Este ensayo tiene por objetivo encontrar la respuesta en frecuencia del micrófono que se desea utilizar para la sonda.

El ensayo fue realizado en el Laboratorio de Acústica de la Pontificia Universidad Católica del Perú donde los materiales utilizados fueron:

- Micrófono 4133 de Bruel & Kjær
- Micrófono Halion OV380A
- Analizador de Sonido Tipo 2012 de Bruel & Kjær
- Un parlante con su amplificador

El diagrama de bloques de la experiencia es el mismo que el de la figura 6 puesto que la experiencia es bastante similar. La disposición de los elementos se muestra en la figura 17.



Fuente: Creación propia

Figura 17

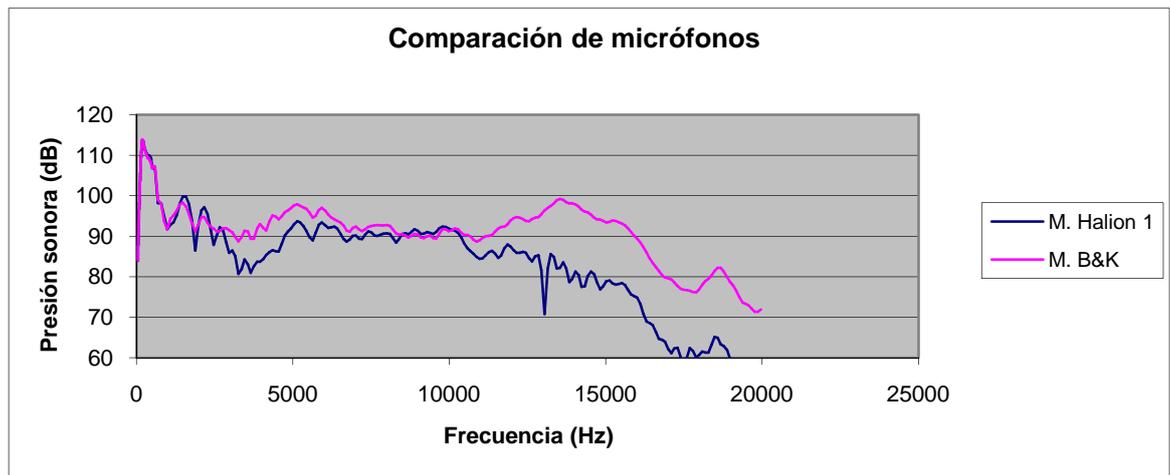
Disposición de elementos para corrección de magnitud

Se realizan dos ensayos por cada micrófono debido al rango de frecuencias, una en campo cercano y otra en campo lejano, las razones de esto fueron expuestas en la sección anterior, el parlante utilizado fue el mismo que para el ensayo de corrección de fase, por lo cual la frecuencia límite en campo cercano es de 832.5Hz.

A partir de las señales de entrada y de salida se obtiene la respuesta en frecuencia con la FFT del parlante más el primer micrófono. A continuación se repite la misma experiencia utilizando el segundo micrófono. Finalmente la figura 18 muestra la comparación entre las respuestas en frecuencia del parlante obtenida por el micrófono Bruel & Kjær (B&K) y el micrófono Halion respectivamente.

Gracias a esta figura se puede determinar en qué rangos es necesario corregir el micrófono Halion. En primer lugar se observa que en bajas frecuencias

aproximadamente hasta 1kHz la respuesta es prácticamente idéntica en ambos micrófonos, mientras que a partir de los 10kHz la diferencia entre las magnitudes del micrófono B&K y el Halion supera los 10dB, a partir de este punto ya no es recomendable realizar una corrección de magnitud puesto que se estaría amplificando el ruido. Para efectos de la tesis es más que suficiente trabajar hasta los 5kHz.



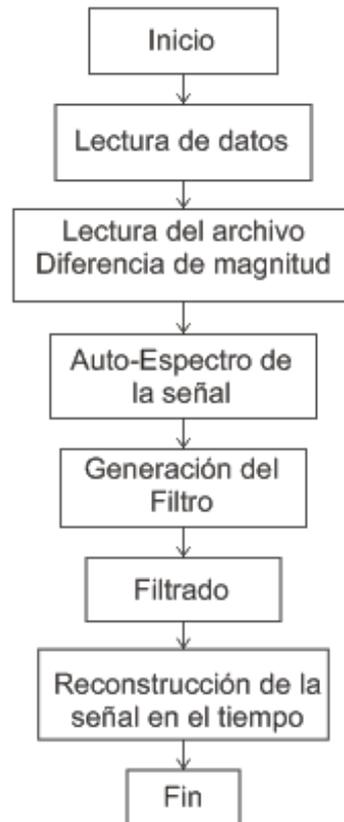
Fuente: Creación propia

Figura 18

Comparación de magnitudes M. B&K y M. Halion

Una vez terminados los ensayos y sabiendo los requerimientos de corrección se procede a la segunda etapa, la cual consiste en corregir la magnitud por medio de software, para objeto de la tesis se ha utilizado Matlab para el procedimiento, el cual debe efectuarse cada vez que se graba una señal con ese micrófono específico, si se desea utilizar otro micrófono debe realizarse el ensayo nuevamente. Para la corrección se generó un código (Apéndice C) el cual tiene como objetivo filtrar la señal grabada por el micrófono utilizado en la primera etapa de tal manera que el resultado sea una señal muy parecida a una grabada por un micrófono de respuesta plana. La figura 19 muestra el diagrama de flujo que sigue el código.

En primer lugar se lee la señal a ser filtrada y el archivo en Excel que contiene los datos obtenidos en el ensayo de laboratorio, con la diferencia que estos datos leídos corresponden únicamente a la respuesta en frecuencia de uno de los micrófonos, esto se obtuvo al restar en dB los datos obtenidos por el micrófono Halion y los datos del micrófono B&K.



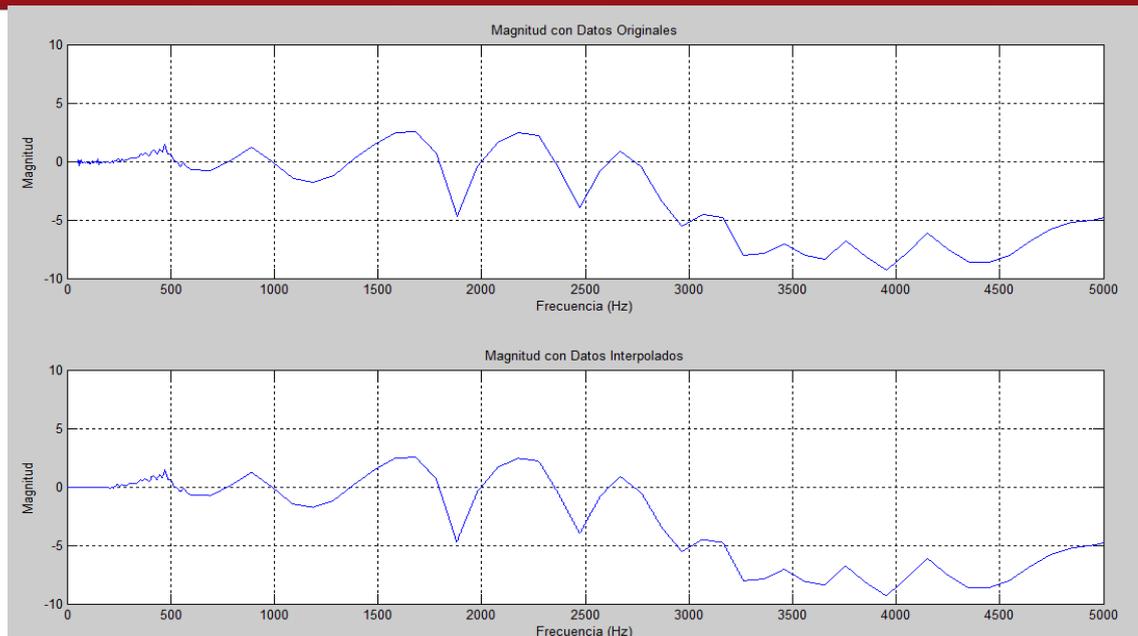
Fuente: Creación propia

Figura 19

Diagrama de bloques corrección de magnitud

Después se obtiene el auto-espectro de la señal por medio de la FFT. A continuación se realiza un procedimiento prácticamente idéntico a la sección de corrección de fase, donde se interpolan los datos de magnitud para que encajen con las frecuencias del auto-espectro.

La figura 20 en su parte superior muestra los datos obtenidos en el laboratorio, mientras que la parte inferior muestra los datos obtenidos por la interpolación. Como se observa en esta figura la corrección empieza a partir de los 200Hz aproximadamente puesto que para frecuencias menores que esta, la magnitud es prácticamente plana y no requiere ningún tipo de corrección, mientras que para frecuencias mayores a 5kHz tampoco se realizan correcciones ya que el rango válido de medición de los micrófonos ya corregidos es de 20Hz a 5kHz, adaptado para medir la intensidad sonora de la voz humana.



