

RAE

1. **TIPO DE DOCUMENTO:** Trabajo de grado para optar por el título de INGENIERO DE SONIDO.
2. **TÍTULO:** Diseño e implementación de una aplicación móvil para simulación de entorno de escucha estereofónica a través de audífonos.
3. **AUTORES:** Jhon Freddy Barreto Leguízamo, Andrés Mateo Bejarano Garzón, Yonathan Camilo Layton González.
4. **LUGAR:** Bogotá D.C.
5. **FECHA:** Enero de 2020.
6. **PALABRAS CLAVE:** Imagen estereofónica, simulación de crosstalk, calidad sonora, procesamiento digital.
7. **DESCRIPCIÓN DEL TRABAJO:** El presente es un trabajo de grado para optar al título de Ingeniero de Sonido, cuyo objetivo fue diseñar una aplicación móvil que implementara dos métodos diferentes de simulación de crosstalk en audífonos y determinar el impacto de dichos métodos de procesamiento en la calidad sonora de pistas de audio comercial reproducidas desde la aplicación y escuchadas a través de audífonos.
8. **LÍNEAS DE INVESTIGACIÓN:** Línea de Investigación de la USB: Acústica.
9. **METODOLOGÍA:** El enfoque de esta investigación es empírico-analítico, ya que se busca diseñar un sistema de simulación de crosstalk para dispositivos móviles Android y, según los antecedentes, este tipo de aplicativos no han sido desarrollados a profundidad para esta plataforma. Se pretende observar los resultados de las evaluaciones subjetivas para medir el impacto que el procesamiento realizado por la aplicación tiene sobre la calidad de la imagen estereofónica en pistas de audio escuchadas a través de audífonos. Las pruebas de calidad sonora fueron de dos tipos: evaluación de preferencia con pares comparados y evaluaciones de atributos sonoros mediante una escala semántica.
10. **CONCLUSIONES:** La aplicación móvil desarrollada en Unity permitió la lectura, procesamiento y reproducción de archivos de audio en formato WAV presentes en la memoria interna del dispositivo. Fue posible implementar en la aplicación los dos métodos de simulación de crosstalk de mayor preferencia en la primera etapa. Los resultados obtenidos en las evaluaciones de calidad sonora de la aplicación desarrollada indicaron que los métodos de simulación de crosstalk implementados no mejoraron la calidad sonora del audio original de una manera significativa. Si bien, observando el consolidado de datos obtenidos con estudiantes, la calidad general del audio procesado mediante el método de Manor fue mayor que en los demás, estas diferencias no fueron estadísticamente significativas. Para oyentes de mayor experiencia, el audio original fue el de mayor calidad general en cada uno de los atributos evaluados.

Diseño e Implementación de una Aplicación Móvil para Simulación de Entorno de Escucha Estereofónica a través de Audífonos

Jhon Freddy Barreto Leguízamo

Andrés Mateo Bejarano Garzón

Yonathan Camilo Layton González

Universidad de San Buenaventura, sede Bogotá

Facultad de Ingeniería

Programa de Ingeniería de Sonido

Bogotá, Colombia

2020

**Diseño e Implementación de una Aplicación Móvil para Simulación de Entorno de
Escucha Estereofónica a través de Audífonos**

Jhon Freddy Barreto Leguízamo

Andrés Mateo Bejarano Garzón

Yonathan Camilo Layton González

Trabajo de grado presentado como requisito para optar al título de Ingeniero de Sonido

Tutor: Luis Fernando Hermida Cadena

Línea de Investigación:

Calidad Sonora

Universidad de San Buenaventura, sede Bogotá

Facultad de Ingeniería

Programa de Ingeniería de Sonido

Bogotá, Colombia

2020

Notas de aceptación

Firma del presidente del Jurado

Firma del jurado

Firma del jurado

Bogotá, enero de 2020.

Agradecimientos

Agradezco en primer lugar a mis padres Rafael Barreto y Beatriz Leguizamo por apoyarme de manera emocional y económica lo largo de esta carrera, en las decisiones que tuve que tomar, los imprevistos que surgieron y obstáculos que solo nosotros sabemos lo duro que pudieron llegar a ser, les estaré eternamente agradecido por la vida, por cada día confiar y creer en mí y en mis expectativas, ya que sin ellos no hubiese sido posible cumplir esta etapa de mi vida. Un agradecimiento también a los profesores que me brindaron conocimiento sobre la vida y se convirtieron en una inspiración para este período. A mis amigos, que me brindaron su apoyo, amistad y consejo en los momentos necesarios. Finalmente, un agradecimiento especial a Laura Sánchez que, en momentos de cansancio y preocupación, estuvo ahí para apoyarme, creer en mí y recordarme la razón y pasión con la que emprendí esta meta.

Gracias a la vida por este nuevo triunfo y gracias a todas las personas que me apoyaron y creyeron en la realización de esta tesis.

Jhon Fredy Barreto Leguízamo.

Agradezco a mis padres Harry Bejarano y Nidia Garzón por todo su apoyo durante el transcurso de mi carrera, brindándome todo de manera incondicional. Les dedico todo el empeño y dedicación que puse para este proyecto. Son ustedes mi mayor motivación para lograr grandes cosas en la vida, entre ellas mi título de Ingeniero de Sonido. A mi abuela Oliva Morales, por haberme apoyado en los momentos más difíciles para mi hogar, permitiéndome llegar a donde estoy. A mis compañeros Jhon Barreto y Yonathan Layton por su arduo trabajo y desempeño, no solo en el desarrollo de este trabajo sino en cada una de las materias en las cuales tuve el privilegio de trabajar a su lado. Más que ingenieros, son personas ejemplares y no pude haber estado en un mejor grupo. A Luz Marina Lazcano, por recibirme como a un hijo y darme el regalo más valioso, enseñándome vivir siempre desde el silencio interior y la unidad. Finalmente, a todas las personas que creyeron en mí y que me ayudaron a ser el hombre que soy ahora.

Andrés Mateo Bejarano Garzón.

Agradezco a mis padres (Oscar Layton y Diocelina González) por la oportunidad de darme una formación tanto en los primeros años hasta mi formación de pregrado, sin dejar de lado los valores y principios impuestos por ellos. Un agradecimiento importante para mis amigos (Mateo Bejarano y Jhon Barreto), quienes contribuyeron con la mayor disposición para que este proyecto se llevara a cabalidad.

Yonathan Camilo Layton González.

Agradecemos a todos los profesores, estudiantes e ingenieros que contribuyeron con el desarrollo de este proyecto brindándonos su tiempo en la participación de las pruebas de evaluación. A nuestro tutor Luis Fernando Hermida, por su dirección, orientación, dedicación, apoyo y conocimiento brindado para el desarrollo correcto de este trabajo. Ha sido un placer para nosotros poder trabajar en este proyecto bajo su dirección. Agradecemos a Andrés Araujo, profesional de la dirección de investigación, por sus excelentes críticas constructivas en la redacción de este documento, análisis cualitativos de la muestra y sus diferentes puntos de vistas en la ejecución del proyecto. A Manuel Torres, director del programa de ingeniería de sonido, por su apoyo para el desarrollo de las pruebas de preferencia complementarias. Finalmente, a Darío Páez por ser de gran ayuda en la implementación de unos de los métodos de simulación.

Jhon Freddy Barreto Leguízamo,

Andrés Mateo Bejarano Garzón,

Yonathan Camilo Layton González.

Resumen

La manera en la que los atributos sonoros en un producto musical son interpretados depende del dispositivo de reproducción. La imagen estereofónica de una pista de audio es percibida de manera distinta al ser escuchada mediante altavoces o audífonos. La gran mayoría de trabajos musicales son producidos con ayuda de un par de altavoces. Debido a esto, se han desarrollado varios métodos de procesamiento digital con el fin de simular este entorno de escucha en audífonos. Este proyecto se realizó en tres etapas. En primer lugar, se llevó a cabo una prueba de preferencia mediante pares comparados entre 4 métodos de simulación de crosstalk. Los dos métodos de mayor preferencia fueron implementados en una aplicación móvil. Finalmente, se evaluaron cuatro atributos sonoros, mediante una escala semántica, de tres pistas de audio reproducidas a través de la aplicación desarrollada. Dichas pruebas se realizaron con dos grupos de evaluadores, uno con mayor experiencia en el audio que el otro. Los resultados indican que los métodos de simulación de crosstalk implementados deterioraron el timbre del audio original. Además, se hallaron diferencias entre los tipos de prueba, ya que el método denominado “earGoggles” tuvo la mayor preferencia mediante pares comparados, mientras que el audio original tuvo las calificaciones más altas de calidad general en las evaluaciones mediante la escala semántica.

Palabras clave: Imagen estereofónica, Simulación de crosstalk, Calidad sonora, Procesamiento digital, Pares comparados, Escala semántica.

Abstract

The way sound attributes in a musical product are interpreted depends on the playback device. The stereo imaging of an audio track is perceived differently when heard through speakers or headphones. Most musical products are produced using a pair of speakers. Therefore, several digital processing methods have been developed in order to simulate this listening environment in headphones. This project was divided in three stages. First, a preference test was carried out using the compared pairs method between four crosstalk simulation methods. The two most preferred methods were implemented in a mobile application. Finally, four sound attributes were evaluated, using a semantic scale, for three audio tracks reproduced in the developed application. These tests were conducted with two juries, one with more audio experience than the other. Results showed that crosstalk simulation methods damaged the timbre of the original audio. In addition, differences were found between the types of tests, since the method called “earGoggles” had the highest preference ratings through compared pairs, while the original audio had the highest overall quality scores in the evaluations using the semantic scale.

Keywords: Stereo imaging, crosstalk simulation, sound quality, digital processing, Compared pairs, Semantic scale.

Contenido

1	Planteamiento del problema	3
1.1	Descripción de la situación actual	3
1.1.1	Consumo digital de música	3
1.1.2	Uso de dispositivos móviles	7
1.1.3	Uso de audífonos	7
1.1.4	Problemática Técnica	7
1.1.5	Pregunta Problema	9
1.2	Estado del Arte	9
1.2.1	Antecedentes locales	9
1.2.2	Antecedentes nacionales	10
1.2.3	Antecedentes internacionales	10
1.3	Objetivos	13
1.3.1	Objetivo General	13
1.3.2	Objetivos Específicos	13
1.4	Justificación.....	13
1.5	Alcances y Limitaciones	15
1.5.1	Alcances	15
1.5.2	Limitaciones	15
2	Marcos de Referencia.....	16
2.1	Marco Conceptual	16
2.1.1	Percepción Binaural	16
2.1.2	Imagen Estereofónica	18
2.1.3	Crosstalk.....	19
2.1.4	Calidad Sonora	21
2.1.5	Procesamiento Digital de Señales	22
2.1.6	Análisis Estadístico de Datos	27
2.2	Marco Teórico	31
2.2.1	Métodos de Simulación de Crosstalk	31
2.2.2	Programación en Unity.....	34
2.3	Marco Normativo	35
2.3.1	RECOMMENDATION ITU-R BS.1284-1: General methods for the subjective assessment of sound quality.	35
3	Metodología	38

3.1	Enfoque de Investigación	38
3.2	Etapas de la Metodología	38
3.2.1	Evaluación de los métodos	38
3.2.2	Desarrollo de la aplicación móvil.....	39
3.2.3	Evaluación del producto.....	39
3.3	Variables	40
3.3.1	Variables dependientes.....	40
3.3.2	Variable independiente.....	40
3.4	Hipótesis.....	40
4	Desarrollo Ingenieril	41
4.1	Evaluación de Preferencia para Métodos de Simulación de Crosstalk	41
4.1.1	Desarrollo y aplicación de los métodos de simulación de crosstalk.....	41
4.1.2	Diseño y aplicación de pruebas de preferencia	53
4.1.3	Aplicación de la Ley de Juicio Comparativo	56
4.2	Desarrollo de la aplicación móvil.....	60
4.2.1	Plataforma de desarrollo.....	60
4.2.2	Estructura del software.....	61
4.2.3	Funcionamiento de la aplicación.....	64
4.3	Evaluaciones calidad sonora de la aplicación móvil	67
4.4	Evaluaciones calidad sonora adicionales a oyentes expertos	69
5	Discusión y Análisis de Resultados.....	71
5.1	Resultados de las pruebas de preferencia de los métodos de simulación de crosstalk	71
5.1.1	Resultados de pruebas de preferencia para Symphony No. 5	71
5.1.2	Resultados de pruebas de preferencia para Money	71
5.1.3	Resultados de pruebas de preferencia para la canción “Con Calma”	72
5.1.4	Análisis del consolidado de datos para las pruebas de preferencia	74
5.2	Resultados de las pruebas de calidad sonora de la aplicación móvil.....	75
5.2.1	Resultados de pruebas de calidad sonora de la aplicación móvil realizadas a estudiantes	75
5.2.2	Resultados de pruebas de calidad sonora de la aplicación móvil realizadas a oyentes expertos	85
5.2.3	Comparación de pruebas de calidad sonora entre estudiantes y oyentes expertos. ...	92
6	Conclusiones y recomendaciones.....	95
6.1	Conclusiones	95
6.2	Recomendaciones a futuro	99

7	Referencias.....	101
8	Anexos.....	105
8.1	Anexo A	105
8.1.1	Tablas de distribución normal Z.....	105
8.1.2	Protocolo para pruebas de preferencia	107
8.1.3	Consentimiento de la prueba de preferencia por medio de pares comparados.....	108
8.1.4	Resultados y aplicación de la ley Thurstone para prueba de preferencia.....	109
8.2	Anexo B: Anexos de prueba de calidad sonora de la aplicación móvil.	114
8.2.1	Consentimiento informado para participantes de prueba de calidad sonora	114
8.2.2	Protocolo de aplicación para pruebas de calidad sonora	115
8.2.3	Imágenes de Google Form para prueba de calidad sonora.....	116
8.3	Anexo C	117

Lista de Figuras

Figura 1. Crecimiento de los ingresos del streaming entre 2013 y 2017. Recuperado de: (International Federation of the Phonographic Industry, 2018),.....	4
Figura 2. Crecimiento de los ingresos por la industria entre 1999 y 2017. Recuperado de (International Federation of the Phonographic Industry, 2018),.....	5
Figura 3. Ingresos segmentados en el año 2017. Recuperado de (International Federation of the Phonographic Industry, 2018),.....	6
Figura 4. Efecto de la separación de los oídos y la sombra de la cabeza en el proceso de percepción binaural. Recuperado de: (Willert, 2006),.....	16
Figura 5. Diferencia de nivel entre cada oído en función de la frecuencia. Recuperado de (Zölzer, 2002),.....	17
Figura 6. Simulación de espacialidad sonora a partir de técnicas de microfoneo estéreo. Recuperado de: (Bartlett, 1991),.....	19
Figura 7. Diferencia entre escucha estereofónica (izquierda), bifónica (centro) y bifónica con simulación de crosstalk (derecha). Recuperado de: (Manor, Martens, & Cabrera, 2012),.....	21
Figura 8. Representación una señal analógica $x(t)$ a través de una señal digital $x(n)$ a intervalos de tiempo T . Recuperado de: (Zölzer, 2002),.....	23
Figura 9. Diagrama de bloques de un filtro digital. Recuperado de (Kuo, Lee, & Tian, 2006),.....	24
Figura 10. Respuesta en frecuencia de los filtros para canal directo (izquierda) y canal opuesto (derecha) del método Balanced Stereo Widening Network. Recuperado de: (Kirkeby, 2002),.....	32
Figura 11. Respuesta en frecuencia de los filtros para campo cercano y campo lejano. Recuperado de: (Manor E. , Martens, Marui, & Cabrera, 2015),.....	32
Figura 12. Ciclo de desarrollo de software. Recuperado de: (Ortega, Orjuela, & Cortes, 2008).....	39
Figura 13. Diagrama de bloques de procesamiento digital mediante el método de Kirkeby.	42
Figura 14. Respuesta en frecuencia (azul) y fase (rojo) de filtro FIR pasa-bajas diseñado en MATLAB.....	43
Figura 15. Filtrado de señal a partir de filtro FIR diseñado, implementando el comando <i>filter()</i> de MATLAB.....	43
Figura 16. Generación de canales originales retrasados 112 muestras.	44
Figura 17. Retraso de 0.7ms de canales filtrados.....	44
Figura 18. Creación de canales de salida para audio estereofónico procesado mediante el método de Kirkeby.....	45
Figura 19. Diagrama de bloques de procesamiento digital mediante el método de Manor.	45
Figura 20. Diagrama de bloques de procesamiento digital mediante el método de Gilchrest.	48
Figura 21. Diagrama de bloques de procesamiento digital mediante el método de Bolaños.....	50
Figura 22. Ubicación de la cabeza binaural en el sweetspot del estudio de Mastering.....	50
Figura 23. Ubicación de un micrófono DPA dentro de la oreja de la cabeza binaural.	51