Fuente: Creación propia

Figura 20

Comparación entre datos originales e interpolados de magnitud

Posteriormente para la generación del filtro se parte de la premisa que un filtro analizado en frecuencia es un vector de números complejos que está comprendido por magnitud y fase en un rango de frecuencias. Para este caso en particular el rango de frecuencias es el mismo que el obtenido en el auto-espectro de la señal, la magnitud será la respuesta de magnitud del micrófono y la fase es siempre cero.

Esto se puede representar como la ecuación (3.5), es decir un vector de números complejos donde cada componente tiene la forma:

$$Z = re^{i\phi} \quad (3.5)$$

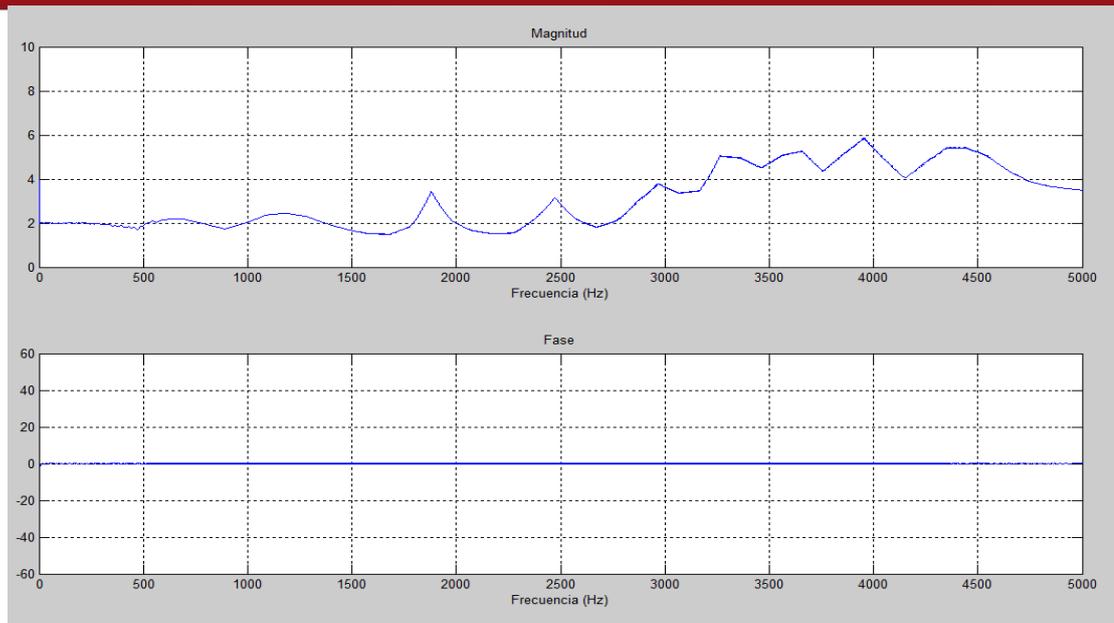
donde:

Z : Número complejo

r : Magnitud (respuesta en magnitud del micrófono)

ϕ : Fase (cero para todas las frecuencias)

Enseguida se procede al filtrado el cual se realiza por medio de la división del auto-espectro de la señal entre el vector que contiene al filtro, lo que da como resultado una variación en la magnitud y la conservación de la fase. Después a estos nuevos datos se les aplica la Transformada Inversa de Fourier (IFFT) y se almacenan en un archivo de formato .wav que reconstruye la señal en el tiempo.



Fuente: Creación propia

Figura 21
Respuesta en frecuencia del filtro

Por último la comprobación de la respuesta del filtro se observa al dividir en frecuencia la señal reconstruida con la señal original. La figura 21 demuestra que el filtro tiene una respuesta de magnitud muy parecida a la inversa del micrófono con la cual fue grabada la señal y su fase es cero en el rango de frecuencias de 0 a 5kHz.

A pesar que se pueden realizar estas correcciones de magnitud y fase para ajustar las características dinámicas de los micrófonos para que funcionen como una sonda de intensidad sonora, es recomendable que se utilicen micrófonos de buena calidad para poder garantizar una estabilidad en las mediciones que se realicen.

3.4 Medición de intensidad sonora

Con la sonda completa es posible realizar la medición de intensidad sonora basada en la información entregada por los dos micrófonos que será procesada por medio de un código implementado en Matlab. Para el caso de la tesis el cálculo de intensidad sonora sigue el diagrama de bloques que muestra la figura 22, el cual representa la ecuación (3.6) demostrada en el apéndice A.

$$I = -\frac{P_1 + P_2}{2\rho\Delta r} \int (P_1 - P_2) dt \quad (3.6)$$

donde:

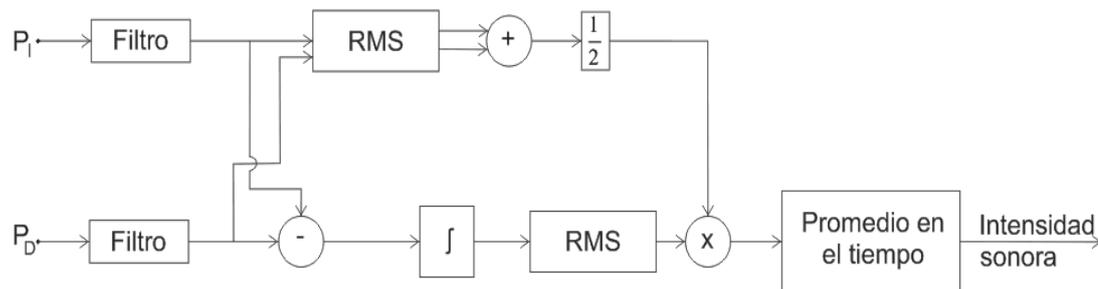
I : Intensidad sonora

P_1 : Presión del micrófono 1

P_2 : Presión del micrófono 2

ρ : Densidad del aire

Δr : Distancia de separación de los micrófonos



Fuente: Creación propia

Figura 22

Diagrama de bloques de intensidad sonora

En primer lugar cada señal proveniente de los micrófonos pasa por un filtro pasa-banda digital, donde su ancho de banda y frecuencia de corte son variables, esto tiene como objetivo centrar el estudio de la intensidad sonora en un rango de frecuencias deseado. En seguida, ambas presiones de los micrófonos son sumadas y divididas entre dos para obtener la presión promedio en el centro de la sonda. A continuación se realiza la resta de presiones entre ambos micrófonos y se le aplica la integral que da como resultado la velocidad de partículas. Estos dos resultados son multiplicados para obtener la intensidad sonora en el tiempo, la cual se convierte en un único valor al realizarse un promedio en el tiempo.

Además en la figura 20 se muestra dos bloques RMS, los cuales representan al valor cuadrático medio, el apéndice D explica el funcionamiento de estos bloques.

3.5 Algoritmo para la localización de una fuente sonora tomando como referencia su mayor nivel de intensidad

La última parte del diseño e implementación de la tesis consiste en unir todas las partes realizadas anteriormente en un solo código (Apéndice E), que además

incluya el procedimiento para hallar el ángulo de rotación de la cabeza respecto a la sonda, este ángulo se obtiene por medio de la ecuación (3.1) y finalmente dependiendo del ángulo que se tenga por resultado se le asignará una zona correspondiente, la cual determinará cuanta corrección debe realizarse al micrófono del expositor, para mejorar la inteligibilidad de sus palabras. La tabla 1 muestra la asignación de zonas según el ángulo de rotación de la cabeza del orador.

Tabla 1
Asignación de zonas dependiendo del ángulo

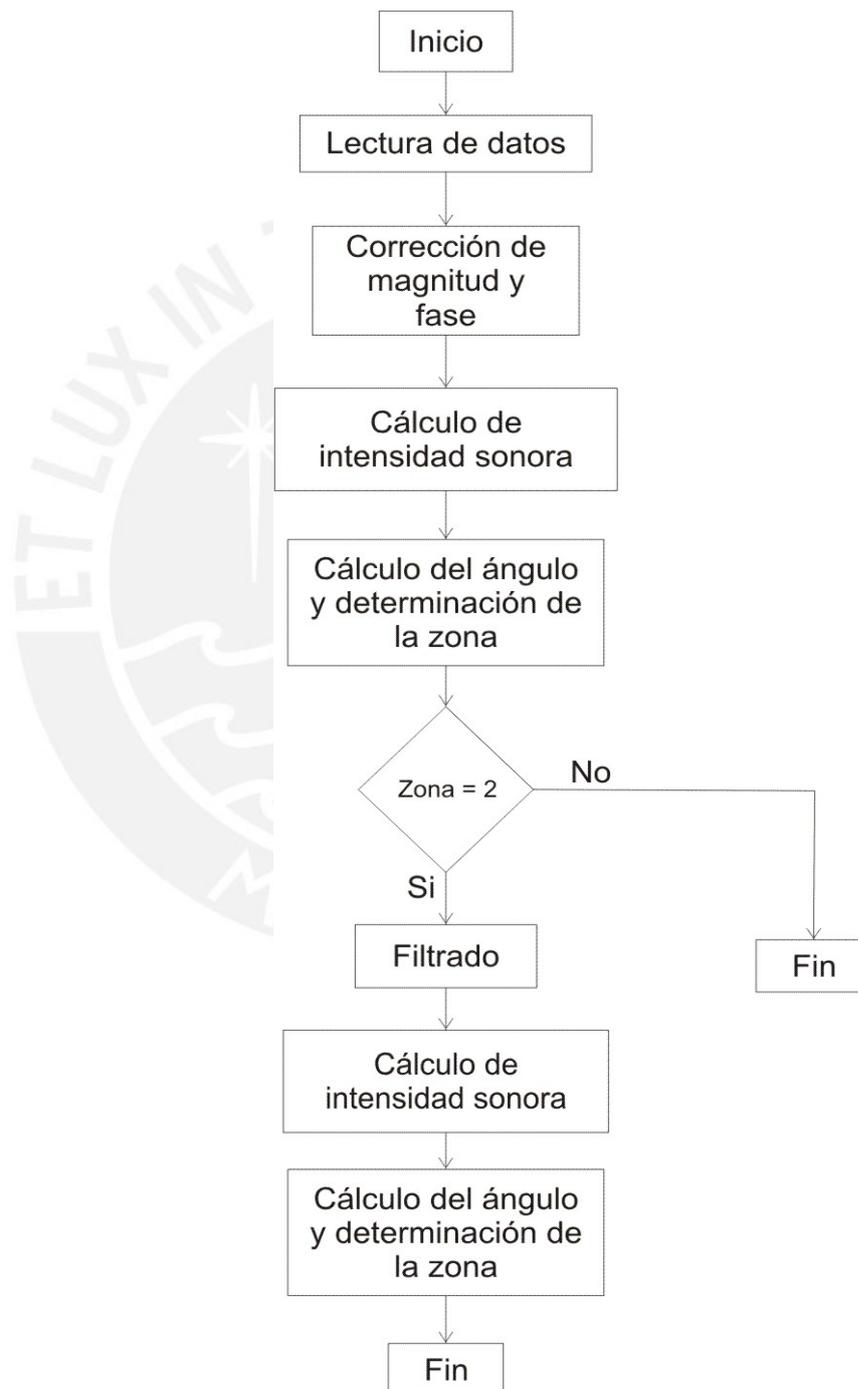
Zonas	Rango de ángulos
-4	-90° a -60°
-3	-60° a -45°
-2	-45° a -30°
-1	-30° a 0°
1	0° a 30°
2	30° a 45°
3	45° a 60°
4	60° a 90°

Según la Tabla 1 se puede diferenciar 8 zonas en total, cuatro para cada lado de la sonda, el signo negativo indica que se encuentra para el lado derecho de la sonda mientras que el signo positivo indica que se encuentra al lado izquierdo de la misma. Esta diferenciación de zonas principalmente es producto de la interacción del patrón de radiación de la voz y el patrón bidireccional de la sonda de intensidad.

Como se observó cuando se mostraron los patrones de radiación de la voz no existe diferenciación de los niveles de intensidad en el rango de 0° a 30° por lo que solamente se puede indicar que la persona se encuentra hablando en ese rango. Por otra parte según los diagramas polares del patrón de radiación se ha seleccionado las zonas 2, 3 y 4 debido a su caída progresiva en magnitud respecto a la zona 1.

La figura 23 muestra el diagrama de flujo del programa principal el cual inicia leyendo los datos a analizar, los cuales pasan por la corrección de magnitud y fase para cada micrófono, a continuación los datos corregidos pasan por un filtro pasa-banda de frecuencia central de 250Hz y un ancho de banda de 100Hz, y posteriormente se obtiene el valor de intensidad sonora con lo cual se calcula el

ángulo de rotación de la cabeza, en base a este ángulo se hace una primera diferenciación de tres zonas, si la zona encontrada es la primera o tercera el programa finaliza allí, caso contrario se pasa la señal original por otro filtro pasabanda de frecuencia central 500Hz con ancho de banda de 100Hz, lo que nos permitirá descomponer la zona 2 en 2 zonas diferentes, lo cual dará un total de 4 zonas de diferenciación por lado dependiendo del ángulo encontrado.



Fuente: Creación propia

Figura 23

Diagrama de flujo del programa principal

El programa reconoce en qué lado se encuentra la fuente sonora distinguiendo la diferencia de presiones de los micrófonos, la presión será mayor en el micrófono de la derecha si la fuente se encuentra en ese lado y le asigna un valor negativo, caso contrario el valor es positivo.

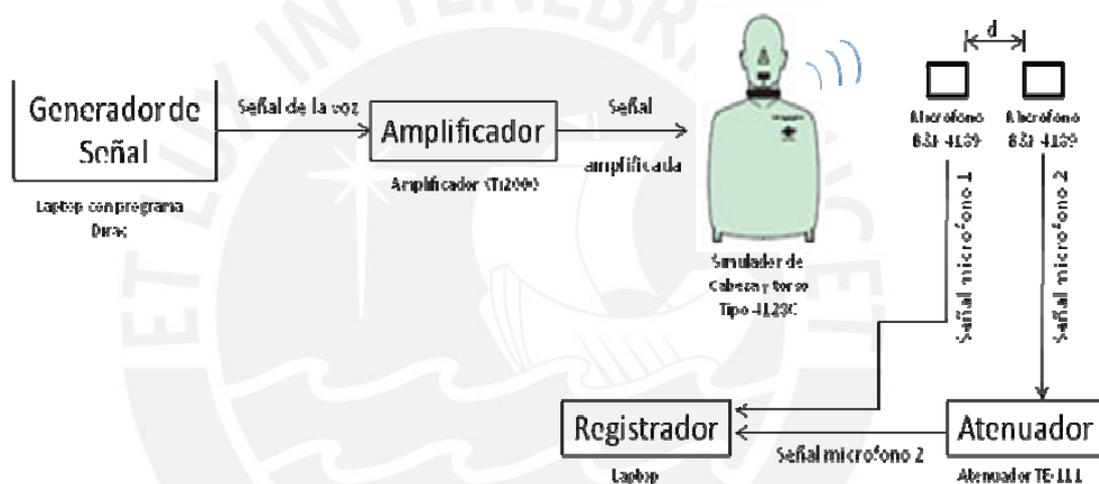
Por último, se eligieron esos valores de frecuencias para los filtros porque la voz tiene mayor potencia en esas frecuencias y segundo porque permiten diferenciar claramente ciertas zonas dependiendo del patrón de radiación. No se tomó en cuenta la frecuencia de 1kHz puesto que su intensidad se mantiene dentro del rango de ángulos y por encima de esta frecuencia la potencia de la voz es muy baja, lo cual no es recomendable para analizar.



Capítulo 4: “Pruebas y resultados del diseño”

4.1 Pruebas

Para comprobar que el diseño de la sonda y el programa de determinación de la posición de la voz tomando como punto de referencia el nivel máximo de su patrón de radiación funcionan adecuadamente se instaló en el laboratorio de acústica un montaje de instrumentos (la figura 24 muestra el diagrama de montaje de los mismos). Éste permite simular los movimientos de rotación de la cabeza respecto a un punto, lo cual se logra colocando el Simulador de Cabeza y Torso tipo 4128C de B&K sobre un goniómetro, el cual le permitirá girar sobre su propio eje.



Fuente: Creación propia

Figura 24

Diagrama de bloques de la experiencia

Los instrumentos utilizados en la experiencia fueron:

- Simulador de Cabeza y Torso tipo 4128C de B&K
- 2 micrófonos 4189 de B&K
- 1 Goniómetro
- 1 Amplificador XTi2000 de CROWN
- 1 Laptop IBM con el programa DIRAC
- 1 Laptop HP Pavillion DV6000 para el registro de las señales
- 1 Atenuador de décadas TE-111 de Tech Instruments

La figura 24 muestran los instrumentos y las conexiones entre los instrumentos, donde la señal de la voz sale de la Laptop IBM la cual entra al amplificador y luego al simulador de torso y cabeza que emitirá la señal, esta será recibida por el arreglo de los dos micrófonos donde uno estará conectado directamente a la Laptop HP y el otro pasará por el atenuador y entrará a la computadora, donde ambas señales serán registradas.

Se decidió utilizar el simulador de torso y cabeza puesto que es un maniquí que posee una estructura de oreja y boca que reproduce de forma realista las propiedades acústicas de un humano adulto promedio, lo cual es ideal para mediciones de dispositivos de conferencias, entre otras aplicaciones.