Figura 24. Diagrama de conexionado empleado para la medición de respuesta al impulso binaural dentro del estudio de Mastering.	52
Figura 25. Respuestas al impulso binaurales obtenidas de la medición.....	52
Figura 26. Respuestas al impulso binaurales ecualizadas por un oyente experto.	53
Figura 27. Ejemplo de prueba de pares comparados en la interfaz gráfica desarrollada en MATLAB.....	55
Figura 28. Diagrama de conexionado empleado para la realización de las pruebas de preferencia entre métodos de simulación de crosstalk.	56
Figura 29. Resultado del escalamiento de preferencia para el tema Symphony No. 5.	59
Figura 30. Diagrama de flujo de la aplicación desarrollada.....	61
Figura 31. Clases implementadas en el GameObject 'AudioPlayer'.	62
Figura 32. Interfaz gráfica de la aplicación desarrollada.	63
Figura 33. Diagrama de bloques para las clases implementadas en la aplicación.	64
Figura 34. Diagrama de conexionado de la prueba de calidad sonora.	68
Figura 35. Oyente experto realizando prueba de calidad sonora.	70
Figura 36. Resultado del escalamiento para el tema Money.....	72
Figura 37. Resultado del escalamiento para el tema Con Calma.	73
Figura 38. Escala de preferencia obtenida de los datos consolidadas para las tres canciones evaluadas.....	74
Figura 39. Instrucciones en formulario Google Form para las pruebas de evaluación de calidad sonora.....	116
Figura 40. Instrucciones para uso de la aplicación móvil en el formulario de Google Form.	116
Figura 41. Ejemplo de evaluación de calidad sonora en formulario Google Form.....	117

Lista de Tablas

Tabla 1. Matriz de frecuencia observadas de los diferentes métodos de procesamiento para el tema Symphony No. 5.	57
Tabla 2. Matriz de proporciones ordenada de los datos para el tema Symphony No. 5.....	58
Tabla 3. Matriz de proporciones normalizadas Z y escala de preferencia (gris) para la canción Symphony No. 5.	59
Tabla 4. Panel adicional de oyentes expertos.....	70
Tabla 5. Prueba de Shapiro-Wilk para los datos obtenidos con estudiantes.	76
Tabla 6. Estadísticos descriptivos de las evaluaciones de calidad sonora con estudiantes para el tema Symphony No. 5.	77
Tabla 7. Estadísticos de Kruskal-Wallis de las evaluaciones de calidad sonora con estudiantes para el tema Symphony No. 5.	78
Tabla 8. Estadísticos descriptivos de las evaluaciones de calidad sonora con estudiantes para el tema Money.	79
Tabla 9. Estadísticos de Kruskal-Wallis de las evaluaciones de calidad sonora con estudiantes para el tema Money.	80
Tabla 10. Estadísticos descriptivos de las evaluaciones de calidad sonora con estudiantes para el tema Con Calma.	81
Tabla 11. Estadísticos de Kruskal-Wallis de las evaluaciones de calidad sonora con estudiantes para el tema Con Calma.	82
Tabla 12. Estadísticos descriptivos de las evaluaciones de calidad sonora con estudiantes para los tres temas evaluados.	83
Tabla 13. Estadísticos de Kruskal-Wallis de las evaluaciones de calidad sonora con estudiantes para los tres temas evaluados.....	84
Tabla 14. Prueba de Shapiro-Wilk para los datos obtenidos con oyentes expertos.	85
Tabla 15. Estadísticos descriptivos de las evaluaciones de calidad sonora con oyentes expertos para el tema Symphony No 5.....	86
Tabla 16. Estadísticos de Kruskal-Wallis de las evaluaciones de calidad sonora con oyentes expertos para el tema Symphony No. 5.	87
Tabla 17. Estadísticos descriptivos de las evaluaciones de calidad sonora con oyentes expertos para el tema Money.	88
Tabla 18. Estadísticos de Kruskal-Wallis de las evaluaciones de calidad sonora con oyentes expertos para el tema Money.	89
Tabla 19. Estadísticos descriptivos de las evaluaciones de calidad sonora con oyentes expertos para el tema Con Calma.....	90
Tabla 20. Estadísticos de Kruskal-Wallis de las evaluaciones de calidad sonora con oyentes expertos para el tema Con Calma.....	90
Tabla 21. Estadísticos descriptivos de las evaluaciones de calidad sonora con oyentes expertos para los tres temas.....	91
Tabla 22. Estadísticos de Kruskal-Wallis de las evaluaciones de calidad sonora con oyentes expertos para las tres canciones.	92

Tabla 23. Comparación de medias y desviación entre estudiantes y oyentes expertos	93
Tabla 24. Comparación de estadísticos de Kruskal-Wallis entre datos consolidados de estudiantes y oyentes expertos	94
Tabla 25. Tabla de protocolo de aplicación de pruebas de preferencia.....	107
Tabla 26. Matriz de frecuencia observada de los diferentes métodos de procesamiento para el tema Symphony No 5.	109
Tabla 27. Matriz de proporciones ordenada de los datos para el tema Symphony No. 5.....	109
Tabla 28. Matriz de proporciones normalizadas Z para tema Symphony No. 5.	110
Tabla 29. Matriz de frecuencia observada de los diferentes métodos de procesamiento para el tema Con Calma.	110
Tabla 30. Matriz de proporciones ordenada de los datos para Con Calma.	110
Tabla 31. Matriz de proporciones normalizadas Z para el tema Con Calma.	111
Tabla 32. Matriz de frecuencia observada de los diferentes métodos de procesamiento para el tema Money.	111
Tabla 33. Matriz de proporciones ordenada de los datos para el tema Money.	111
Tabla 34. Matriz de proporciones normalizadas Z para el tema Money.	112
Tabla 35. Matriz de frecuencia observada de los diferentes métodos de procesamiento para las tres canciones	112
Tabla 36. Matriz de proporciones ordenada de los datos consolidados para los tres temas.....	112
Tabla 37. Matriz de proporciones normalizadas Z para los tres temas.	113
Tabla 38. Procedimiento para aplicación de pruebas de calidad sonora de la aplicación móvil.	115

Lista de Ecuaciones

Ecuación 1. Ecuación de entrada y salida de un filtro digital.....	24
Ecuación 2. Transformada Z	24
Ecuación 3. Función de transferencia de un filtro digital.....	25
Ecuación 4. Transformada de Laplace de una función de transferencia análoga	25
Ecuación 5. Cambio de variable para pasar de plano s a plano Z	25
Ecuación 6. Magnitud del número complejo Z	25
Ecuación 7. Ángulo del número complejo Z.....	26
Ecuación 8. Transformada bilineal.....	26
Ecuación 9. Convolución discreta	26
Ecuación 10. Media muestral	27
Ecuación 11. Varianza muestral	28
Ecuación 12. Número de series para presentación por pares comparados	29
Ecuación 13. Coeficiente AB para filtro shelving.....	46
Ecuación 14. Coeficiente V0 para filtro shelving	46
Ecuación 15. Coeficiente H0 para filtro shelving.....	46
Ecuación 16. Ecuación de entrada y salida de función $y_1(n)$	46
Ecuación 17. Ecuación de entrada y salida de filtro shelving implementado.....	46
Ecuación 18. Función de transferencia implementada para filtro shelving implementado en el método de Manor	46
Ecuación 19. Función de transferencia de un filtro análogo low-pass de primer orden.....	47
Ecuación 20. Función de transferencia del filtro low-pass implementado en el método de Manor	47
Ecuación 21. Función de transferencia del filtro Head Model implementado en el método de Gilcrest	48
Ecuación 22. Función de transferencia de un filtro análogo low-pass de segundo orden	48
Ecuación 23. Función de transferencia de un filtro análogo high-pass de segundo orden	49
Ecuación 24. Función de transferencia del filtro low-pass implementado en el método de Gilcrest	49
Ecuación 25. Función de transferencia del filtro high-pass implementado en el método de Gilcrest	49
Ecuación 26. Cálculo de la matriz de proporciones	58
Ecuación 27. Ecuación de entrada y salida para el canal L implementada en la aplicación para el método de Manor.....	65
Ecuación 28. Ecuación de entrada y salida para el canal R implementada en la aplicación para el método de Manor.....	65
Ecuación 29. Ecuación de entrada y salida para el canal L implementada en la aplicación para el método de Gilcrest.....	66
Ecuación 30. Ecuación de entrada y salida para el canal R implementada en la aplicación para el método de Gilcrest.....	66

Introducción

Con el auge de las plataformas de streaming musical en los últimos años, junto con el uso masivo de teléfonos inteligentes y audífonos, un alto porcentaje de la población actual usa estos dispositivos para escuchar música mediante streaming. Por otro lado, muchos de los procesos de mezcla de las pistas de audio comercial son realizados mediante el uso de dos altavoces de alta fidelidad, cada uno reproduciendo un canal independiente. Esto se conoce como escucha estereofónica, en la cual cada oído escucha ambos altavoces. El fenómeno que se presenta cuando el oído derecho recibe información del oído izquierdo, y viceversa, es conocido como crosstalk. Al escuchar la misma mezcla mediante audífonos, cada oído recibe información de solamente un canal, generando ausencia de crosstalk. Por ende, la imagen estereofónica y la experiencia general de escucha difieren al realizarse mediante audífonos o altavoces. Por esta razón se implementan sistemas de procesamiento de señal para simulación de crosstalk. De esta manera la escucha a través de audífonos es más parecida a la de un entorno estereofónico con altavoces, acorde a la intencionalidad original que se tenía durante el proceso de mezcla.

En este proyecto se desarrolló una aplicación móvil implementando métodos de procesamiento digital en pistas de audio comercial, con el fin de simular el fenómeno del crosstalk en la escucha a través de audífonos. Además, se realizaron evaluaciones de calidad sonora de pistas reproducidas mediante la aplicación desarrollada, determinando el impacto de la simulación de crosstalk en la percepción de algunos parámetros subjetivos del audio.

Se realizó un proceso de evaluación de preferencia entre cuatro métodos de simulación de crosstalk y el audio original, mediante el método de pares comparados, haciendo uso de tres temas musicales, cada uno de diferente género. Los dos métodos de

procesamiento con mayor preferencia fueron implementados en una aplicación para dispositivos Android. Por último, se evaluaron parámetros subjetivos como la impresión estereofónica, la impresión espacial, el timbre y la calidad general de las tres pistas musicales reproducidas a través de la aplicación. En las pruebas de preferencia, uno de los métodos de simulación de crosstalk logró mejorar la calidad con respecto al audio original. Sin embargo, en los resultados de la evaluación de calidad sonora con la aplicación móvil, se observa que el audio original sigue siendo el de mayores calificaciones, por encima de los audios procesados.

1 Planteamiento del problema

1.1 Descripción de la situación actual

1.1.1 Consumo digital de música

El streaming, también conocido como *live broadcasting* o *webcasting*, es una manera de llevar datos de una fuente a un receptor, en este caso haciéndolo en tiempo real y de manera continua (Austerberry, 2005). Se entiende por tiempo real que no hay ningún procesamiento intermedio entre la fuente y el receptor diferente al *buffering* requerido para la comunicación digital. Gracias a las tecnologías digitales modernas, la comunicación entre los dispositivos receptores (PC, celulares, etc.) y los servidores puede realizarse de manera bidireccional, mejorando la interactividad. Las plataformas de streaming de música permiten a las personas elegir las canciones que quieren escuchar sin necesidad de almacenar los archivos de audio en la memoria del dispositivo.

En la actualidad, el consumo de música en formato digital ha experimentado un crecimiento con respecto a años anteriores. La industria de la música ha tenido una evolución significativa en el área del mercado digital (International Federation of the Phonographic Industry, 2018), indicando que las plataformas de streaming son un proyecto con gran potencial económico, con un crecimiento global del 8.1% en 2017 y un ingreso monetario de US\$17.300 millones. Esto es debido principalmente al desarrollo, accesibilidad y optimización de los dispositivos móviles (smartphones), y a su vez la adaptación a diferentes plataformas de reproducción en streaming como Spotify, Deezer, Soundcloud, Shazam, YouTube, etc. En la Figura 1 se describe el crecimiento económico de las plataformas de

streaming en los últimos años. Se puede observar que desde el año 2015 los ingresos aumentaron drásticamente en relación con los periodos anteriores.

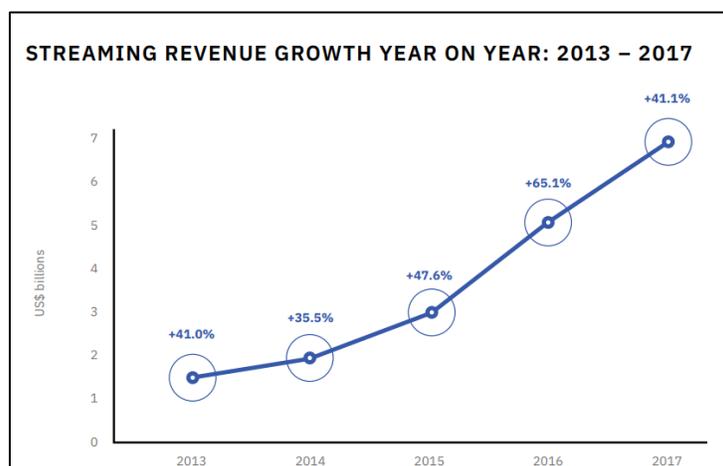


Figura 1. Crecimiento de los ingresos del streaming entre 2013 y 2017. Recuperado de: (International Federation of the Phonographic Industry, 2018).

En términos globales, las cifras significativas para el año 2017 con respecto al año 2016 fueron:

- Crecimiento de los ingresos globales: + 8.1%.
- Participación digital de los ingresos globales: +54%.
- Crecimiento de los ingresos digitales: + 19.1%.
- Crecimiento en ingresos de transmisión: + 41.1%.
- Ingresos físicos: -5.4%.
- Descargar ingresos: -20.5%.

La Figura 2 describe los ingresos por billones de dólares en los últimos años desde 1999. Los porcentajes de ventas en físico describen un decaimiento desde 1999. Sin embargo,

a partir de 2004, surgió la escucha de música en formato digital, esto haciendo referencia a la música pregrabada sin tener en cuenta el uso de streaming. Los ingresos por parte del formato digital se incrementan a través del tiempo y de igual manera el consumo en formato físico disminuye. A partir de 2015 se percibe un incremento significativo en ingresos por parte del servicio de streaming.