Por otro lado, la señal utilizada para las pruebas es un Secuencia de Longitud Máxima (MLS, por sus siglas en inglés) con un filtro que simula el espectro de frecuencias de la voz de un hombre promedio, la cual está estandarizada para pruebas de inteligibilidad del habla según la norma IEC 60268-16. Esta señal se obtuvo del programa DIRAC de la compañía B&K, el objetivo principal de elegir esta señal y el maniquí como objetos de pruebas es reducir al mínimo los errores producidos por la fuente, y centrarse únicamente en los problemas encontrados por el algoritmo de la tesis, de esta manera se logra una uniformidad en todas las pruebas que se realicen.



Fuente: Creación propia

Figura 25a
Prueba de calibración



Figura 25b
Prueba de calibración

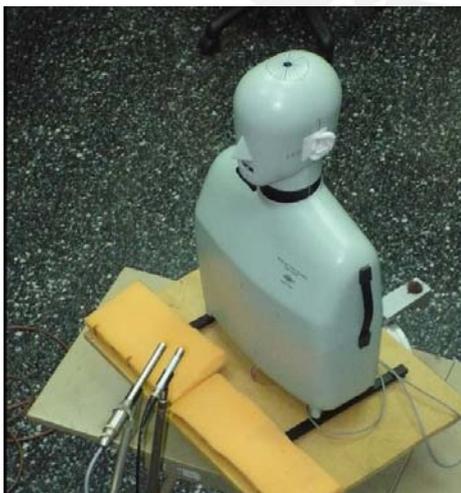
Para comenzar la prueba se colocó el arreglo de micrófonos en frente del maniquí a una distancia de 50cm aproximadamente, mientras que los micrófonos se colocaron

uno al lado del otro con una separación de 3cm. entre ellos. El primer paso es el proceso de calibración el cual va a brindar el valor de referencia para hacer los cálculos de las siguientes mediciones. Este valor se obtiene cuando el maniquí está exactamente en frente de los micrófonos y emite una señal la cual viene a ser el valor máximo de intensidad de su patrón de radiación (Figura 25).

Posteriormente la prueba prosigue girando el torso 5° y emitiendo la misma señal por medio de la boca del maniquí, la cual será captada por los micrófonos y almacenada en formato .wav con el programa Audacity 1.3, donde la frecuencia de muestreo de la señal es de 16kHz. Luego se procede a girar nuevamente el torso en 5° con ayuda del goniómetro, y se repite la experiencia hasta llegar a un ángulo de 90° que es el valor máximo que se ha tomado en cuenta para las pruebas.

En cada prueba la señal es guardada con el nombre del ángulo de giro del goniómetro para su posterior comparación con los resultados del programa. Los movimientos de giro del torso simulan la rotación de la cabeza de una persona en frente de un micrófono sin moverse de su lugar. Se tomó el caso máximo de 90° puesto que se asume que es el peor escenario que se da cuando el expositor dirige la palabra a un colega que está a su costado.

La figura 26 muestra el torso en una de las posiciones durante la realización de las pruebas.



Fuente: Creación propia

Figura 26a

Torso y arreglo de micrófonos



Fuente: Creación propia

Figura 26b

Torso y arreglo de micrófonos

4.2 Resultados

Los siguientes resultados son el producto de cinco pruebas realizadas con el método descrito en la sección anterior, donde se hicieron mediciones cada 5°, las cuales fueron almacenadas con el nombre de los grados indicados por el goniómetro y posteriormente fueron procesadas. La tabla 2 muestra los valores de los ángulos obtenidos después del procesamiento.

Tabla 2
Tabla de resultados general

Ángulo (°) según el goniómetro	Zonas	Prueba 1 ángulo (°)	Prueba 2 ángulo (°)	Prueba 3 ángulo (°)	Prueba 4 ángulo (°)	Prueba 5 ángulo (°)
0	1	Referencia	Referencia	Referencia	Referencia	Referencia
5		mayor	18	12	23	3
10		mayor	mayor	12	16	19
15		28	12	21	6	19
20		9	20	11	13	27
25		33	23	34	39	28
30		30	23	47	44	32
35	2 y 3	17	40	39	42	28
40		44	46	49	47	36
45		44	53	52	49	40
50		54	54	57	56	49
55		56	58	57	56	51
60	4	55	60	61	54	51
65		62	65	65	60	61
70		62	63	66	62	63
75		66	68	69	63	66
80		62	72	71	67	68
85		68	72	71	67	68
90		72	73	73	68	70

Esta tabla muestra los datos obtenidos después del procesamiento realizado con la ayuda del primer filtro con frecuencia central de 250Hz, donde se reconocen tres zonas: la primera es la zona 1 entre 0° y 30°, la segunda el conjunto de la zonas 2 y 3 comprendido entre los ángulos de 30° a 60° y la tercera es la zona 4 de los ángulos entre 60° y 90°. Si el programa detecta que se encuentra en el conjunto de la zona 2 y 3, se pasa al segundo proceso de filtrado con una frecuencia central de 500Hz donde se discierne si la rotación de la cabeza pertenece a la zona 2 o a la zona 3, los resultados de ello se muestran en la tabla 3.

Tabla 3

Tabla de resultados zonas 2 y 3

Ángulo (°) según el goniómetro	Zonas	Prueba 1 ángulo (°)	Prueba 2 ángulo (°)	Prueba 3 ángulo (°)	Prueba 4 ángulo (°)	Prueba 5 ángulo (°)
0	1	Referencia	Referencia	Referencia	Referencia	Referencia
35	2	43	25	28	41	-
40		45	33	43	27	26
45		53	47	45	34	38
50	3	50	50	54	46	46
55		60	56	58	49	49
60		60	55	57	54	52

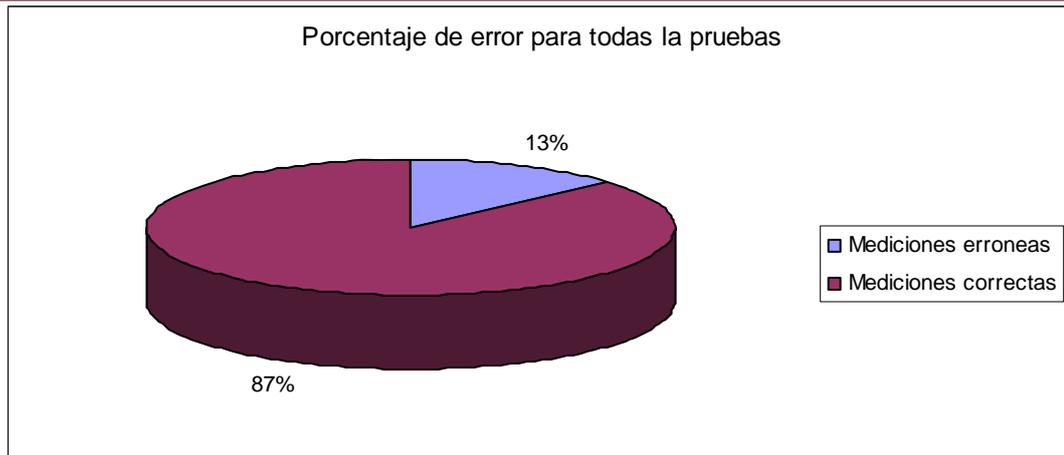
Los valores en color azul de la tabla 3 muestran que el ángulo encontrado es menor que 30°, lo cual podría considerarse como un error, sin embargo en la primera parte del procesamiento ya se ha determinado que estos valores pertenecían a la zona 2 o a la zona 3, al ser menores que 45° se les asigna inmediatamente la zona 2. Mientras que los valores en color rojo muestran los puntos donde el programa ha cometido un error. La tabla 4 muestra el error producido en cada prueba independientemente de las otras y el error si se toma en consideración todas las mediciones, mientras que la figura 27 muestra un gráfico circular con el porcentaje de error producido en todas las pruebas.

Tabla 4

Tabla de los errores

N° de prueba	N° de muestras	N° de errores	Error (%)
1	18	3	16.67
2	18	1	5.56
3	18	3	16.67
4	18	3	16.67
5	18	2	11.11
Todas	90	12	13.33

Finalmente estos son los resultados que se obtuvieron para el movimiento del maniquí hacia el lado derecho de la sonda de intensidad, no se muestran resultados para el lado izquierdo puesto que el patrón de radiación de la voz es simétrico por lo que los resultados en este lado serían muy similares a los ya presentados con la única diferencia que las zonas tendrían un valor negativo que indica el lado que se encuentra la boca respecto a la sonda.



Fuente: Creación propia

Figura 27

Gráfico circular que indica el porcentaje de error

4.3 Presupuesto

La tabla 5 presenta el presupuesto de la ejecución del proyecto.

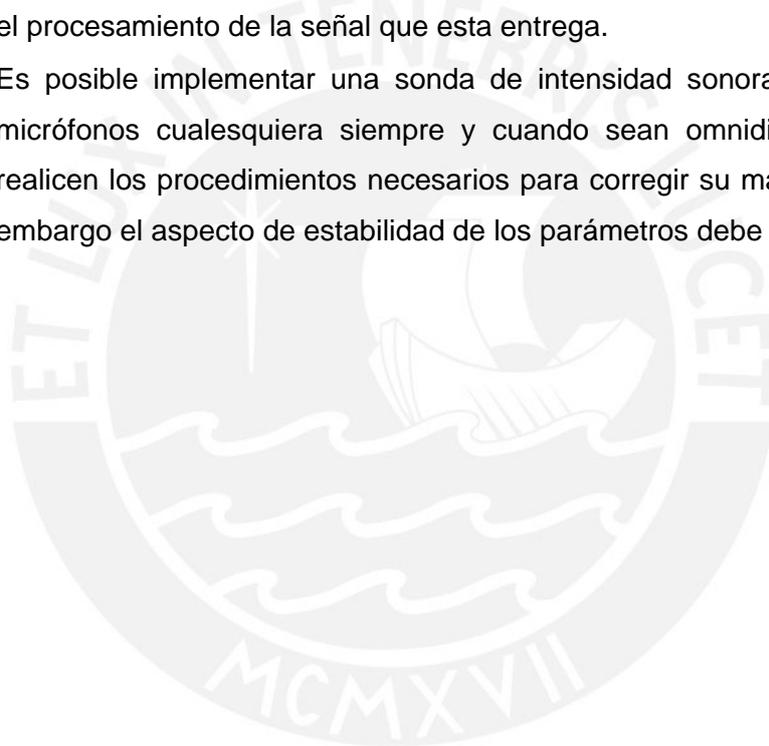
Tabla 5
Presupuesto del proyecto

Materiales	Precio unitario (\$)	Cantidad	Costo (\$)	Costo (S/.)
Computadora	600.00	1	600.00	1,668.00
Mediciones	500.00	6	3,000.00	8,340.00
Honorarios	3,500.00	1	3,500.00	9,730.00
Total			7,100.00	19,738.00

Para el presupuesto del proyecto se han tomado en cuenta diferentes consideraciones. En primer lugar el tipo de cambio es de S/. 2.78, en segundo lugar el costo de las mediciones realizadas se han obtenido como si se tratase de un servicio a terceros realizado por la Pontificia Universidad Católica del Perú, en donde se pueden utilizar todos los equipos del laboratorio de acústica con un costo de \$500 por día. Los honorarios es un único pago realizado a la persona encargada del proyecto el cual tiene una duración de 3 meses aproximadamente. Por último se asume la compra de una computadora la cual será utilizada para realizar toda la investigación y creación del programa para el proyecto.

CONCLUSIONES

- Se diseñó e implementó un arreglo de dos micrófonos que cumplen con los requisitos de una sonda de intensidad sonora, con lo cual se cumple el primer objetivo específico de la tesis.
- Se implementó un algoritmo de localización de la mayor intensidad de la voz, el cual da como resultado un ángulo sólido respecto a un punto origen, el cual se obtiene de un arreglo de dos micrófonos, cumpliendo de esta manera con el segundo objetivo específico de la tesis.
- Se ha logrado cumplir con el objetivo general de determinar la dirección de mayor intensidad de la voz por medio de una sonda de intensidad sonora y el procesamiento de la señal que esta entrega.
- Es posible implementar una sonda de intensidad sonora a partir de dos micrófonos cualesquiera siempre y cuando sean omnidireccionales y se realicen los procedimientos necesarios para corregir su magnitud y fase sin embargo el aspecto de estabilidad de los parámetros debe tenerse presente.



RECOMENDACIONES

- Si se desea utilizar otro tipo de micrófonos para armar la sonda se debe de verificar que éstos tengan respuesta plana en el rango de frecuencias que se desea medir, además deben de ser omnidireccionales, caso contrario el rango efectivo de ángulos que es capaz de medir la sonda se reducirá y la estabilidad se verá comprometida.
- La corrección de la magnitud y fase es estrictamente necesaria si los micrófonos no son de respuesta plana ni de diferencia de fases lo más cercano a cero respectivamente.
- Es necesario hacer un estudio que permita determinar los momentos en los cuales la voz se encuentre presente dentro de una señal de audio, puesto que el sistema asume que la voz siempre se encuentra presente. Esto puede ser realizado como una siguiente tesis para continuar con el proyecto.
- El sistema aun requiere de un mayor estudio puesto que para efectos de la tesis sólo se está contemplando los movimientos de la cabeza del orador en el eje horizontal, sin embargo también existen movimientos en el eje vertical y de adelante hacia atrás.
- Debido al costo de la implementación del sistema se le ha diseñado para salas de conferencia de cierto nivel donde se supone que se guarda un respeto y silencio por el orador que está exponiendo, por lo cual el sistema no contempla soluciones al ruido de fondo puesto que supone que estos son mínimos. Si se deseará implementarlo para otros lugares sería necesario añadir un filtro de reducción de ruido.
- La segunda parte del proyecto que implica la ecualización del micrófono por el cual habla el expositor, debe tener en cuenta que la corrección debe realizarse de manera progresiva para que el cambio generado en los niveles de intensidad sonora no sean demasiado bruscos, los cuales podrían generar molestias al público en general.

Bibliografía

- [1] Brüel & Kjær Company
s/a “*Spherical Beamforming*”
Fecha de consulta: 21/05/09
<<http://www.bksv.com/products/pulseanalyzerplatform/pulsesolutionsoverview/acousticapplications/sphericalbeamforming.aspx>>
- [2] J. D. Maynard, E. G. Williams, Y. Lee
“*Nearfield acoustic holography: I. Theory of generalized holography and the development of NAH*”. J. Acoust. Soc. Am.
1985, Vol. 78, n°4, pp. 1395-1413
- [3] K.B.Ginn, J.Hald
2005 “*Combined NAH and Beamforming using the same microphone array*”. Paper presentado en el Forum Acoustic. Budapest.
- [4] Brüel & Kjær Company
2008 “*Pulse Array-based Noise Source Identification Solutions*”.
Fecha de consulta: 21/05/09
<<http://www.bksv.com/doc/bp2144.pdf>>
- [5] *Diccionario Acústico*
Fecha de consulta: 21/05/09
<<http://diccionario-audio.buscamix.com/>>
- [6] Océano Grupo Editorial
1995 “*Diccionario Enciclopédico Ilustrado*”
- [7] Davis, Gary
“*The Sound Reinforcement Handbook*”
1989, 2da ed.
- [8] Jon R. Sank
“*Microphones*”. J.Audio Eng. Soc
1985, Vol 33. N° 7/8, pp. 514-547
- [9] Brüel & Kjær Company
1982 “*Technical Review: Sound Intensity*”
Fecha de consulta: 22/07/09
<<http://www.bksv.com/doc/TechnicalReview1982-3.pdf>>

APÉNDICE A

Intensidad sonora es un vector que describe la cantidad y la dirección del flujo de energía acústica en una posición determinada, su unidad es W/m^2 . La intensidad sonora es igual al promedio en el tiempo del producto de la velocidad de partículas instantánea y la presión instantánea en la misma posición (ecuación (A.1))

$$I = \overline{p \cdot u} \quad (A.1)$$

donde:

I : Intensidad sonora

p : Presión sonora

u : Velocidad de partículas

La presión sonora se puede medir con cualquier tipo de micrófono, sin embargo medir la velocidad de partículas directamente no es sencillo. Es por este motivo que se asocia la velocidad de partículas a un gradiente de presión con la ecuación linealizada de Euler (A.2), la cual es la segunda ley de Newton aplicada a fluidos.

$$a = -\frac{1}{\rho} \text{grad } p \quad (A.2)$$

donde:

a : Aceleración

ρ : Densidad

p : Presión sonora

La ecuación (A.2) dice que la aceleración es igual a un gradiente de presiones dividido por la densidad del fluido en el cual las ondas se propagan, esta ecuación también se puede escribir de la siguiente manera:

$$\frac{\partial u}{\partial t} = -\frac{1}{\rho} \frac{\partial p}{\partial r} \quad (A.3)$$

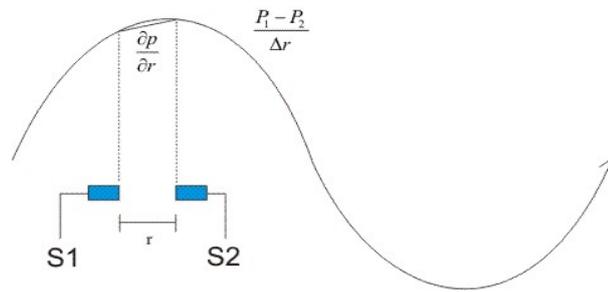
$$u = -\frac{1}{\rho} \int \frac{P_1 - P_2}{\Delta r} dt \quad (A.4)$$

donde:

u : Velocidad de partículas

P_1 : Presión sonora del micrófono 1

P_2 : Presión sonora del micrófono 2



Fuente: Creación propia

Figura 28

Modelo de Intensidad Sonora

También es necesario hallar la presión promedio entre los sensores 1 y 2 de la Figura 28, esto se realiza por medio de la ecuación (A.5).