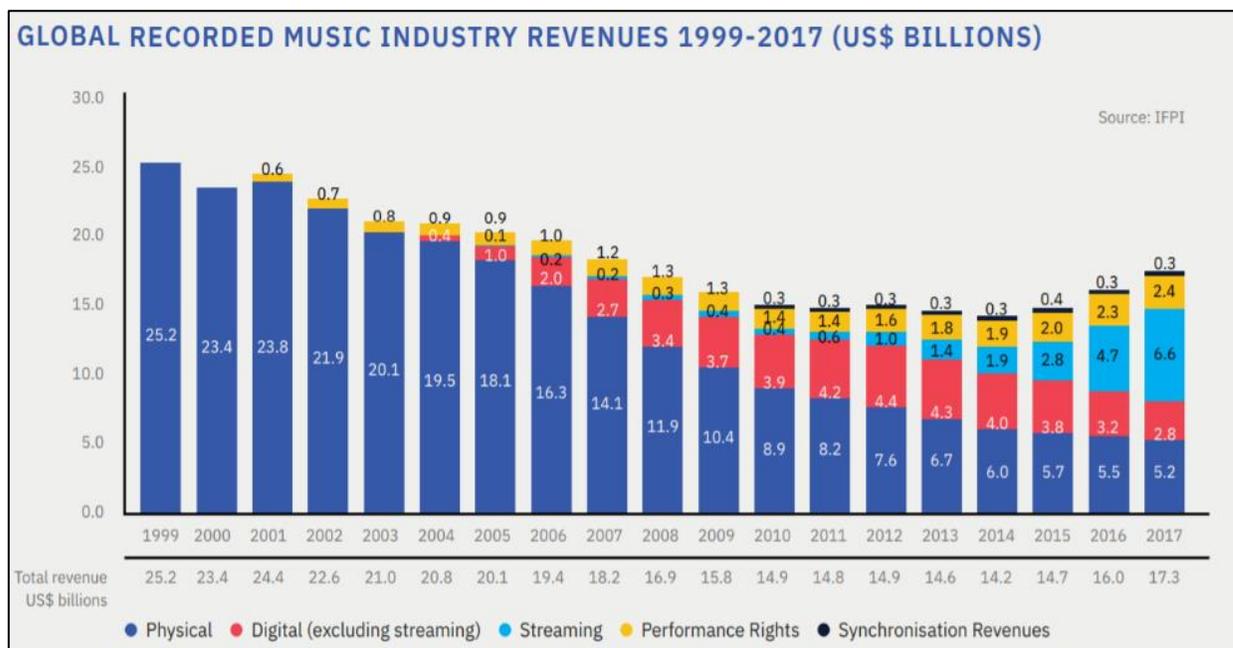


Figura 2. Crecimiento de los ingresos por la industria entre 1999 y 2017. Recuperado de (International Federation of the Phonographic Industry, 2018).

Para el año 2017, los ingresos generados por la industria musical fueron en realidad muy segmentados. No obstante, se puede observar en la Figura 3, que el mayor porcentaje de ingresos está dado gracias a servicios de streaming. En relación con años anteriores, este valor es mucho más significativo y demuestra la evolución y desarrollo que ha tenido este segmento de la industria musical en un muy corto tiempo.

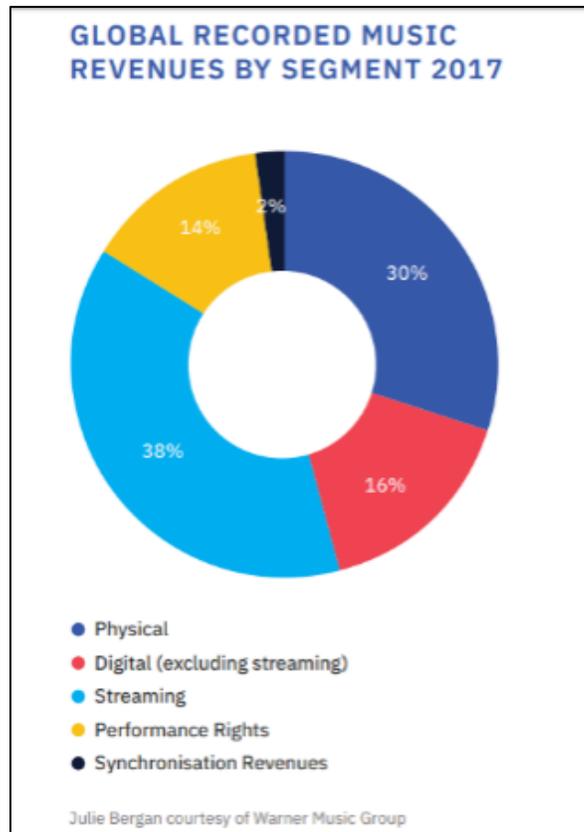


Figura 3. Ingresos segmentados en el año 2017. Recuperado de (International Federation of the Phonographic Industry, 2018).

En Latinoamérica se ha registrado un incremento en el consumo de plataformas de streaming. Los ingresos digitales aumentaron un 17,7 %; más de tres veces el promedio mundial a diferencia de otros continentes como Asia y Australia (5.4 %). Las estadísticas advierten que países latinos como Perú (21.7 %), Chile (14.3 %) y Colombia (10,5 %) tuvieron una mayor contribución en el consumo de música.

Por otra parte, el Departamento Administrativo Nacional de Estadística (2016) realizó un estudio con 29.938 colombianos, de los cuales el 52 % afirmaron haber consumido música grabada en aquella semana. Entre ellas, los valores más altos de consumo se encontraron en los rangos de edad de 12 a 25 años (69.4 %) y de 26 a 40 años (57.4 %). Los datos segmentados muestran un mayor porcentaje de consumo de música en las regiones

Andina (62.3 %) Pacífica (57.6 %), mientras que el valor más bajo se obtuvo en la Región Atlántica (48.6%).

1.1.2 Uso de dispositivos móviles

En cuanto al uso de dispositivos móviles, el estudio realizado en Colombia por Deloitte (2018) con 943 personas entre 16 y 44 años determinó que el 96 % tiene acceso a un smartphone en su hogar o lugar de trabajo. Clasificando el uso del smartphone según el tipo de actividad que realizan, se obtuvo un aumento en cada una de las actividades, con respecto a las cifras de 2016. Al evaluar el porqué del uso del celular, se obtuvo que el 33% de los encuestados lo usan para escuchar música por streaming, mostrando un incremento frente al 28% del 2016.

1.1.3 Uso de audífonos

El uso de audífonos ha tenido un aumento notable. Según el portal web de estadísticas Statista (2017a), al evaluar la frecuencia del uso de audífonos en personas de los Estados Unidos, se muestra que el 31 % de los encuestados los usan a diario, y el 21 % lo hacen varios días a la semana. En cuanto al propósito de su uso, el 87% de los encuestados afirman que usan audífonos principalmente para escuchar música (Statista, 2017b).

1.1.4 Problemática Técnica

A diferencia de la localización espacial proveniente por una sola fuente, que se basa en diferencias interaurales de tiempo e intensidad, la espacialidad del sonido al escuchar por dos altavoces se basa en el precedence effect, en donde el cerebro determina la localización según el retraso que haya entre los frentes de onda cuando dos fuentes emiten versiones muy

similares del mismo sonido, tendiendo a ubicar la fuente virtual hacia la dirección del frente de onda que se percibe primero (Rumsey & McCormick, 2009). Además, diferencias de 15 a 18 dB o de 0.5 ms a 1.5 ms entre ambos canales generan psicoacústicamente una fuente que se ubica completamente hacia un costado.

Escuchar música a través de audífonos tiene ciertas desventajas. Según Bauer (1965), el uso prolongado de audífonos genera fatiga en el oyente, debido a la lateralización de la señal, en la cual hay fuentes que suenan únicamente en un solo oído, algo que no sucede en un entorno de escucha con altavoces. Se generan diferencias de nivel aún mayores que en situaciones de escucha de fuentes externas.

Los productores musicales generalmente basan su juicio de mezcla usando sistemas de monitores de alta fidelidad (Nacach, 2014). Al escuchar el audio a través de monitores en estudios tratados acústicamente, cada oído recibe sonido de ambos canales, mientras que, al usar audífonos, cada oído recibe únicamente sonido de un solo canal, por lo que la imagen estereofónica y la intención inicial de la mezcla se pierden, causando una interiorización no natural del sonido en la cabeza. La solución propuesta por Bauer (1965) se basó en generar artificialmente crosstalk o sangrado entre ambos canales de la señal estereofónica propio de la escucha natural. Con las técnicas de grabación con cabeza binaural, se obtiene una señal similar a la que escucharía una persona ubicada en un punto específico. Sin embargo, al tener crosstalk desde la captura y escuchar dicha señal a través de altavoces se generaría un crosstalk adicional propio de la escucha estereofónica, dañando la espacialidad que se buscaba lograr con la técnica. Por otro lado, la escucha a través de audífonos genera localización principalmente dentro de la cabeza del oyente. Al no tener la posibilidad de

ayudas visuales o movimientos de cabeza, se hace más difícil la ubicación de las fuentes en dirección posterior o frontal y en términos de elevación y distanciamiento.

1.1.5 Pregunta Problema

¿Cómo simular la espacialidad de un entorno de escucha con altavoces a través de audífonos en pistas de audio reproducidas en un dispositivo móvil?

1.2 Estado del Arte

1.2.1 Antecedentes locales

1.2.1.1 Desarrollo de un stand-alone software para simular el comportamiento de una cabeza binaural (Arévalo & Rodríguez, 2014)

El propósito de este trabajo fue obtener la respuesta al impulso de una cabeza binaural, con el fin de poder simular la espacialidad sonora que perciben las personas y de manera digital, sin tener que recurrir a este tipo de instrumentos, ya que estos suelen ser bastante costosos. El proyecto fue desarrollado para trabajar como una plataforma Stand-alone o VST, lo que minimiza las posibilidades de aplicación en un DAW arbitrario, únicamente con DAW que implemente plugins como Re-Wire.

1.2.1.2 Evaluación de la Percepción binaural de un Espacio Tridimensional Definido Implementando un Plugin VST (Hernández & Mosquera, 2015).

Este proyecto de grado buscó evaluar la percepción binaural de un entorno sonoro simulado a través de un plugin VST. Para esto se midió la respuesta al impulso en 36 direcciones de una cabeza binaural y se realizó un algoritmo que permitiera ubicar una fuente en el plano horizontal, generando espacialidad.

1.2.2 Antecedentes nacionales

1.2.2.1 Algoritmo de reproducción de fuentes de sonido virtuales en un espacio virtual tridimensional a través de un sistema binaural (Flórez & Gómez, 2016).

Por medio de interpolación radial, se buscó realizar un algoritmo que lograra simular la existencia de fuentes virtuales dentro de un ambiente o contexto esférico, a través de una estimación de respuesta al impulso, y de esta manera poder desplazar la ubicación de las fuentes en una trayectoria esférica. En cuanto a estimación de respuesta al impulso mencionada anteriormente, se calculó la respuesta al impulso en cualquier punto a partir de la interpolación esférica haciendo uso de una base de datos de respuesta al impulso. Se logró crear un ambiente virtual donde la fuente sonora se desplaza sobre una superficie esférica.

1.2.3 Antecedentes internacionales

1.2.3.1 Round Robin Subjective Evaluation of Stereo Enhancement Systems for Headphones (Lorho, Isherwood, Zacharov, & Huopaniemi, 2002).

En este estudio se realizó una evaluación de la calidad sonora de 9 sistemas de procesamiento de señal estereofónica para reproducción en audífonos, entre los cuales se encontraban 4 plug-ins, 1 stand-alone, 2 prototipos y una tarjeta de audio. Se reprodujeron 6 temas musicales de diferentes géneros a través de cada algoritmo, al igual que sin procesamiento. Se concluyó que ninguno de los sistemas de procesamiento logró mejorar la calidad del audio con respecto a la señal original.

1.2.3.2 Utilization of Filtering Effects in Stereo Headphone Devices to Enhance Spatialization of Source Around a Listener (USA Patente n° US 7,536,021 B2, 2009)

Esta patente muestra un método eficiente de simulación de entorno acústico, en el cual la señal de audio parece venir desde afuera. Para esto se usaron filtros y funciones de transferencia, así como una matriz para enviar las señales al canal respectivo.

1.2.3.3 Preferred Spatial Post-Processing of Popular Stereophonic Music for Headphone Reproduction (Manor, Martens, & Cabrera, 2012).

En este estudio se realizaron pruebas de preferencia entre varios métodos de simulación de crosstalk, en las cuales se concluyó que el método más aceptado fue el basado en simulación de Nearfield Crosstalk a baja frecuencia, simulando diferencias interaurales de intensidad y tiempo.

1.2.3.4 The Duplex Panner: Comparative Testing and Applications of an Enhanced Stereo Panning Technique for Headphone-Reproduced Commercial Audio (Nacach, 2014).

Este proyecto buscó acoplar técnicas de Ambiophonic y Audio Binaural para conseguir una simulación de crosstalk entre los canales de una señal estéreo. De esta manera se evita la necesidad de un sweet spot para estudios de mezcla con cancelación de crosstalk y se resuelven las limitaciones de individualización debido a las diferencias entre los canales auditivos de las personas. Las principales características de este método son decorrelación, paneo temporal y HRTFs.

1.2.3.5 Nearfield Crosstalk Increases Listener Preferences for Headphone-Reproduced Stereophonic Imagery (Manor E. , Martens, Marui, & Cabrera, 2015).

En este estudio se realizó un estudio de preferencia, en el cual se presentaron extractos de piezas musicales populares, uno sin edición digital y otros procesados mediante varios programas de simulación de crosstalk. Los resultados de dicho estudio desembocaron en el desarrollo de una técnica de simulación llamada Nearfield Crosstalk Simulation. La implementación de esta técnica generó mayor aceptación que el audio sin procesamiento. En cuanto a limitaciones de este proyecto, se indica que podría desarrollarse un modelo de predicción para ayudar en el desarrollo posterior de soluciones de simulación de diafonía que deberían producir resultados que probablemente sean preferibles a las versiones originales no procesadas de una gama de programas musicales. Sin embargo, los resultados mostrados en este modelo demuestran la relación compleja entre los atributos auditivos y las preferencias psicoacústicas. Además, conducen a la conclusión de que se requerirá una investigación adicional para que el modelo de predicción encuentre un uso práctico.

1.2.3.6 Headphone Stereo Enhancement Using Equalized Binaural Responses to Preserve Headphone Sound Quality (Bolaños, Pilkki, & Mäkivirta, 2016).

Este artículo propone que el motivo por el cual los oyentes tienden a preferir escuchar el audio original en lugar del procesado para simular una imagen estereofónica similar a la generada por altavoces se debe a la coloración que generan los filtros usados en las HRTFs. Los autores proponen el diseño de filtros que mantengan la fase original, ecualizando la amplitud, de tal manera que se preserve la calidad sonora de los audífonos. El proyecto tiene una limitación en cuanto a la escucha totalmente objetiva de las personas, por lo que esto

afectó las preferencias que se obtuvieron entre la señal original y la señal ya procesada por medio del diseño del filtro.

1.3 Objetivos

1.3.1 Objetivo General

Diseñar e implementar una aplicación para dispositivos Android capaz de simular el fenómeno de crosstalk propio de un entorno de escucha estereofónico con altavoces a través de audífonos en pistas de audio comercial.

1.3.2 Objetivos Específicos

Realizar un proceso comparativo, a partir de pruebas de preferencia, entre cuatro métodos de procesamiento de señal para simulación de crosstalk.