$$p = \frac{P_1 + P_2}{2} \quad (\text{A.5})$$

Finalmente reemplazando las ecuaciones (A.4) y (A.5) en la ecuación (A.1) se tiene que intensidad sonora es:

$$I = -\frac{P_1 + P_2}{2\rho\Delta r} \int (P_1 - P_2) dt \quad (\text{A.6})$$

donde:

Δr : Distancia de separación entre los micrófonos

Esta ecuación se puede resolver por medio de filtros e integradores de manera analógica o utilizando un analizador de FFT y transformando la ecuación (A.6) a la ecuación (A.7).

$$I = -\frac{1}{\rho w \Delta r} \text{Im} G_{12} \quad (\text{A.7})$$

donde:

w: frecuencia angular

$\text{Im} G_{12}$: es la parte imaginaria del espectro de correlación de los puntos 1 y 2.

Apéndice B

Código en Matlab para la corrección de la fase de dos micrófonos

```

clear all
close all
clc
[A,Fs,Nbits]=wavread('C:\universidad\Tesis\fase\Barrido 5k.wav');
A=double(A);
I=A(:,1);
D=A(:,2);      %Se lee la señal proveniente del archivo *.wav
L=length(I);
NFFT = 2^nextpow2(L); % Se obtiene la potencia de 2 más cercana al
                    % número de datos de la señal
Y = fft(I,NFFT)/L;   % Se obtiene el auto-espectro de la señal por
                    % medio de la FFT
f = Fs/2*linspace(0,1,NFFT/2+1);

% Se gráfica la magnitud y fase de la señal.
magnitud=2*abs(Y(1:NFFT/2+1));
subplot(2,1,1);
plot(f,2*abs(Y(1:NFFT/2+1)));
title('Magnitud');
xlabel('Frecuencia (Hz)');
ylabel('Magnitud');
grid

fase=180/pi*angle(Y(1:NFFT/2+1));
subplot(2,1,2);
plot(f,180/pi*angle(Y(1:NFFT/2+1)));
title('Fase');
xlabel('Frecuencia (Hz)');
ylabel('Angulo (°)');
grid

%%
%Creación del nuevo Vector fase de corrección de fase
%Se logra al leer los datos obtenidos en el ensayo del laboratorio y
%realizando la interpolación correspondiente.
[datos,txt,raw]=xlsread('C:\universidad\Tesis\correcciones\Diferencia
a de fases.xml');
const=datos(1);
partida=find(f>const,1);
limite=length(f);
ndatos=zeros(limite-partida,1);
%Interpolación simple
for j=partida:limite

    valor=f(j);
    estado=find(datos(:,1)>f(j),1);
    aux=-(datos(estado)-valor)/(datos(estado)-datos(estado-1));
    ndatos(j)=aux*(datos(estado,4)-datos(estado-
1,4))+datos(estado,4);
end
%Se grafica la fase del archivo de corrección original y lo obtenido
%por la interpolación
figure
subplot(2,1,1)
plot(datos(:,1),datos(:,4))
axis([0 5000 -60 60]);
title('Fase con Datos Originales');

```

```

xlabel('Frecuencia (Hz)');
ylabel('Angulo (°)');
grid
subplot(2,1,2)
plot(f,ndatos)
axis([0 5000 -60 60]);
title('Fase con Datos Interpolados');
xlabel('Frecuencia (Hz)');
ylabel('Angulo (°)');
grid
%%
%CORRECCION DE FASE
nfase=fase-ndatos; %Se restan las fases de la señal a corregir con
                  %los datos obtenidos al laboratorio, para
                  %obtener un nuevo vector con fase corregida

Z=zeros(limite,1);
%Se crea el nuevo vector de la señal a partir de la magnitud que no
%ha sido modificada y el nuevo vector de fase corregido
for j=1:limite
    Z(j)=magnitud(j)*L.*exp(1i*pi*nfase(j)/180);
end
sconst=ifft(Z,NFFT); %Se regresa al dominio del tiempo por medio
                    %de la Transformada Inversa de Fourier

sol(:,1)=sconst(1:length(I));
sol(:,2)=sconst(1:length(I));
%Se convierte la señal al formato .wav
wavwrite(sol,Fs,Nbits,'C:\universidad\Tesis\fase\tono prueba.wav')
%%
%Comprobación del Filtro
B=wavread('C:\universidad\Tesis\fase\tono prueba.wav');
J=B(:,1);
FJ = fft(B,NFFT)/L;
dif=zeros(L,1);
%Para obtener la respuesta en el filtro y saber que se ha corregido
%la fase como se desea, se divide la señal filtrada entre la señal
%original en el dominio de la frecuencia.
for k=1:L
    dif(k)=FJ(k)/Y(k);
end
%Se grafica la respuesta del filtro
figure
subplot(2,1,1);
plot(f,abs(dif(1:NFFT/2+1)))
axis([0 5000 0 2]);
title('Magnitud');
xlabel('Frecuencia (Hz)');
grid;
subplot(2,1,2);
plot(f,180/pi*angle(dif(1:NFFT/2+1)))
axis([0 5000 -60 60]);
title('Fase');
xlabel('Frecuencia (Hz)');
grid;

```

Apéndice C

Código en Matlab para la corrección de magnitud de un micrófono.

```

clear all
close all
clc
[A,Fs,Nbits]=wavread('C:\universidad\Tesis\fase\Barrido 5k.wav');
A=double(A);
I=A(:,1);
D=A(:,2); %Se lee la señal proveniente del archivo *.wav
L=length(I);
NFFT = 2^nextpow2(L); % Se obtiene la potencia de 2 más cercana al
                        % número de datos de la señal
Y = fft(I,NFFT)/L; % Se obtiene el auto-espectro de la señal por
                  % medio de la FFT
f = Fs/2*linspace(0,1,NFFT/2+1);

% Se gráfica la magnitud y fase de la señal.
magnitud=2*abs(Y(1:NFFT/2+1));
subplot(2,1,1);
plot(f,2*abs(Y(1:NFFT/2+1)));
title('Magnitud');
xlabel('Frecuencia (Hz)');
ylabel('Magnitud');
grid

fase=180/pi*angle(Y(1:NFFT/2+1));
subplot(2,1,2);
plot(f,180/pi*angle(Y(1:NFFT/2+1)));
title('Fase');
xlabel('Frecuencia (Hz)');
ylabel('Angulo (°)');
grid

%%
%Creación del nuevo Vector fase de corrección de magnitud
%Se logra al leer los datos obtenidos en el ensayo del laboratorio y
%realizando la interpolación correspondiente.
[datos,txt,raw]=xlsread('C:\universidad\Tesis\correcciones\magnitudc
h1.xml');
const=find(datos(:,1)>200,1);
const=datos(const);
partida=find(f>const,1);
limite=length(f);
ndatos=zeros(limite-partida,1);
%interpolación simple
for j=partida:limite
    valor=f(j);
    estado=find(datos(:,1)>f(j),1);
    aux=-(datos(estado)-valor)/(datos(estado)-datos(estado-1));
    ndatos(j)=aux*(datos(estado,5)-datos(estado-
1,5))+datos(estado,5);
end
%Graficamos la magnitud del archivo de corrección original y lo
%obtenido por la interpolación
figure
subplot(2,1,1)
plot(datos(:,1),datos(:,5))
axis([0 5000 -10 10]);
title('Magnitud con Datos Originales');