Desarrollar una aplicación móvil para dispositivos Android que implemente los dos métodos de procesamiento evaluados con mayor preferencia.

Evaluar la calidad sonora de pistas de audio procesadas por la aplicación desarrollada, mediante pruebas de evaluación subjetiva.

1.4 Justificación

La gran mayoría de los procesos de mezcla de temas musicales son llevados a cabo escuchando la pista a través de monitores de alta fidelidad (Nacach, 2014). Al escuchar la misma mezcla a través de audífonos no se presenta el fenómeno del crosstalk, por lo que la imagen estereofónica original se ve alterada, perdiendo la intencionalidad de la mezcla. Por otro lado, teniendo en cuenta los estudios realizados por Bauer (1965), la escucha a través de audífonos genera una interiorización del sonido no natural, ubicándolo dentro de la cabeza, además, se produce lateralización no natural, debido a la ausencia de diferencias interaurales

de tiempo e intensidad, propias de un entorno de escucha de fuentes externas. Estos factores producen fatiga con el uso prolongado de audífonos, implementando un sistema que añada el crosstalk en la escucha por audífonos se puede simular un entorno de escucha más realista, manteniendo la imagen estereofónica original del audio sin generar fatiga en el oyente.

Las encuestas realizadas con respecto al uso masivo de audífonos (87 % los usa para escuchar música), así como de smartphones (96 % tienen acceso a dichos dispositivos), muestra que una de las maneras más comunes de escuchar música actualmente es haciendo uso de ambos dispositivos. Teniendo en cuenta las problemáticas que se presentan a nivel perceptual al escuchar pistas de audio mediante audífonos, es necesario diseñar una aplicación que permita mejorar la calidad del audio que se reproduce en el celular y que es escuchado a través de los audífonos.

Se han desarrollado varios métodos de procesamiento digital de señal con el fin de simular crosstalk en audífonos, los cuales han logrado mejorar la calidad sonora del audio. Sin embargo, ninguno fue implementado propiamente en una aplicación para dispositivos móviles.

Hay funciones internas del sistema operativo (como el caso de Android) que simulan espacialidad en el audio que se reproduce en el dispositivo móvil, pero se manejan mediante parámetros arbitrarios que no especifican qué tipo de procesamiento se está realizando a la señal, ni tampoco se han evaluaciones de calidad sonora de dichas funciones.

Por lo anterior, se evidencia la necesidad del desarrollo de una herramienta capaz de simular la escucha con altavoces a través de audífonos en dispositivos móviles. Teniendo en cuenta las cifras analizadas previamente, es necesario resolver las problemáticas presentes en la escucha con audífonos, tales como fatiga auditiva, pérdida de la intencionalidad inicial de la mezcla para pistas de audio comercial, deficiencias en la espacialidad de la imagen

estereofónica cuando es reproducida por audífonos, lo que provoca una sensación de escucha de un entorno no natural.

1.5 Alcances y Limitaciones

1.5.1 Alcances

El desarrollo exitoso del proyecto sería una solución a las problemáticas de fatiga por lateralización no natural y de alteración de la imagen estereofónica en las mezclas musicales, generando un entorno de escucha más acorde a la intención original de la mezcla mientras es reproducida a través de audífonos.

Se crearía una aplicación móvil que implemente los métodos de simulación de crosstalk y que permita la reproducción de archivos de audio estereofónico almacenadas en la memoria interna del dispositivo móvil.

1.5.2 Limitaciones

El uso de esta aplicación es únicamente para dispositivos con sistema operativo Android. Además, el desempeño adecuado depende de la capacidad de procesamiento de este.

La aplicación se limita a reproducir archivos de audio guardados en la memoria del dispositivo. Como una mejora a futuro, se puede configurar para procesar audio en tiempo real procedente de plataformas de streaming.

El grado en el que el oyente logra percibir el efecto del procesamiento depende del tipo de transductor, así como de la calidad sonora del mismo. La mejor manera de escuchar el efecto del procesamiento del audio es a través de audífonos over-ear.

2 Marcos de Referencia

2.1 Marco Conceptual

2.1.1 Percepción Binaural

Dada la composición binaural del sistema auditivo humano, es posible determinar la dirección de la cual proviene un evento sonoro y el espacio que ocupa, esto en función de su posición y del contenido espectral del sonido. Esta capacidad está ligada a 3 fenómenos importantes (Bartlett, 1991): la diferencia en los tiempos de llegada de una onda sonora a ambos oídos, diferencias de nivel entre los eventos sonidos percibidos y el filtrado de peine generado por el oído externo. Tanto las diferencias de tiempo como de intensidad permiten localización en el eje horizontal, mientras que el oído externo permite discernir entre fuentes ubicadas en frente o detrás del oyente (Gelfand, 1981). Al tener un evento sonoro ubicado más hacia un costado, estas diferencias permiten saber de qué dirección proviene (Willert, 2006). La Figura 4 muestra el efecto de la separación de los oídos y la sombra acústica de la cabeza a la hora de hallar diferencias entre los sonidos que se perciben por cada canal auditivo.

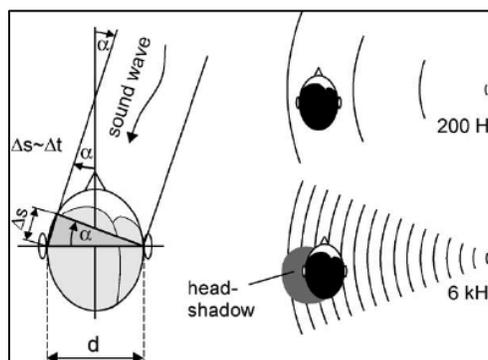


Figura 4. Efecto de la separación de los oídos y la sombra de la cabeza en el proceso de percepción binaural. Recuperado de: (Willert, 2006).

2.1.1.1 Diferencias interaurales de tiempo

Las diferencias interaurales de tiempo, son más evidentes en el espectro de baja frecuencia, debido a que la longitud de onda de dichas frecuencias no tiene interacción directa con el tamaño de la cabeza (Blauert, 1997). Las diferencias son percibidas a partir de la llegada del primer frente de onda, lo que se conoce como “efecto de precedencia” (Zölzer, 2002).

2.1.1.2 Diferencias interaurales de intensidad

Las diferencias interaurales se comportan de manera diferente en función de la frecuencia (Fastl & Zwicker, 2007). Por lo tanto, las diferencias interaurales de intensidad son mayores para los componentes de alta frecuencia, como se ilustra en la Figura 5.

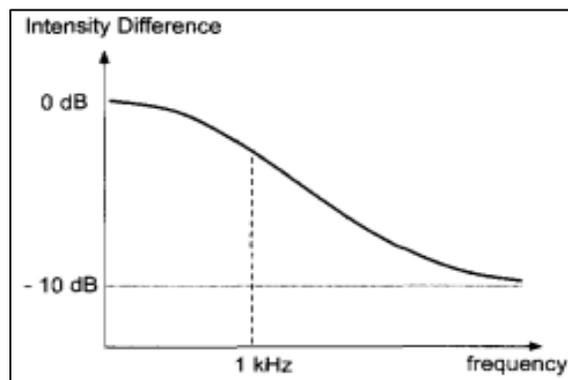


Figura 5. Diferencia de nivel entre cada oído en función de la frecuencia. Recuperado de (Zölzer, 2002).

Debido al tamaño de la cabeza y la sombra acústica que se genera sobre el oído contrario a la dirección de la cual proviene el sonido, las diferencias interaurales de intensidad son más notorias a partir de 1kHz.

2.1.2 Imagen Estereofónica

Los inicios del audio estereofónico datan de 1881 cuando Clement Ader implementó un sistema de dos canales para telecomunicaciones (Hertz, 1981). Se hacían transmisiones de la Ópera de París a través de un par de teléfonos, reproduciendo estas señales en pequeños cuartos de invitados.

Debido a la capacidad de interpretación visual que tienen los seres humanos y a las características espaciales del sonido es posible generar una representación mental de los eventos sonoros (Blauert, 1997). Las condiciones del entorno de escucha también influyen en la percepción espacial. Por ejemplo, es más difícil determinar la posición de una fuente en un entorno reverberante que en uno anecoico. Debido al componente espacial o visual del fenómeno de percepción auditiva, el ser humano ha aprendido a inferir atributos espaciales de las fuentes en función de lo que escucha.

Al tener varios eventos sonoros en simultáneo, cada uno de ellos es procesado por el sistema auditivo, generando atributos espaciales para cada sonido. Esta espacialidad es simulada de diferentes maneras en los procesos de captura, mezcla y reproducción del audio. La imagen estereofónica está compuesta por los atributos espaciales de diferentes eventos sonoros reproducidos desde un audio estereofónico, (Shirley, Kendrick, & Claire, 2007). Esta imagen es un componente netamente subjetivo de la experiencia de escucha, y depende de las técnicas de captura y mezcla usadas para la creación del audio, así como de las características del sistema de reproducción.

Desde el proceso de captura, se implementan técnicas de microfoneo en estéreo con el fin de simular de la manera más fiel posible la ubicación de los eventos sonoros presentes

en un espacio (Bartlett, 1991), como lo muestra la Figura 6. Técnicas como el par A/B, M/S, ORTF o grabación con una cabeza binaural son algunas de las maneras más populares de generar espacialidad desde el proceso de captura de una señal sonora.

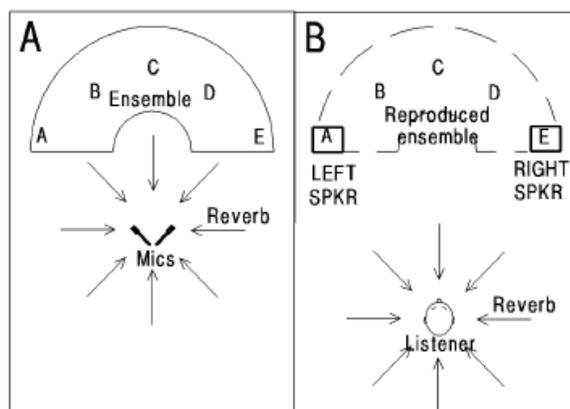


Figura 6. Simulación de espacialidad sonora a partir de técnicas de microfoneo estéreo. Recuperado de: (Bartlett, 1991).

La imagen estereofónica está asociada a la imagen fantasma (Theile, 1991), la cual se genera con la ubicación de dos altavoces, cada uno reproduciendo un canal de manera independiente. La ubicación de las fuentes en la imagen estereofónica se esparce a lo largo de la distancia que hay entre ambas fuentes, dependiendo de la intensidad de cada evento sonoro. Dicha imagen es diferente al reproducir el mismo audio a través de audífonos, debido a las diferencias interaurales presentes en este caso.

2.1.3 Crosstalk

El crosstalk es el fenómeno presente en la escucha de fuentes externas, generado por diferencias interaurales de tiempo e intensidad entre el sonido percibido por ambos oídos

(Lorho, Isherwood, Zacharov, & Huopaniemi, 2002). El crosstalk se da porque una fuente externa se escucha a través de cada oído.

2.1.3.1 Escucha estereofónica

La escucha estereofónica se produce mediante la implementación de un par de altavoces, cada uno reproduciendo un canal independiente dentro de una mezcla (Manor, Martens, & Cabrera, 2012). La externalización de las fuentes sonoras genera crosstalk en el sonido percibido, debido a que cada oído escucha ambos altavoces.

2.1.3.2 Escucha bifónica

La escucha bifónica se obtiene a través del uso de audífonos para reproducir dos canales de una mezcla estereofónica. La diferencia con la escucha estereofónica es la ausencia de crosstalk, ya que cada oído recibe información de únicamente una fuente (Manor, Martens, & Cabrera, 2012). Los métodos de simulación de crosstalk en audífonos hacen posible la obtención de una escucha estereofónica en audífonos, generando a través de procesamiento de señal las diferencias interaurales presentes en un entorno de escucha con un par de altavoces. La Figura 7 ejemplifica la diferencia entre la escucha estereofónica y bifónica, así como el efecto de la simulación del crosstalk en audífonos.

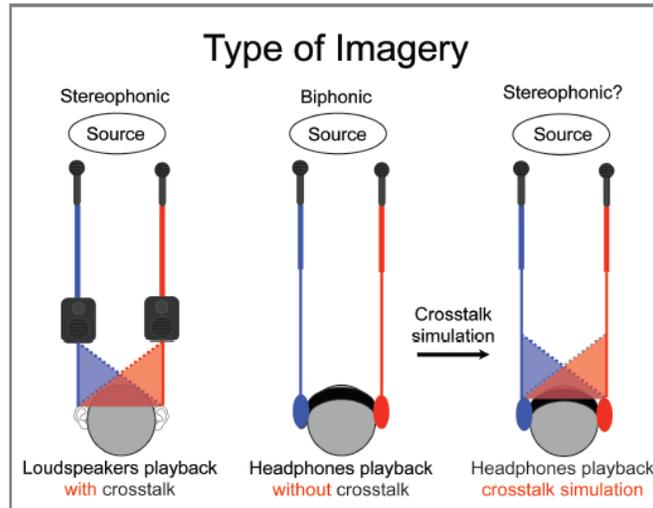


Figura 7. Diferencia entre escucha estereofónica (izquierda), bifónica (centro) y bifónica con simulación de crosstalk (derecha). Recuperado de: (Manor, Martens, & Cabrera, 2012).

2.1.4 Calidad Sonora

La calidad sonora se define como “adecuación de un sonido en el contexto de un objetivo técnico específico y/o tarea” (Kang, 2006). Sin embargo, Blauert (2012), propone que la calidad sonora es un fenómeno muy complejo, ya que la percepción se puede analizar desde un punto de vista epistemológico basado en el cuerpo, alma y mente frente a una situación o contexto de evaluación. Por lo tanto, la percepción se divide en sentimientos (alegría, miedo, etc), cosas (percepciones sensoriales) y conceptos (combinación de sentimientos y cosas). Además, Blauert (2012) advierte que las evaluaciones de calidad sonora deben realizarse por capas, dependiendo del tipo de información que se requiere abstraer. Según el tipo de parámetros que se estén evaluando, se puede definir cómo es la percepción del oyente ante un evento sonoro específico, así como la adecuación de dicho evento sonoro a un contexto específico. Para este proyecto, se hacen evaluaciones subjetivas, por lo que la calidad del audio es definida en función de las reacciones de los evaluadores.

2.1.5 Procesamiento Digital de Señales

Un sistema de procesamiento digital de señal se define como el conjunto de algoritmos usados para calcular una secuencia de números en la salida a partir de una secuencia de números entrante (Zölzer, 2002). Dichos algoritmos tienen parámetros que definen su funcionamiento. Las secuencias de entrada y salida son señales digitales que representan señales sonoras análogas.

2.1.5.1 Señales digitales

Una señal digital es una representación de una señal análoga mediante una serie de números que equivalen al valor de una magnitud física medido a intervalos equidistantes de tiempo (Zölzer, 2002). El tiempo T que hay entre dos muestras se define como el periodo de muestreo, cuyo inverso equivale a la frecuencia de muestreo f_s del conversor análogo digital expresada en ciclos por segundo o Hertz. La frecuencia máxima a la cual la señal análoga puede ser representada en una señal digital equivale a la mitad de la frecuencia de muestreo. Por otro lado, la magnitud de la señal digital en cada muestra está definida por la profundidad en bits BD , la cual permite cuantizar la magnitud física medida entre $2^{BD} - 1$ valores digitales. Dicho rango también puede expresarse en valores entre -1 y 1. La Figura 8 ilustra la representación digital de una señal análoga.

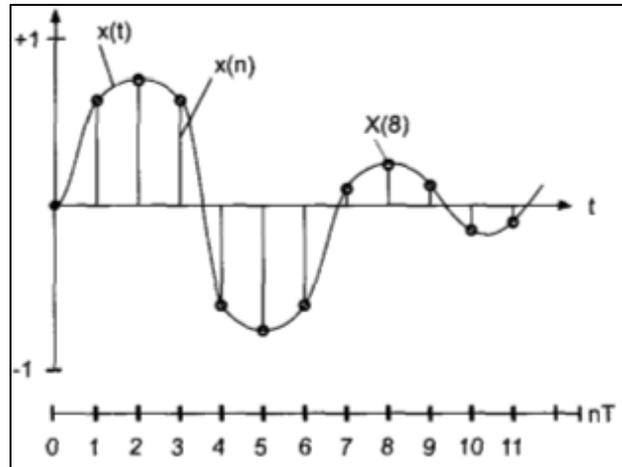


Figura 8. Representación una señal analógica $x(t)$ a través de una señal digital $x(n)$ a intervalos de tiempo T . Recuperado de: (Zölzer, 2002).

2.1.5.2 Filtros digitales

Un filtro digital es un algoritmo que modifica el contenido en frecuencia de una señal digital (Tyril & Pedersen, 2001). Está definido por una serie de coeficientes, los cuales pueden ser de realimentación o retroalimentación. Por ende, hay dos tipos de filtros digitales: de respuesta al impulso finita (FIR) y de respuesta al impulso infinita (IIR). Los filtros FIR hacen uso solamente de coeficientes de realimentación, es decir, la señal digital de salida está definida únicamente por los valores de la señal de entrada. Por otro lado, los filtros IIR generan una señal de salida en función de los valores de las señales tanto de entrada como de salida en muestras anteriores. La Figura 9 muestra el diagrama de bloques de un filtro digital con L cantidad de coeficientes de realimentación b_l , M cantidad de coeficientes de retroalimentación a_m y con bloques z^{-1} que simbolizan un retraso de 1 muestra. Para los filtros FIR, los coeficientes retroalimentación son iguales a 0.

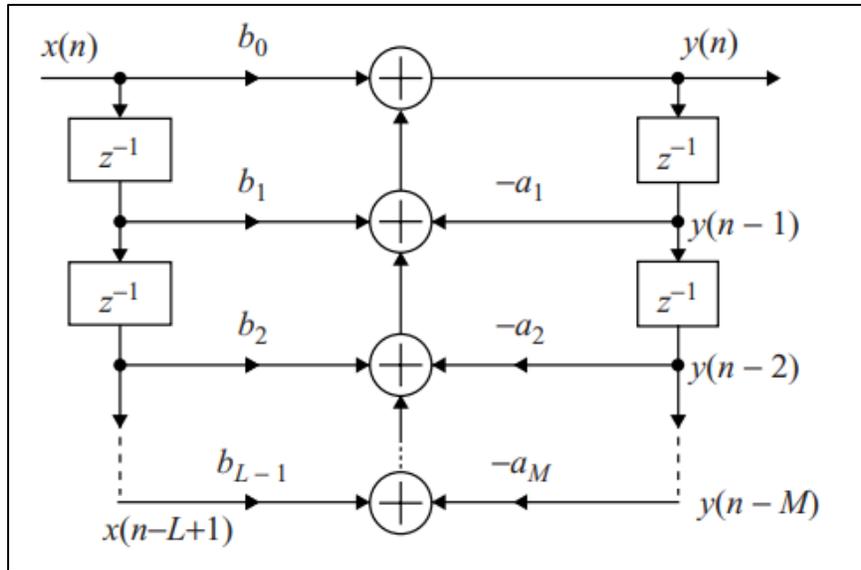


Figura 9. Diagrama de bloques de un filtro digital. Recuperado de (Kuo, Lee, & Tian, 2006).

Teniendo esto en cuenta, se puede representar la señal de salida $y(n)$ de un filtro digital en función de su señal de entrada $x(n)$ mediante la Ecuación 1.

$$y(n) = \sum_{l=0}^{L-1} b_l x(n-l) - \sum_{m=1}^M a_m y(n-m) \quad \text{Ecuación 1}$$

2.1.5.3 Función de transferencia de filtros digitales

La transformada Z de una señal discreta $x(n)$ se haya mediante la Ecuación 2 (Graf, 2000).

$$X(z) = \sum_{n=0}^{\infty} x(n)z^{-n} \quad \text{Ecuación 2}$$

Al realizar la transformada Z de la ecuación de entrada y salida de un filtro digital, se puede expresar su función de transferencia mediante la Ecuación 3.

$$H(z) = \frac{Y(z)}{X(z)} = \frac{\sum_{l=0}^{L-1} b_l z^{-l}}{1 + \sum_{m=1}^M a_m z^{-m}}$$

Ecuación 3

2.1.5.4 Transformada bilineal

Es un procedimiento mediante el cual se puede representar una función de transferencia análoga a través de una serie de coeficientes de realimentación y retroalimentación, es decir, permite transformar un filtro del dominio análogo al dominio digital (Barsainya, Rawat, & Mahendra, 2016). Esto se realiza debido a la diferencia en los rangos de frecuencia de cada dominio. Mientras que las frecuencias análogas no tienen ningún valor máximo, la información de las frecuencias digitales se limita hasta la mitad de la frecuencia de muestreo usada.

La función de transferencia de un filtro análogo es igual a la transformada de Laplace de su respuesta al impulso, como se muestra en la Ecuación 4 (Schiff, 2013).

Ecuación 4

$$H(s) = \frac{Y(s)}{X(s)} = \int_0^{\infty} h(t) e^{-st} \quad ; \quad s = \sigma + j\Omega$$

Según Kuo (2006) la transformada Z de una señal digital $x(nT)$ puede verse como su transformada de Laplace, mediante el cambio de variable de la Ecuación 5. Se puede representar la amplitud del número complejo mediante la Ecuación 6 y el ángulo mediante la Ecuación 7.

Ecuación 5

$$z = e^{sT} = e^{\sigma T} e^{j\Omega T} = |z| e^{j\omega}$$

Ecuación 6

$$|z| = e^{\sigma T}$$

$$\omega = \Omega T$$

Para transformar un filtro análogo al dominio digital se realiza el cambio de variable que se muestra en la Ecuación 8.

$$s = \frac{2}{T} \left(\frac{z - 1}{z + 1} \right)$$

2.1.5.5 Convolución discreta

Para un sistema digital lineal e invariante en el tiempo, se puede representar la señal de salida $y(n)$ del sistema en función de la señal de entrada $x(n)$ según la Ecuación 9 (Chu, 2008).

$$y(n) = x(n) * h(n) = h(n) * x(n) = \sum_{l=-\infty}^{\infty} x(l) h(n - l) = \sum_{l=-\infty}^{\infty} h(l) x(n - l)$$

siendo $h(n)$ la respuesta al impulso de dicho sistema y ‘*’ el operador que simboliza la convolución de ambas señales. Como se observa en la ecuación, la convolución tiene la propiedad conmutativa, es decir, no importa el orden en el que se operen las señales. Sin embargo, se suele tomar la señal de menor tamaño como $h(n)$.

2.1.5.6 Respuesta al impulso

La respuesta al impulso de un sistema lineal e invariante en el tiempo equivale a la señal de salida que se obtiene cuando a dicho sistema se le ingresa un impulso como señal de entrada (Oppenheim, Willsky, & Nawab, 1998). De esta manera es posible caracterizar un

sistema, ya que cualquier señal de salida obtenida de este será igual a la convolución de la señal de entrada con la respuesta al impulso.

2.1.6 Análisis Estadístico de Datos

2.1.6.1 Estadística Descriptiva

La estadística descriptiva es la técnica matemática que obtiene, organiza, presenta y describe un conjunto de datos con el propósito de facilitar el uso, generalmente con el apoyo de tablas, medidas numéricas o gráficas (Rodríguez Ojeda, 2007). Es posible hallar medidas de posición central que definen el valor alrededor del cual se concentran los datos u observaciones según los valores (Dicovski Riobóo, 2008). La media muestral es el resultado de la suma de todas las observaciones dividido por el número total de datos, como se muestra en la Ecuación 10.

$$Media(X) = \bar{x} = \frac{(X_1 + X_2 + \dots + X_n)}{N} \quad \text{Ecuación 10}$$

Donde:

X_n : El conjunto de observaciones.

N: Número total de observaciones.

Las medidas de dispersión son números que proveen información adicional acerca del comportamiento de los datos, describiendo numéricamente su dispersión. (Walpole, Myers, Myers, & Keying, 2012). La varianza muestral es la cuantificación de la distancia de los datos con respecto a la media establecida, como ilustra la Ecuación 11.

$$S^2 = \frac{\sum_{i=1}^n (X_i - \bar{x})^2}{N - 1}$$

Donde:

S^2 : Es Varianza Muestral

X_i : El valor del cual se quiere hallar la distancia a la media

\bar{x} : Media Muestral

N: Número total de observaciones

La desviación estándar es el valor que indica qué tan dispersos están los datos con respecto a la media, se define como la raíz cuadrada positiva de la variancia. Se obtiene al calcular la raíz cuadrada de la variancia.

La distribución normal de datos determina la tendencia de observaciones teniendo en cuenta la media muestral. Es decir, una distribución normal de datos es una distribución con forma de campana simétrica donde las desviaciones estándar sucesivas con respecto a la media establecen valores de referencia para estimar el porcentaje de observaciones de los datos. (Walpole, Myers , Myers, & Keying, 2012).

2.1.6.2 Ley de Juicio Comparativo

Esta ley fue propuesta por el psicólogo Louis Leon Thurstone (Morosini, 2012). Aquí se realiza un escalamiento subjetivo por medio de métodos estadísticos y se valora la actitud personal del individuo sometido a la prueba. La ley de juicio comparativo advierte que cada vez que se le presenta al sujeto de prueba una serie de estímulos, ocurre un proceso de “discriminación”, de manera que cada estímulo tiene un proceso subjetivo y también una

valoración subjetiva. Así, por cada estímulo se tendrá un proceso y una valoración diferente. Otro factor importante para tener en cuenta es la “distribución discriminativa”, el cual se hace relación al valor subjetivo de mayor frecuencia en el estímulo. Para que en la distribución haya una variabilidad es necesario tener en cuenta la desviación típica o bien llamada “dispersión discriminativa”. Por ejemplo, si a un sujeto se le presenta una serie de estímulos con la finalidad de que escoja uno u otro, adicionando una valoración subjetiva de acuerdo a un parámetro determinado, cada valor escogido es el resultado del proceso discriminativo. Existen formas adecuadas de presentar la serie de estímulos. Si llega a ser muy grande la serie, es necesario aplicar el método de pares comparados o bien llamado comparaciones binarias.

Dependiendo de la cantidad de estímulos se obtendrán las posibles combinaciones binarias para que el sujeto pueda presentar la prueba, teniendo la libertad de escoger el estímulo de su preferencia para tener como resultado un valor subjetivo que domine en la escogencia.

En la Ecuación 12 se presenta el número de series de pares para realizar las posibles combinaciones:

Ecuación 12

$$\#Series = N * \left(\frac{N - 1}{2}\right)$$

Para realizar el escalamiento correctamente, es necesario realizar unas matrices con los resultados: matriz de frecuencia observada, matriz de proporciones y matriz de puntuaciones típicas (Rey & Gil, 2013). La matriz de frecuencia observada presenta todos los estímulos y contiene la frecuencia es escogencia de cada par comparativo. Para la matriz de proporciones está dada por la división de cada casilla de la matriz de frecuencia observada

entre el valor de sujetos encuestados. La matriz de puntuaciones típicas se obtiene al aplicar una normalización a partir de una tabla de distribución normal (distribución Z) a la matriz de proporciones.

Por último, para la presentación de datos se realiza una representación gráfica de escala donde están los valores escalares promediados.

2.1.6.3 Prueba de Normalidad de Shapiro-Wilk

La prueba de Shapiro-Wilk permite contrastar la distribución de un conjunto de datos con una distribución normal (Surhone, M., Timpledon, M., & Marseken, 2010). De esta manera es posible determinar si el conjunto de datos analizado posee o no una distribución normal. Se basa en la regresión de los datos sobre los valores teóricos de una distribución hipotetizada, en donde su estadístico W representa el cociente de dos estimaciones de la varianza de una distribución normal (Pedrosa, Juarros-Basterretxea, Robles-Fernández, Basteiro, & García-Cueto, 2015). Se usa generalmente cuando se tiene un número de muestras menor a 30.

2.1.6.4 Prueba H de Kruskal-Wallis

Introducida en 1952 por William H. Kruskal y W. Allen Wallis, es un modelo de análisis de pruebas no paramétricas (no se asume que su distribución es normal) para medir la igualdad de las medias en el análisis de varianza de un factor (Walpole, Myers, Myers, & Keying, 2012). Aquí se asume que los datos vienen de la misma distribución, no necesariamente normal. Esta prueba permita hallar un valor de significancia el cual, si es menor a 0.05, significa que hay diferencias estadísticamente significativas entre las variables evaluadas. Sin embargo, al evaluar varios grupos, no define específicamente cuál es el que

predomina estadísticamente sobre los otros. El cálculo de esta prueba aplicado en el presente estudio fue realizado mediante el software IBM SPSS 25.

2.2 Marco Teórico

2.2.1 Métodos de Simulación de Crosstalk

Se ha propuesto la simulación de crosstalk como una solución para generar en audífonos un entorno de escucha propio de altavoces. De esta manera se podrían resolver las problemáticas relacionadas con fatiga al uso prolongado de audífonos y alteración de la imagen estereofónica de la mezcla. Se seleccionaron 4 métodos de simulación de crosstalk para este estudio.

2.2.1.1 Balanced Stereo Widening Network

Este método se denomina balanceado debido a que, a pesar de realizar suma de la señal original con los filtros que simulan el crosstalk, mantiene la amplitud original (Kirkeby, 2002). Según esto, cada filtro fue atenuado con una ganancia de 0.5. Se basa en la influencia de la cabeza a la hora de filtrar señales provenientes del lado contrario al oído. Para frecuencias menores a 1kHz la cabeza es pequeña en comparación a las longitudes de onda, por lo que no hay ninguna alteración. Entre 1.5kHz y 2kHz se produce una atenuación que aumenta linealmente. Después de 2kHz se asume una atenuación constante. El filtro H_d (izquierda) de la Figura 10 corresponde al procesamiento del canal para el oído directo, mientras que el filtro H_x (derecha) procesa la señal de un canal para enviarla al oído contrario. Los filtros implementados para el procesamiento de señal son filtros FIR de fase lineal (Lian & Lim, 1993), por lo que no hay distorsión de fase a la salida de cada filtro. Para simular las diferencias interaurales de tiempo, el canal filtrado por H_x es retrasado 0.7ms y enviado al canal opuesto.

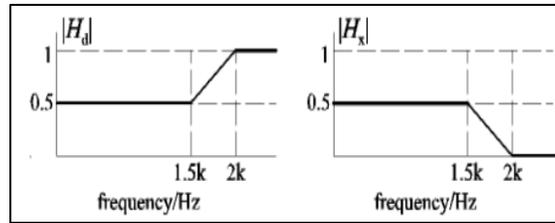


Figura 10. Respuesta en frecuencia de los filtros para canal directo (izquierda) y canal opuesto (derecha) del método Balanced Stereo Widening Network. Recuperado de: (Kirkeby, 2002).