```

```

xlabel('Frecuencia (Hz)');
ylabel('Magnitud');
grid
subplot(2,1,2)
plot(f,ndatos)
axis([0 5000 -10 10]);
title('Magnitud con Datos Interpolados');
xlabel('Frecuencia (Hz)');
ylabel('Magnitud');
grid
%%
%%CORRECCION DE MAGNITUD
nmag=ones(limite,1);
lndatos=zeros(limite,1);
for j=1:limite
    lndatos(j)=20*log10(magnitud(j)/(20*10^-6)); %Se convierte la
% magnitud de la señal en dB
end
nmag=lndatos-ndatos; %Se restan los dB de la señal original con los
% datos obtenidos en el laboratorio, lo que
% equivale a dividirlos.

for j=1:limite
    nmag(j)=20*10^(-6)*10^(nmag(j)/20);
%Se regresan los valores en dB a su magnitud original.
end
Z=zeros(limite,1);
%Se crea el nuevo vector de la señal a partir de la magnitud
% corregida y la fase que no ha sido modificada.
for j=1:limite
    Z(j)=nmag(j)*L.*exp(1i*pi*fase(j)/180);
end
sconst=ifft(Z,NFFT);
sol(:,1)=sconst(1:length(I));
sol(:,2)=sconst(1:length(I));
%Se convierte la señal al formato .wav
wavwrite(sol,Fs,Nbits,'C:\universidad\Tesis\fase\tono prueba.wav')
%%
%Comprobación
B=wavread('C:\universidad\Tesis\fase\tono prueba.wav');
J=B(:,1);
FJ = fft(B,NFFT)/L;
dif=zeros(L,1);
%Para obtener la respuesta en el filtro y saber que se ha corregido
% la magnitud como se desea, se divide la señal filtrada entre la
% señal original en el dominio de la frecuencia.
for k=1:L
    dif(k)=FJ(k)/Y(k);
end
%Se grafica la respuesta en frecuencia del filtro.
figure
subplot(2,1,1);
plot(f,2*abs(dif(1:NFFT/2+1)))
axis([0 5000 0 10]);
title('Magnitud');
xlabel('Frecuencia (Hz)');
grid;
subplot(2,1,2);
plot(f,180/pi*angle(dif(1:NFFT/2+1)))
axis([0 5000 -60 60]);
title('Fase');
xlabel('Frecuencia (Hz)');
grid;

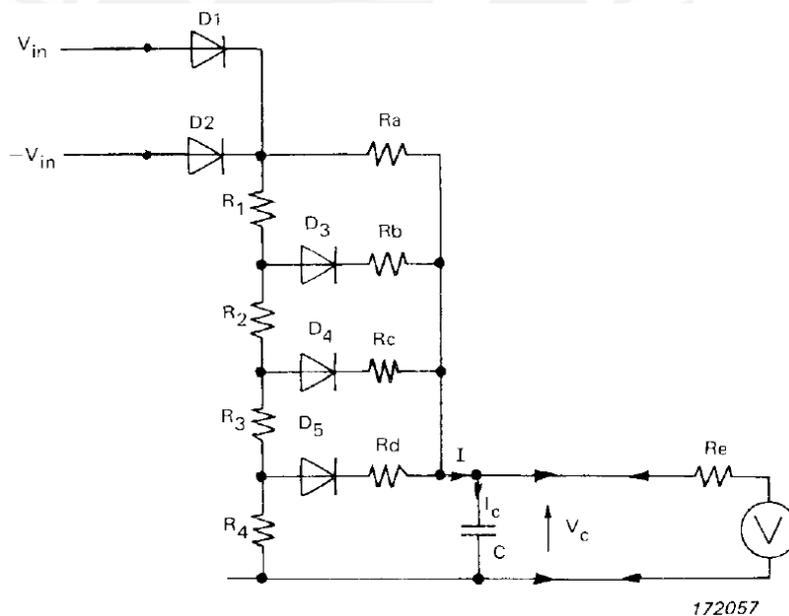
```

Apéndice D

El bloque RMS representa al valor cuadrático medio, conocido por sus siglas en inglés RMS (root mean square). El cual puede calcularse para una serie de valores discretos o para una función de variable continua. Para el caso de la tesis, se utilizará para una función de variable continua (presión sonora), este valor puede ser calculado según la fórmula (D.1) para un intervalo de $T_1 \leq t \leq T_2$.

$$x_{rms} = \sqrt{\frac{1}{T_2 - T_1} \int_{T_1}^{T_2} [f(t)]^2 dt} \quad (D.1)$$

Esta ecuación puede ser implementada por medio de circuitos analógicos, los cuales en su mayoría cumplen con el mismo principio fundamental. La figura 29 muestra el circuito ideal para dicha implementación.



Fuente: Technical Review No. 2 -1972 de B&K

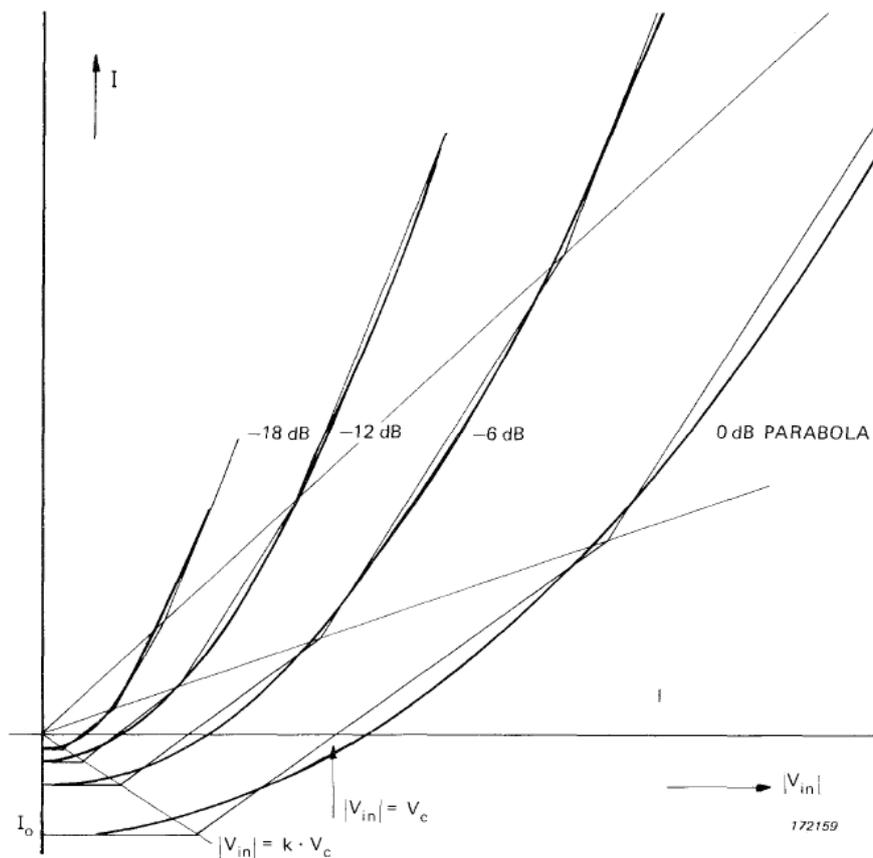
Figura 29

Diagrama esquemático del detector básico de RMS

El principio de funcionamiento del circuito mostrado en la figura 29 es descrito en el "Technical Review: To Advance Techniques in Acoustical, Electrical and Mechanical Measurement" No. 2 de 1972 de la compañía B&K. El cual indica lo siguiente:

Si el voltaje de entrada es inicialmente cero, una corriente negativa I_o fluirá a través de la resistencia $R_a + R_{1-4}$ resultando en un voltaje de polarización de los diodos D_1 y D_2 de kV_c , $k < 1$.

Si luego el valor numérico del voltaje de entrada incrementa, I permanecerá constante e igual a I_o , hasta que el voltaje alcance el valor de kV_c , entonces uno de los diodos D_1 o D_2 conduce e I incrementa linealmente con el incremento del voltaje de entrada (con una pendiente de la curva igual a $1/R_a$).



Fuente: Technical Review No. 2 -1972 de B&K

Figura 30

Los polígonos $I/|V_c|$

Cuando el voltaje de entrada $|V_{in}|=V_c$ entonces $I=0$ como indica la figura 30. Si el voltaje continua incrementando alcanzará un valor donde D_3 empiece a conducir, en donde la pendiente de la curva cambiará a $1/R_a + (1/R_b) (R_{2-4} / R_{1-4})$. Un mayor incremento de voltaje podría causar que D_4 y mucho después D_5 también conduzcan.

La curva resultante $I|V_{in}|$ es un polígono, el cual con una debida selección de las resistencias puede aproximarse a una parábola ligeramente desplazada sobre el eje x.

La carga del capacitor determina su voltaje. La carga depende de la corriente de pérdida I_m a través del medidor y de la corriente del circuito de ley de cuadrados de la entrada, la constante la cual es inversamente proporcional cambiada al voltaje del capacitor. Esta realimentación provee la raíz cuadrada necesaria para la obtención de un voltaje de capacitor linealmente dependiente del valor RMS del voltaje de entrada. La resistencia del medidor y del capacitor, conjunto con las otras resistencias, determinan el tiempo de integración del circuito rectificador total.

Este circuito posee muchas limitaciones, es solamente valido si la señal excede el valor de conducción de los diodos, esto va a depender del tipo de diodo que se utilice. Otra limitación es el factor máximo de cresta que puede ser correctamente medido, esto va a depender del diseño de la parábola, entre más grande sea este factor mayor precisión requerirá la parábola, lo que aumentaría el numero de diodos a utilizar. Además el promedio en el tiempo depende del valor del capacitor y de las resistencias que están conectadas a los diodos. Si se desea tiempos largos de promedio ambos valores de capacitor y de resistencias deben ser altos, lo cual es contraproducente puesto que los capacitores aumentan su tamaño mientras más alto sea su valor, y por otro lado valores altos de resistencias afectarían el desempeño de los diodos a los cuales están conectadas.