2.2.1.2 Nearfield Crosstalk

En este método se busca incluir IIDs a baja frecuencia, lo que se conoce como Nearfield Crosstalk (Manor E. , Martens, Marui, & Cabrera, 2015). El procesamiento para el oído directo tiene una ganancia general de 3dB y el filtro del oído contrario tiene una atenuación de -3dB. De esta manera hay diferencias interaurales de intensidad en todo el rango de frecuencias. La Figura 11 muestra la respuesta en frecuencia de los filtros del método de campo cercano (Nearfield) en contraste con los filtros usados para campo lejano (Farfield). Para este método se empleó un delay de 0.5ms para las diferencias interaurales de tiempo.

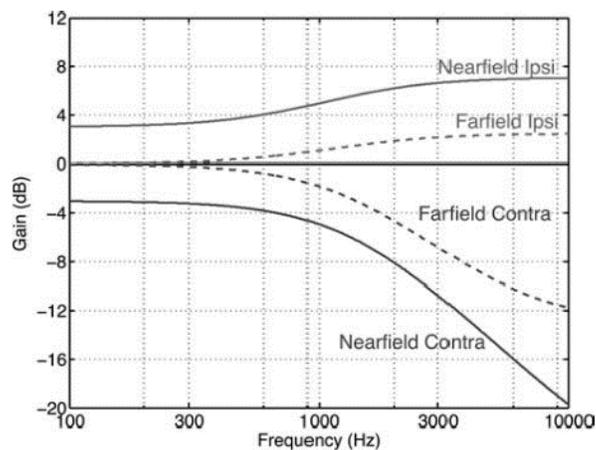


Figura 11. Respuesta en frecuencia de los filtros para campo cercano y campo lejano. Recuperado de: (Manor E. , Martens, Marui, & Cabrera, 2015).

2.2.1.3 earGoggles

Para este método, denominado ‘earGoggles’ (Gilchrest, 2016), se toma como referencia la técnica de microfoneo M/S. La señal Side se obtiene de la resta de ambos canales de la señal estéreo. Esta señal no tiene componentes del audio paneados al centro de la imagen estéreo y contiene mucha información sobre la espacialidad. Al incluir la señal Side en ambos canales se mantiene la señal Mid. El método aquí propuesto realiza para cada canal un filtro pasa-bajas con atenuación de -6dB y delay de 20 ms dirigido al oído contrario. La señal Side es procesada mediante un filtro pasa-altas con atenuación de -6dB y delay de 15ms dirigido a cada canal, pero con polaridad invertida en el canal derecho, simulando efectos de sala. Además, la señal Side es procesada mediante un filtro ‘Head Model’, el cual simula la sombra acústica de la cabeza para el oído contrario a la dirección de una fuente sonora, esto en función de su ángulo de incidencia (Zölzer, 2002).

2.2.1.4 Ecuación de respuesta al impulso binaural

Este método busca simular crosstalk en audífonos sin dañar la calidad sonora de los mismos (Bolaños, Pilkki, & Mäkivirta, 2016). Se menciona la posibilidad de tener un filtro binaural cuya respuesta sea ecualizada por un oyente experto con el fin de evitar el desbalance tonal que puede provocar un filtrado de este tipo en la escucha con audífonos, algo que realizaron Fleischmann, Silzle y Plogsties (2011). Por otro lado, Merimaa (2009) propone la adecuación de filtros obtenidos a partir de HRTF para contrarrestar dicha coloración no deseada. Bolaños (2016), por ende, menciona en su estudio la posibilidad de ecualizar una respuesta binaural de una ‘dummy head’ para mejorar la espacialidad del audio estereofónico reproducido en audífonos. Esto se obtiene diseñando un filtrado de señal que reduzca las coloraciones que incluye intrínsecamente un sistema de simulación de crosstalk. Para el

procesamiento de la señal se requirió una Head Related Transfer Function (HRTF) y Binaural Room Impulse Response (BRIR) simultáneamente, siendo la HRTF la respuesta en frecuencia de la BRIR. Se obtuvo la respuesta al impulso de una cabeza binaural dentro de un recinto con características acústicas controladas. Posteriormente, se le pidió a un oyente experto la ecualización de dichos filtros, para que la señal fuese balanceada a su criterio. Al igual que en el método de Kirkeby, se busca mantener la amplitud general del audio.

2.2.2 Programación en Unity

Unity es un motor y editor de videojuegos multiplataforma (Landa, 2013). Su programación se basa en GameObjects, componentes propios de Unity, los cuales pueden ser elementos de interfaz gráfica, figuras 3D, fuentes de iluminación u objetos vacíos. Cada GameObject contiene componentes de diversos tipos (ubicación en el espacio, material de renderización, etc.), los cuales le dan a cada GameObject atributos específicos según las funciones que se requieran.

El componente 'AudioSource' es el encargado de generar audio proveniente de un GameObject, así como de editar su panning, pitch, ganancia y demás atributos sonoros (Geig, 2013). Cada 'AudioSource' recibe un objeto de tipo 'AudioClip', el cual actúa como un contenedor de datos de audio, es decir, un arreglo con muestras de audio, la frecuencia de muestreo, el número de canales, el número de muestras, etc. Cada 'AudioClip' debe ser asignado a un componente 'AudioSource' para permitir su reproducción. Al obtener el arreglo de muestras de un audio estereofónico (2 canales) proveniente de un AudioClip se obtiene un arreglo unidimensional, el cual contiene un dato del canal L seguido de un dato del canal R, siguiendo esta secuencia para cada muestra del audio original.

El lenguaje de programación para desarrollo en Unity es C#, el cual es un lenguaje de alto nivel que pertenece al paquete .NET, un entorno de desarrollo multilenguaje diseñado por Microsoft para aplicaciones de internet (Ceballos Sierra, 2007).

Para ejecutar un código en Unity es necesario asignarlo a un GameObject creado dentro del proyecto. Cada código tiene dos funciones principales:

- *start()*: Se ejecuta al dar inicio a la escena.
- *update()*: Se ejecuta constantemente durante el transcurso de la escena.

2.3 Marco Normativo

2.3.1 RECOMMENDATION ITU-R BS.1284-1: General methods for the subjective assessment of sound quality.

La Unión Internacional de Telecomunicaciones (2003) publicó este documento, en el cual brinda recomendaciones para realizar evaluaciones subjetivas de calidad sonora, de las cuales algunas se muestran a continuación.

2.3.1.1 Diseño del experimento

Es necesario tener un control minucioso sobre los detalles de la prueba, para evitar que los resultados se vean afectados por factores externos (Campbell, Jones, & Glavin, 2009). Está permitido presentar las condiciones de prueba de manera aleatoria, en especial cuando las tareas puedan resultar difíciles para los sujetos de prueba. Un paso importante para aplicar en el experimento es la inclusión, es decir, presentar algunas muestras con el fin de familiarizar al sujeto de prueba con los atributos que se espera que pueda calificar.

2.3.1.2 Selección de los oyentes

Si bien los oyentes inexpertos representan la mayoría de la población de nuestra sociedad, se recomienda recurrir a oyentes expertos en las evaluaciones de calidad sonora (Blauert & Jekosch, 2012). Esto debido a que tienen el mejor criterio para juzgar las muestras de audio, además de que una larga exposición a ciertos fenómenos auditivos puede hacer que algunos oyentes inexpertos deterioren su juicio de los parámetros evaluados. En caso de recurrir a oyentes expertos, el número mínimo de participantes debe ser 10. De lo contrario, se recomienda usar 20 oyentes inexpertos como mínimo.

2.3.1.3 Método de prueba

- **Escalas de apreciación**

Para la evaluación de calidad se puede usar una escala de 1 a 5 (Mala, Insuficiente, Adecuada, Buena, Excelente).

2.3.1.4 Atributos

Los atributos o características que se evalúan dependen de los objetivos de la prueba. A continuación, se muestran los más relevantes para el proyecto.

- **Calidad de Audio Básica**

Incluye todos los atributos que puedan ser calificados en el audio, más no de manera exhaustiva, sino más bien de manera general.

- **Atributos que especifican detalladamente la calidad del sonido estereofónico de dos canales y multicanal**

Para calidad de imagen estereofónica de audio de dos canales se evalúa según los emplazamientos de la imagen sonora y sensaciones de profundidad y realidad del elemento de audio.

2.3.1.5 Material del programa

El contenido general de la prueba no debe ser ni muy llamativa ni muy aburrido, de manera que genere un balance en la experiencia de cada participante (Deutsch, Reimes, Fiebig, & Oehler, 2019). Debido al enfoque del proyecto en pistas de audio comercial, es necesario buscar los temas musicales más exitosos en la actualidad, ya que estos son los que se reproducen con mayor frecuencia dentro del público.

2.3.1.6 Dispositivos de reproducción

Esta sección exige el uso de auriculares que permitan la reproducción de la señal de forma clara, lo más limpia posible para todo tipo de pruebas a realizar. Haciendo uso de audífonos, se pueden generar condiciones de escucha mucho más similares para todos los participantes, sin tener influencia considerable de la acústica del recinto, así como aplicar varias pruebas en simultáneo sin problemas (Crawford-Emery & Lee, 2014).

2.3.1.7 Condiciones de escucha

Se tienen en cuenta factores como el nivel de reproducción, características acústicas de la sala y posicionamiento del participante dentro de esta (Koehl & Paquier, 2013). Debido a que las pruebas realizadas en este proyecto no requieren de altavoces, no se tiene en cuenta la acústica de la sala. Al tener reproducción por audífonos, solo se requiere de una sala con bajo ruido de fondo (preferiblemente ajustado a NR 10) y sin elementos que puedan distraer al participante durante la prueba.

3 Metodología

3.1 Enfoque de Investigación

El enfoque de esta investigación es empírico-analítico (Hernández, Fernández, & Baptista, 2014), ya que se buscó diseñar un sistema de simulación de crosstalk para dispositivos móviles Android y, según los antecedentes, este tipo de aplicativos no han sido desarrollados a profundidad para esta plataforma. Se observaron los resultados de las evaluaciones subjetivas para medir el impacto que el procesamiento realizado por la aplicación tiene sobre la calidad sonora de pistas de audio escuchadas a través de audífonos en un dispositivo móvil.

3.2 Etapas de la Metodología

A continuación, se explican las etapas en las que se organizó la metodología para ser desarrollada.

3.2.1 Evaluación de los métodos

En esta etapa se realizaron evaluaciones subjetivas de calidad sonora de los cuatro métodos de procesamiento de señal para simulación de crosstalk. El proceso de evaluación fue basado en las recomendaciones hechas por la ITU (International Telecommunication Union, 2003) para evaluaciones subjetivas de calidad sonora y análisis de resultados. La prueba aplicada en esta etapa fue de pares comparados. Los diferentes métodos de procesamiento de señal se realizaron con la ayuda del software MATLAB. El análisis estadístico realizado fue una escala de preferencia de Thurstone (Morosini, 2012).

3.2.2 Desarrollo de la aplicación móvil

Para la segunda etapa se seleccionaron los dos métodos de procesamiento con los mejores resultados obtenidos según las evaluaciones de la etapa anterior y fueron implementados en una aplicación móvil. La metodología de desarrollo de software fue la propuesta por Texas Instruments para DSP con Link for Code Composer Studio 3.0, también usada por Ortega, Orjuela y Cortes (2008) para desarrollo de un software de reverberación por convolución. La Figura 12 representa el ciclo de la metodología de desarrollo de software. En primer lugar, se hace una lista con todos los requerimientos del programa. Posteriormente, se realiza el diseño de la interfaz gráfica y de todas las funciones que debe realizar para cumplir los requisitos. Después de esto, se hace la codificación del software. Por último, se verifica el funcionamiento del programa. En caso de que no cumpla los requisitos o tenga algún fallo, se repite este ciclo realizando las correcciones necesarias.

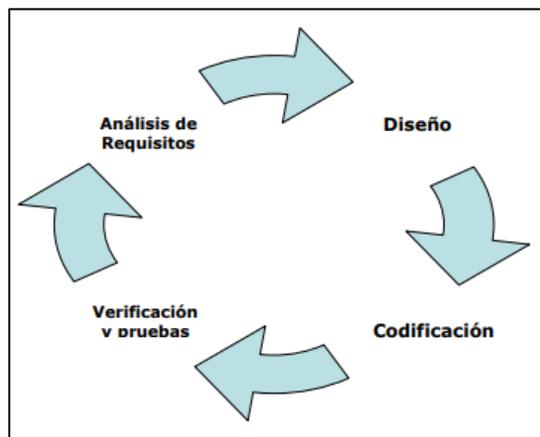


Figura 12. Ciclo de desarrollo de software. Recuperado de: (Ortega, Orjuela, & Cortes, 2008)

3.2.3 Evaluación del producto

Teniendo el producto final del desarrollo del software, se realizaron evaluaciones subjetivas de pistas de audio procesadas por la aplicación y reproducidas a través de

audífonos. Estas evaluaciones también fueron basadas en las recomendaciones de la ITU para este tipo de estudios. El tipo de prueba aplicado en esta tercera etapa fue una escala semántica, con la cual se evaluaron 4 atributos del audio: impresión estereofónica, impresión espacial, timbre y calidad general (Bolaños, Pilkki, & Mäkivirta, 2016).

3.3 Variables

3.3.1 Variables dependientes

- Preferencia.
- Impresión estereofónica.
- Impresión espacial.
- Timbre.
- Calidad general.

3.3.2 Variable independiente

- Método de procesamiento digital del audio estereofónico para simulación de crosstalk en audífonos.
- Género musical del tema presentado.

3.4 Hipótesis

Mediante el desarrollo e implementación de una aplicación móvil para dispositivos Android que aplique la simulación de crosstalk es posible mejorar la calidad sonora de pistas de audio reproducidas a través de audífonos en dichos dispositivos.

4 Desarrollo Ingenieril

En esta sección se presenta el proceso de desarrollo de este proyecto. Para ello, en primer lugar, se seleccionaron los métodos de procesamiento a evaluar a partir de los antecedentes de la sección 1.2. Dichos métodos corresponden a Kirkeby (2002), Manor (2015), Gilchrest (2016) y Bolaños (2016). Se realizó la programación de cada método en MATLAB, con el fin de realizar las pruebas de preferencia entre los cuatro métodos a través de la metodología de pares comparados a 25 oyentes entrenados (estudiantes de Ingeniería de Sonido de Octavo Semestre o mayores). Posteriormente, se implementaron los dos métodos de mayor preferencia en una aplicación para dispositivos Android desarrollada en Unity para reproducir archivos de audio y procesarlos mediante dichos métodos. Por último, se evaluó de calidad sonora de pistas de audio reproducidas a través de la aplicación móvil mediante pruebas de escala semántica a estudiantes de ingeniería de sonido y adicionalmente a oyentes más expertos con trayectoria en producción musical.

4.1 Evaluación de Preferencia para Métodos de Simulación de Crosstalk

4.1.1 Desarrollo y aplicación de los métodos de simulación de crosstalk

En primer lugar, se desarrollaron los algoritmos de procesamiento de señal en MATLAB, para procesar fragmentos de audio mediante los 4 métodos de simulación de crosstalk propuestos anteriormente, con el fin realizar pruebas de preferencia subjetivas. Primero, para la obtención de las muestras de un audio se hizo uso del comando *audioread()*, separando los canales L y R del audio estereofónico.

Debido a la falta de información en la bibliografía (por ejemplo, no se especifica si el filtro es Butterworth, Bessel, etc., ni los coeficientes de este), se diseñaron filtros que cumplieran las necesidades básicas del método, es decir, el tipo de filtro implementado (low-

pass, high-pass, shelving, etc.) y la frecuencia de corte. A continuación, se explica el funcionamiento de cada uno de estos.

4.1.1.1 Método de Kirkeby

La Figura 13 resume el procesamiento digital realizado por el método de Kirkeby, representando los canales de entrada como $xL(n)$ y $xR(n)$ así como las señales de salida $yL(n)$ y $yR(n)$.

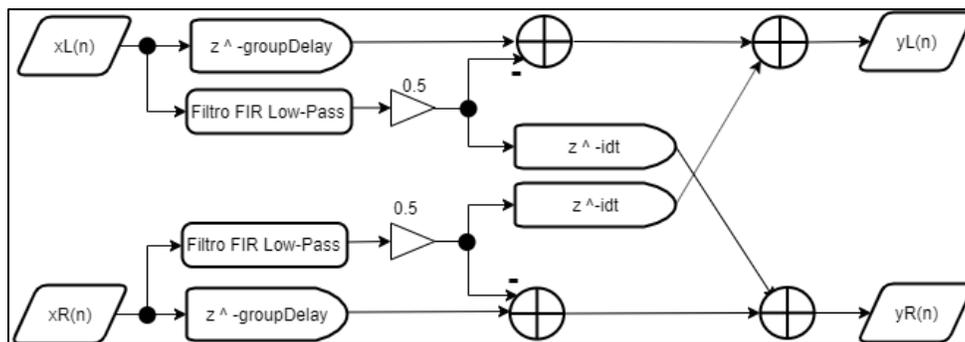


Figura 13. Diagrama de bloques de procesamiento digital mediante el método de Kirkeby.

En este método se usó la herramienta Filter Designer de MATLAB para diseñar un filtro FIR pasa-bajas con una banda de paso hasta 1.5 kHz y una banda de rechazo desde 2kHz a una frecuencia de muestreo de 44.1kHz. La Figura 14 muestra el filtro diseñado para el método de Kirkeby.

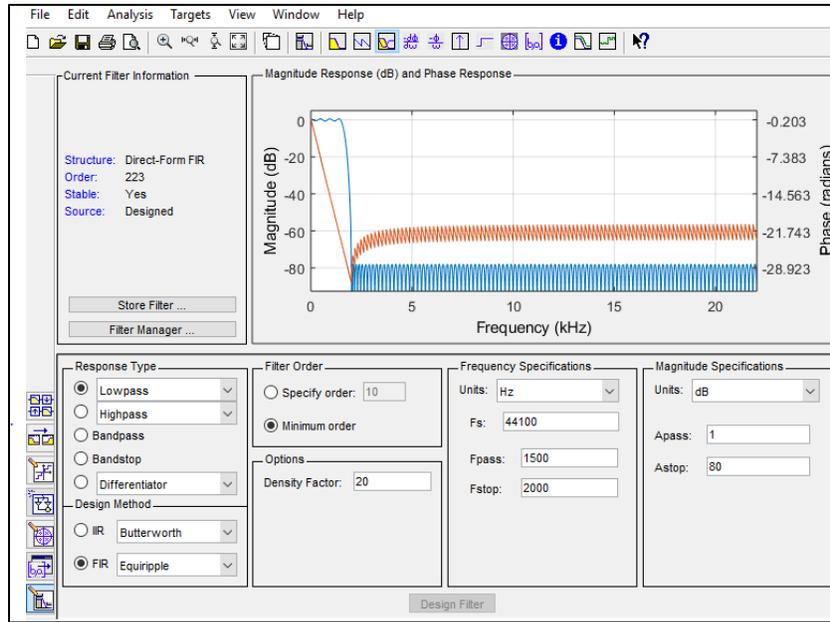


Figura 14. Respuesta en frecuencia (azul) y fase (rojo) de filtro FIR pasa-bajas diseñado en MATLAB.

Los coeficientes obtenidos fueron exportados al código usado para el procesamiento de los audios en la variable *firCoefs*. Los canales L y R del audio fueron filtrados mediante estos coeficientes y atenuados 6dB, como lo indica la Figura 15. El método *filter()* realiza el filtrado de señal mediante el algoritmo de la Ecuación 1. Para filtros FIR, los coeficientes de retroalimentación no existen, por lo que solo se ingresa el número 1 en el segundo parámetro de entrada. Los canales L y R filtrados fueron almacenados en las variables *audioL_FIR* y *audioR_FIR* respectivamente.

```

%% Balanced Stereo Widening Network - Kirkeby

% Filtrado de señal original
audioL_FIR = 0.5 * filter(firCoefs,1,audioL);
audioR_FIR = 0.5 * filter(firCoefs,1,audioR);

```

Figura 15. Filtrado de señal a partir de filtro FIR diseñado, implementando el comando *filter()* de MATLAB.

Se crearon dos arreglos (*delayedAudioL* y *delayedAudioR*) con los canales originales retrasados un número de muestras igual al delay de grupo del filtro FIR diseñado, en este caso, 112 muestras, como lo indica la Figura 16.

```
% Delay de señal original
delayedAudioL = zeros(numDatos,1);
delayedAudioR = zeros(numDatos,1);
delayNGroup = 112; % Group Delay del Filtro

for i = numDatos:-1:delayNGroup+1
    delayedAudioL(i) = audioL(i-delayNGroup);
    delayedAudioR(i) = audioR(i-delayNGroup);
end
```

Figura 16. Generación de canales originales retrasados 112 muestras.

De manera similar, se crearon dos arreglos (*delayedFIRL* y *delayedFIRR*) con los canales filtrados con un retraso equivalente a 0.7ms, es decir, 30 muestras a una frecuencia de muestreo de 44.1 kHz. La Figura 17 representa el retraso generado.

```
% Delay de señal filtrada atenuada
delayedFIRL = zeros(numDatos,1);
delayedFIRR = zeros(numDatos,1);
delayFIRSamples = 30;

for i = numDatos:-1:delayFIRSamples+1
    delayedFIRL(i) = audioL_FIR(i-delayFIRSamples);
    delayedFIRR(i) = audioR_FIR(i-delayFIRSamples);
end
```

Figura 17. Retraso de 0.7ms de canales filtrados.

Por último, se generaron los canales de salida (*outL_K* y *outR_K*) a partir de las señales previamente procesadas, como se ilustra en la Figura 18.

```

% Generacion de canales de salida
outL_K = zeros(numDatos,1);
outR_K = zeros(numDatos,1);

for i = 1:numDatos
    outL_K(i) = (delayedAudioL(i) - audioL_FIR(i) + delayedFIRR(i));
    outR_K(i) = (delayedAudioR(i) - audioR_FIR(i) + delayedFIRL(i));
end

```

Figura 18. Creación de canales de salida para audio estereofónico procesado mediante el método de Kirkeby.

4.1.1.2 Método de Manor

En este método se usaron dos tipos de filtros: un shelving en alta frecuencia y un pasabajas. En la Figura 19 se observa el diagrama de bloques de este método de procesamiento. El valor de diferencia de tiempo entre cada canal fue de 0.5 ms, es decir, 22 muestras a 44.1 kHz.

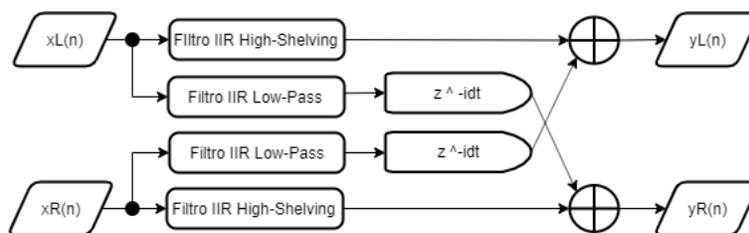


Figura 19. Diagrama de bloques de procesamiento digital mediante el método de Manor.

El filtro Shelving fue diseñado a una frecuencia de corte de 1kHz, con una ganancia de 3dB a frecuencias más bajas y 6B a frecuencias más altas. Por otro lado, el filtro pasabajas fue diseñado a una frecuencia de 1.7 kHz. En este método no se especificaba el uso de filtros FIR o IIR, por lo que se eligió implementar filtros IIR, cuya respuesta en frecuencia fuese similar a la de los filtros usados en la bibliografía. La obtención de la ecuación de entrada y salida del filtro shelving de alta frecuencia se realiza mediante los cálculos realizados por Zölzer (2002) y representados de la Ecuación 13 hasta la Ecuación 17.

$$A_B = \frac{\tan\left(\frac{\pi f_c}{f_s}\right) - 1}{\tan\left(\frac{\pi f_c}{f_s}\right) + 1} \quad \text{Ecuación 13}$$

$$V_0 = 10^{G/20}; \quad \text{Ecuación 14}$$

$$H_0 = V_0 - 1; \quad \text{Ecuación 15}$$

$$y_1(n) = A_B x(n) + x(n-1) - A_B y_1(n-1); \quad \text{Ecuación 16}$$

$$y(n) = \frac{H_0}{2} [x(n) - y_1(n)] + x(n) \quad \text{Ecuación 17}$$

teniendo que:

- f_c es la frecuencia de corte del filtro, en este caso, 1000 Hz,
- f_s es la frecuencia de muestreo del audio, en este caso, 44.1 kHz,
- $x(n)$ la señal de entrada en su muestra n ,
- $y(n)$ la señal de salida en su muestra n ,
- G el nivel en dB de ganancia para alta frecuencia, en este caso, +3dB.

La Ecuación 3 indica la obtención de la función de transferencia para un filtro digital a partir de sus coeficientes. Según esto, se puede hallar la función de transferencia H_{SH} del filtro shelving mediante la Ecuación 18.

$$H_{SH}(z) = 0.707 \frac{\left(1 + \left[\frac{H_0}{2}(1 - A_B)\right]\right) + \left(A_B + \left[\frac{H_0}{2}(A_B - 1)\right]\right) z^{-1}}{1 + A_B z^{-1}} \quad \text{Ecuación 18}$$

Los coeficientes de realimentación (numerador de la función de transferencia) fueron multiplicados por 0.707, correspondiente a una atenuación de -3dB. Se optó por atenuar en lugar de incrementar el nivel (lo que originalmente estipula el método) para no generar saturación en el audio obtenido. En su lugar, las ganancias de los filtros shelving y low-pass

(+3dB para el filtro shelving y -3dB para el filtro low-pass) fueron cambiadas a -3dB y -9dB respectivamente.

Se usó la función de transferencia análoga de un filtro pasa-bajos de primer orden que se representa en la Ecuación 19 para una frecuencia de corte $f_c=1700$ Hz. Se usó este tipo de filtro debido a la pendiente del filtro pasa-bajos de campo cercano que muestra la Figura 11.

$$H_{LP}(s) = \frac{(2\pi f_c)}{s + (2\pi f_c)} \quad \text{Ecuación 19}$$

Posteriormente, se halló la transformada bilineal de esta función de transferencia análoga para hallar su equivalente en el dominio digital mediante la Ecuación 8. Los valores obtenidos se muestran en la Ecuación 20.

$$H_{LP}(z) = 0.3584 \frac{0.0383 + 0.0383z^{-1}}{1 - 0.784z^{-1}} \quad \text{Ecuación 20}$$

Los coeficientes de realimentación de este filtro fueron multiplicados por 0.3584 para obtener una atenuación de -9dB. De esta forma se cambiaron las ganancias originales del método (+3dB para shelving y -3dB para pasa-bajas) a valores que no generaran distorsión en el audio (-3dB para shelving y -9dB para pasa-bajas).

Los procesos de filtrado y retraso de señal se realizaron de manera similar a la implementación del método de Kirkeby en MATLAB.

4.1.1.3 Método de Gilcrest

Este método se diferencia de los demás por hacer uso de la señal Side del audio estereofónico. La Figura 20 representa el diagrama de bloques del procesamiento digital mediante el método de Gilcrest.

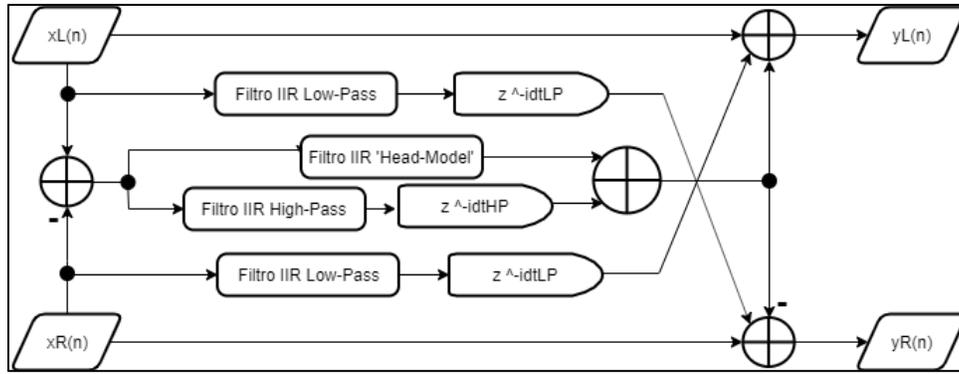


Figura 20. Diagrama de bloques de procesamiento digital mediante el método de Gilcrest.

Se creó un arreglo para la señal Side, correspondiente a la resta punto a punto entre el canal izquierdo y el canal derecho. Se implementó el filtro Head Model (Zölzer, 2002), mediante el cual se procesó la señal Side para un ángulo θ de 30° . Los coeficientes obtenidos para este filtro se observan en la Ecuación 21.

$$H_{HM}(z) = \frac{0.3023 - 0.1293z^{-1}}{1 - 0.827z^{-1}} \quad \text{Ecuación 21}$$

Posteriormente, para los filtros low-pass y high-pass se diseñaron filtros de segundo orden en la forma de filtros análogos, para después realizar la transformada bilineal y obtener los coeficientes de los filtros digitales. La función de transferencia análoga de un filtro pasabajas de segundo orden se da mediante la Ecuación 22 (Tan, 2008).

$$H_{LP}(s) = \frac{(2\pi f_c)^2}{s^2 + \left(\frac{2\pi f_c}{0.707}\right) s + (2\pi f_c)^2} \quad \text{Ecuación 22}$$

Para un filtro pasa-altas de segundo orden su función de se halla como lo indica la Ecuación 23.

$$H_{HP}(s) = \frac{s^2}{s^2 + \left(\frac{2\pi f_c}{0.707}\right)s + (2\pi f_c)^2} \quad \text{Ecuación 23}$$

Los coeficientes calculados del filtro pasa-bajas para una frecuencia de corte de 250 Hz se muestran en la Ecuación 24.

$$H_{LP}(z) = 0.5 \frac{0.0001546 + 0.0003093z^{-1} + 0.0001546z^{-2}}{1 - 1.9496z^{-1} + 0.9509z^{-2}} \quad \text{Ecuación 24}$$

Esta función de transferencia fue multiplicada por 0.5 para generar la atenuación de -6dB propuesta por el método. Por otro lado, para el filtro pasa-altas la función de transferencia digital calculada para 5 kHz fue la mostrada en la Ecuación 25.

$$H_{HP}(z) = 0.5 \frac{0.3066 - 0.6132z^{-1} + 0.3066z^{-2}}{1 - 1.0709z^{-1} + 0.3821z^{-2}} \quad \text{Ecuación 25}$$

Los canales L y R originales fueron procesados mediante el filtro pasa-bajas, y retrasados 20ms. Por otro lado, la señal Side original fue procesada por el filtro pasa-altas, atenuada 6dB y retrasada 15ms. Finalmente se generaron los canales de salida acorde al diagrama de bloques que se representa en la Figura 20.

4.1.1.4 Método de Bolaños

La Figura 21 representa el diagrama de bloques del procesamiento digital para el método de Bolaños, siendo $H(i,j)$ la respuesta al impulso ecualizada y obtenida con el micrófono i y el altavoz j .