Sin embargo para efectos de la tesis, se ha utilizado un bloque RMS de la librería "Statistics" de "Signal and Processing Sources" de Simulink de Matlab. Este bloque permite obtener el valor RMS de una señal continua variable en el tiempo, bajo la fórmula (D.2).

$$y_j = \sqrt{\frac{\sum_{i=1}^M |u_{ij}|^2}{M}}, 1 \leq j \leq N \quad (D.2)$$

También se puede representar según el siguiente código (ec. (D.3)):

$$y = \text{sqrt}(\text{sum}(u.*\text{conj}(u))/\text{size}(u,1)) \quad (D.3)$$

Donde básicamente lo que se hace es aplicar la raíz cuadrada a la suma de los cuadrados de los elementos de una columna y dividirlos entre el tamaño de ésta, lo cual podría interpretarse como el valor RMS de la señal.



Apéndice E

Código en Matlab del programa final para la localización de la mayor intensidad de la voz de un orador en un plano.

```

close all
clear all
clc
referencial=3.0011*10^-9; %Tiene el valor de referencia para el
                        %filtro de 250Hz
referencia2=5.5302*10^-7; %Tiene el valor de referencia para el
                        %filtro de 500Hz

fis=250;
%Se lee la señal y se guarda en una variable, en L se guardan los
%datos almacenados en el canal izquierdo y en R se guardan los datos
%almacenados en el canal derecho
[A,Fs,Nbits]=wavread('C:\universidad\Tesis\pruebas\22-10-
10\Direccionada\laboratorio\final 5\70.wav');
A=double(A);
R=A(:,2);
L=A(:,1);
%A continuación se guardan las constantes que se van a utilizar
%tanto para las funciones como también para las distintas subrutinas
window=2096;
nfft>window;
distancia_sens=0.013; %separación de los sensores en metros
densidad=1.2; %densidad del aire
TRO=0.25; %tiempo de la región de interés
delta=3; %tiempo deseado de análisis en segundos
delta=delta/TRO;
N=length(A);
tiempo=(0:1/Fs:(N-1)/Fs);
angulo2=0;
%FILTRADO DE LA SEÑAL
%Se utiliza un filtro digital pasa-banda con un ancho de banda de
%100Hz con frecuencia fundamental de fis Hz
H=firls(400,[0 (fis-100) (fis-50) (fis+50) (fis+100) 8000]/8000,[0 0
1 1 0 0]);
L=conv(L,H);
R=conv(R,H);

a(:,1)=(0:1/Fs:TRO); %Se crea un vector de tiempo para el
                    %procesamiento

b=a;
for num=1:delta
tempo=TRO*num;
%num=1;
%tempo=TRO;
aux1=floor(tempo*Fs);
aux2=floor((tempo+TRO)*Fs);

a(:,2)=(L(aux1:aux2));
b(:,2)=(R(aux1:aux2)); %en a y b se guardan los valores a
                    %analizar

sim('C:\universidad\Tesis\ultima\conv_rms.mdl',TRO);
%El bloque conv_rms.mdl convierte los valores de a y b en valores
RMS.
%En suma se tiene el valor promedio de la presión sonora RMS que hay
entre los dos micrófonos
suma=(arms+brms)/2;

```

```

resta(:,1)=(0:1/Fs:TRO); %Se realiza la resta de las
                           %presiones
resta(:,2)=-(b(:,2)-a(:,2));

%DETERMINACION DEL SIGNO
signo=brms-arms;
signo=sum(signo)/length(signo);
if signo>0
    bandera=-1; % La variable bandera determina si el sonido
                %proviene de la derecha o de la izquierda
                % Si bandera=-1 el sonido viene del lado derecho
                % Si bandera=1 el sonido proviene del lado
                %izquierdo
else
    bandera=1;
end

sim('C:\universidad\Tesis\ultima\integral.mdl',TRO);
%En el bloque integral.mdl se realiza la integral de la señal resta.
vel(:,1)=(0:1/Fs:TRO);
vel(:,2)=velocidad; %en vel se guarda la velocidad
                    %hallada por medio de la integral de
                    %la diferencia de presiones entre
                    %los dos micrófonos

sim('C:\universidad\Tesis\ultima\vel_rms.mdl',TRO);
%El bloque vel_rms.mdl obtiene el valor RMS de la integral, con lo
%que se obtiene el valor de velocidad de partículas en RMS.
intensidad=velo.*suma; %Se obtiene la intensidad sonora al
                       %multiplicar la
                       %velocidad por el valor promedio de
                       %presiones

%Se realiza el promedio en el tiempo para hallar la intensidad
%sonora
ayuda=length(intensidad);
total=(intensidad(1)+intensidad(2))/2;
for k=3:ayuda
    total=(total+intensidad(k))/2;
end
if total>0
    total=total*-1;
end
%En final se tiene la intensidad sonora en dB
final(num)=-total/(2*densidad*distancia_sens);
finalog(num)=bandera*10*log10(final(num)/10^-12);

end
%Filtrado de información, consiste en eliminar el valor más alto y
%más bajo de las muestras para mejorar el promedio hecho en el
%tiempo.
maximo=max(finalog);
minimo=min(finalog);
resultado1=(sum(finalog)-maximo-minimo)/(length(finalog)-2);
if resultado1<0
    oj1=10^(-resultado1/10)*10^(-12);
else
    oj1=10^(resultado1/10)*10^(-12);
end

angulo1=90-asind(oj1/referencial);

```

%En angulo1 se obtiene el valor del ángulo obtenido por el primer filtro con el cual se hace la división de las 3 primeras zonas: La primera entre 0° y 30°, la segunda entre 30° y 60°, y la tercera entre 60° y 90°.

```
if angulo1<31
    zona=1;
else
    if angulo1<61
        zona=2;
    else
        zona=4;
    end
end
```

%En el caso que la zona determinada sea la segunda se pasa a un segundo proceso de filtrado con un valor del filtro de 500Hz como frecuencia central, a partir de este punto se repite el procedimiento de la parte anterior.

```
if zona==2
```

```
    fis=500;
    R=A(:,2);
    L=A(:,1);
```

%FILTRADO DE LA SEÑAL

%Se utiliza un filtro digital pasa-banda con un ancho de banda de 100Hz con frecuencia fundamental de fis Hz

```
H=firls(400,[0 (fis-100) (fis-50) (fis+50) (fis+100) 8000]/8000,[0 0
1 1 0 0]);
L=conv(L,H);
R=conv(R,H);
```

%CALCULO DE LA INTENSIDAD

```
a(:,1)=(0:1/Fs:TRO);
b=a;
```

```
for num=1:delta
    tempo=TRO*num;
```

```
aux1=floor(tempo*Fs);
aux2=floor((tempo+TRO)*Fs);
```

```
a(:,2)=(L(aux1:aux2));
b(:,2)=(R(aux1:aux2));
sim('C:\universidad\Tesis\ultima\conv_rms.mdl',TRO);
suma=(arms+brms)/2;
resta(:,1)=(0:1/Fs:TRO);
resta(:,2)=-(b(:,2)-a(:,2));
```

%DETERMINACION DEL SIGNO

```
signo=brms-arms;
signo=sum(signo)/length(signo);
```

```
if signo>0
    bandera=-1;
```

```
else
    bandera=1;
```

```
end
```

```

sim('C:\universidad\Tesis\ultima\integral.mdl',TRO);
vel(:,1)=(0:1/Fs:TRO);
vel(:,2)=velocidad;

sim('C:\universidad\Tesis\ultima\vel_rms.mdl',TRO);
intensidad=velo.*suma;
ayuda=length(intensidad);
total=(intensidad(1)+intensidad(2))/2;
for k=3:ayuda
    total=(total+intensidad(k))/2;
end
if total>0
    total=total*-1;
end
final(num)=-total/(2*densidad*distancia_sens);
finalog(num)=bandera*10*log10(final(num)/10^-12);

end
%Filtrado de información
maximo=max(finalog);
minimo=min(finalog);
resultado2=(sum(finalog)-maximo-minimo)/(length(finalog)-2);
if resultado2<0
    ojo2=10^(-resultado2/10)*10^(-12);
else
    ojo2=10^(resultado2/10)*10^(-12);
end
angulo2=90-asind(ojo2/referencia2);
%En ángulo2 se obtiene el ángulo por el segundo filtro, el cual
%ayuda a discernir si la voz se encuentra en la zona 2(30° a 45°) o
%la zona 3(45° a 60°).

    if angulo2>46
        zona=3;
    else
        zona=2;
    end

end
%Se muestran los resultados en pantalla.
bandera*zona
bandera*angulo1
bandera*angulo2

```