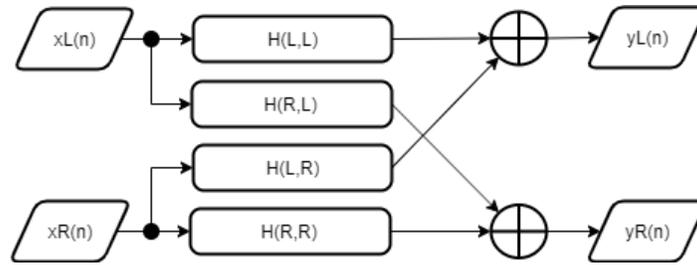


Figura 21. Diagrama de bloques de procesamiento digital mediante el método de Bolaños.

En este método se realizó una medición de respuesta al impulso binaural en el estudio de Mastering de la Universidad de San Buenaventura Sede Bogotá, donde se ubicó la cabeza binaural en el Sweet Spot, como se nota en la Figura 22. Se eligió esta locación debido a que es un recinto acústicamente tratado (Bolaños, Pilkki, & Mäkivirta, 2016).



Figura 22. Ubicación de la cabeza binaural en el sweet spot del estudio de Mastering.

Se capturaron dos respuestas al impulso binaurales: excitando monitor izquierdo y excitando el monitor derecho. Cada una fue capturada por medio de cada canal de la cabeza binaural. Debido a que no funcionaban los preamplificadores de la cabeza binaural, se usaron dos micrófonos DPA 4060 ubicados en las orejas de la cabeza, como lo ilustra la Figura 23. Dichos micrófonos fueron calibrados a 94 dB mediante un pistófono Svantek.



Figura 23. Ubicación de un micrófono DPA dentro de la oreja de la cabeza binaural.

Se usó el software EASERA para generar un barrido sinusoidal desde 20 Hz hasta 20 kHz con duración de 3s. Esta señal fue reproducida por el sistema de monitores presente en el estudio de Mastering, el cual constó de un subwoofer ADAM SUB 10 MK2 y dos monitores ADAM A77X. Cada respuesta al impulso binaural fue exportada a formato WAV a una frecuencia de muestreo de 48 kHz. Posterior a esto, se cambió la frecuencia de muestreo a 44.1kHz mediante MATLAB con el comando *resample()*. Además, cada archivo de audio fue recortado a 40ms, debido a que respuestas al impulso binaurales de mayor tamaño no tienen un mayor efecto en la espacialidad generada (Crawford-Emery & Lee, 2014). En la Figura 24 se puede observar el diagrama de conexionado empleado para esta medición. LA interfaz RME 4000 recibió la señal de los micrófonos, mientras que la interfaz AVOCET se empleó para enviar señal análoga al sistema de altavoces del estudio.

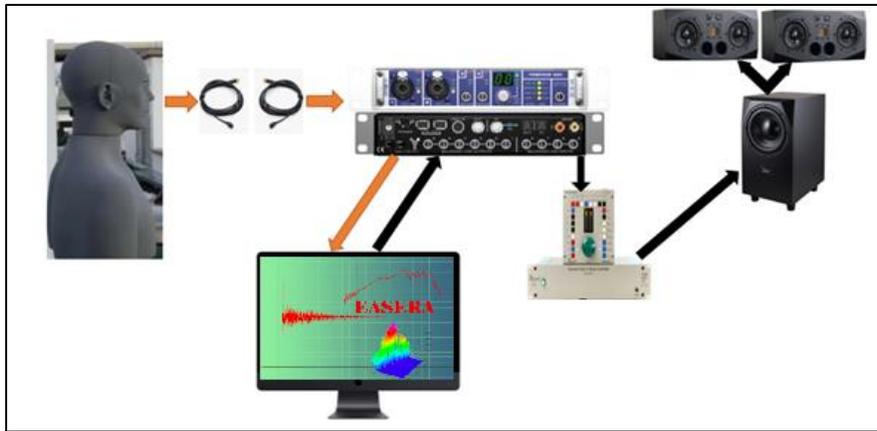


Figura 24. Diagrama de conexionado empleado para la medición de respuesta al impulso binaural dentro del estudio de Mastering.

En la Figura 25 se ven las respuestas al impulso binaurales obtenidas de la medición en el estudio de Mastering, de arriba hacia abajo según la combinación altavoz-micrófono: altavoz izquierdo – micrófono izquierdo (1), altavoz izquierdo– micrófono derecho (2), altavoz derecho – micrófono izquierdo (3), y altavoz derecho – micrófono derecho (4).

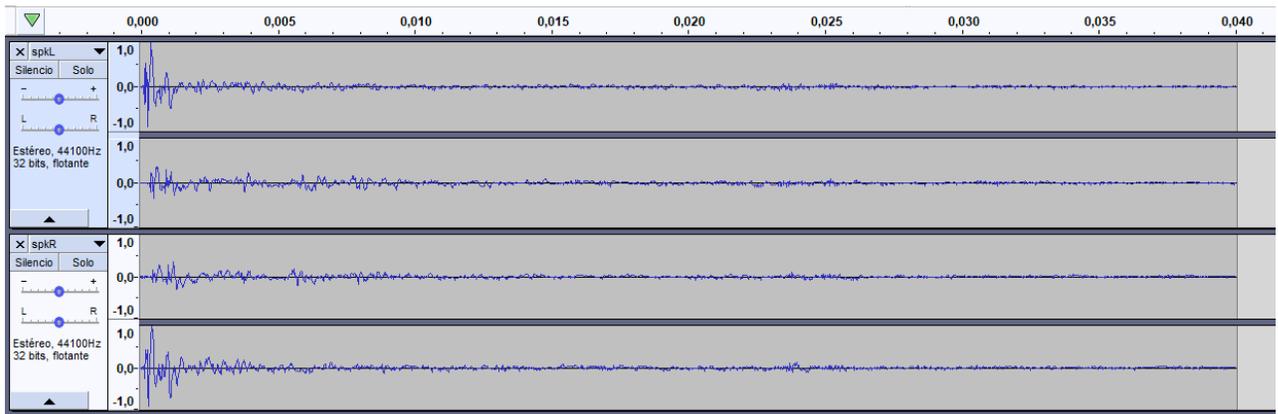


Figura 25. Respuestas al impulso binaurales obtenidas de la medición.

Con cada respuesta al impulso binaural, se procedió a realizar la convolución discreta (Ecuación 9) de dichas respuestas con el tema musical “*Here Comes The Sun*” de The Beatles, el cual se obtuvo del disco original en formato WAV (muestreo de 44.1kHz y

profundidad de 16 bits). El audio procesado se obtuvo mediante la convolución de cada canal (L y R) del audio con la respuesta al impulso binaural obtenida en cada altavoz (izquierdo y derecho) respectivamente.

Teniendo el audio procesado, fue importado a una sesión en Pro Tools. Se realizó una ecualización para generar un balance tonal en el audio procesado con ayuda de un oyente experto, usando audífonos Audio Technica M40X. Finalizado esto, se usó el preset de procesamiento generado por el oyente experto para realizar el bounce de ambas respuestas al impulso binaurales ecualizadas. La Figura 26 ilustra ambas respuestas al impulso binaurales tras su ecualización.

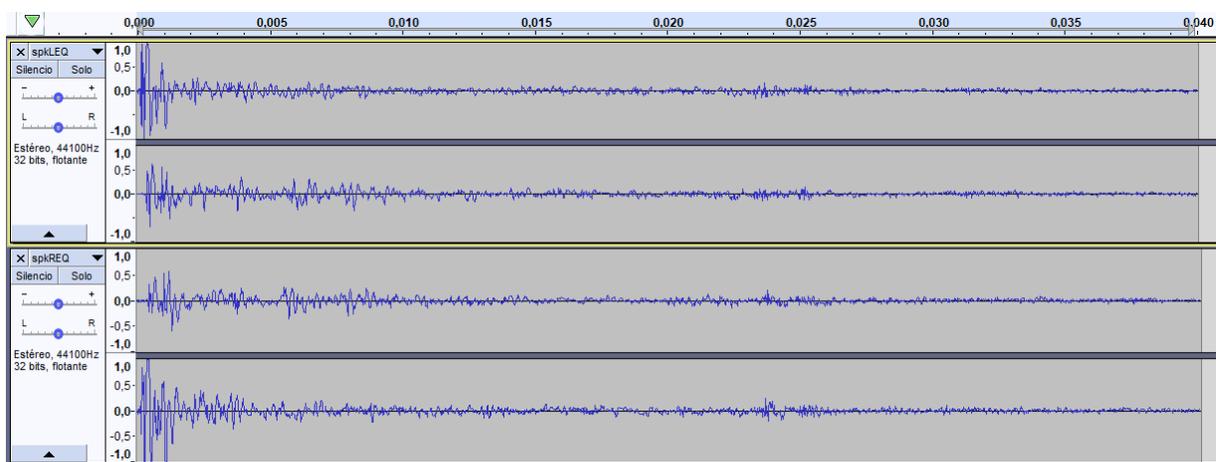


Figura 26. Respuestas al impulso binaurales ecualizadas por un oyente experto.

4.1.2 Diseño y aplicación de pruebas de preferencia

Para las evaluaciones de preferencia se escogieron 3 temas musicales, cada uno de un género distinto teniendo en cuenta aspectos como calidad, características propias de su género y también la popularidad en el año 2019:

- **Ludwig van Beethoven, *Symphony No. 5 in C minor Part 1, Classical Best of Beethoven (2009)***: Esta canción data de inicios del siglo XIX y fue obra del compositor alemán Beethoven, uno de los más importantes en el contexto de la música clásica. El disco que contiene este archivo de audio data de una grabación realizada en 2009 (Cooper, 2008).
- **Pink Floyd, *Money, The Dark Side of The Moon (1973)***: Este álbum es considerado como uno de los mejores, no solo de su género sino de toda la historia de la música (Reising, 2005).
- **Daddy Yankee, *Con Calma (Sencillo) (2019)***: Este sencillo se encontraba en el primer puesto de la lista de Billboard Latino durante las semanas en las que se inició este proyecto (Billboard, 2019).

Se optó por estas tres canciones para lograr una variedad en los géneros musicales (música clásica sinfónica, rock progresivo y urbano) al tener en cuenta los antecedentes (Lorho, Isherwood, Zacharov, & Huopaniemi, 2002), permitiendo evaluar cómo la calidad sonora de cada género se ve afectada por los diferentes métodos de procesamiento. Teniendo los archivos originales de cada tema en formato wav (44.1kHz, 16bits) se extrajo un fragmento de 15 segundos de cada uno. Éstos fueron normalizados a -18dBFS RMS mediante Audacity. Además, a cada uno se le aplicó un Fade In y Fade Out de 0.5 segundos para suavizar su reproducción durante la prueba.

En total, se usaron 15 fragmentos de audios, 5 audios por cada tema (4 procesados y original). Se diseñó una interfaz gráfica en MATLAB, en la cual se aplicó el método de pares comparados (ver Ley de Juicio Comparativo). En primer lugar, la interfaz daba al usuario

una breve introducción al concepto del crosstalk y su simulación en audífonos, para después mostrar los audios por pares comparados. La prueba constaba de 3 fases, en cada una se presentó uno de los temas musicales escogidos. Los audios fueron mostrados en parejas, facilitando la ejecución de la prueba a tan solo decidir cuál de los dos audios presentados era de mayor preferencia para el oyente. Tanto las combinaciones por pares como el orden de presentación de los temas fueron aleatorios. La Figura 27 ejemplifica una prueba de pares comparados en la interfaz de MATLAB, en la cual el participante debía presionar el botón correspondiente al audio de su preferencia.

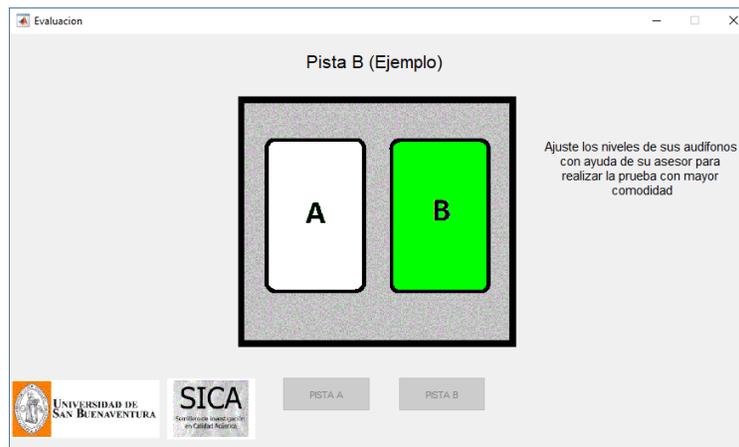


Figura 27. Ejemplo de prueba de pares comparados en la interfaz gráfica desarrollada en MATLAB.

Las pruebas de preferencia se realizaron en los estudios de grabación de la Universidad de San Buenaventura Sede Bogotá, los cuales presentan bajos niveles de ruido de fondo y un ambiente controlado, permitiendo a los participantes de la prueba mayor concentración. El jurado que tomó la prueba fue de un total de 25 personas, todos ellos estudiantes de Ingeniería de Sonido, cursando Octavo Semestre o superiores. Cada una de las personas leyó y aceptó el consentimiento de la prueba, en el cual se indicaban los términos y condiciones, el objetivo de la prueba, la manera en cómo iban a ser tratados sus datos y las

tareas que debería realizar mientras desarrollaba la prueba. En promedio cada prueba tenía una duración de 30 a 40 minutos.

En la Figura 28 se observa el diagrama de conexionado empleado en la prueba de preferencia, donde el sujeto de prueba usaba audífonos Audio Technica M40X conectados a una interfaz de audio Behringer UMC- 202 HD. Cada oyente hacía uso de un computador portátil para la realización de la prueba, a través de la interfaz gráfica desarrollada en MATLAB. Es importante mencionar que cada persona regulaba el nivel de ganancia de los audífonos según su confort y necesidad para tener una escucha con un nivel adecuado.



Figura 28. Diagrama de conexionado empleado para la realización de las pruebas de preferencia entre métodos de simulación de crosstalk.

4.1.3 Aplicación de la Ley de Juicio Comparativo

Con el fin de realizar una prueba en donde se presenten el menor número de variables posibles, para obtener datos de preferencia de un método de procesamiento sobre otro, se optó por implementar el método de escalamiento de Thurstone (ver Ley de Juicio Comparativo) usando el juicio de pares comparativos.

En esta sección se explica la manera en la que se implementó el método de comparación binaria, mostrando el procedimiento llevado a cabo para la canción Symphony No. 5, sabiendo que el mismo procedimiento y análisis de datos fue realizado para las otras dos canciones (Money y Con Calma), cuyo análisis se puede ver en los Resultados y aplicación de la ley Thurstone para prueba de preferencia más adelante.

En la Tabla 1 se muestra la matriz de frecuencias observadas ordenada según la suma de las frecuencias por cada columna vertical, donde se ubicaron los datos de preferencia de un método respecto a otro a través de la prueba de preferencia. En cada celda se encuentra la frecuencia con la que fue seleccionado cada método al compararse contra otro. Al existir un par de métodos iguales, se escribe con cero (línea diagonal) ya que no es válido hacer juicios de igualdad.

Tabla 1. Matriz de frecuencia observadas de los diferentes métodos de procesamiento para el tema Symphony No. 5.

Symphony No 5					
	2	5	3	1	4
1	4	9	15	0	12
2	0	10	15	21	18
3	10	8	0	10	16
4	7	9	9	13	0
5	15	0	17	16	16
Total	36	36	56	60	62

De lo anterior:

- 1 representa la canción original sin procesamiento.
- 2 representa el método de Kirkeby.
- 3 representa el método de Manor.
- 4 representa el método de Gilchrest.
- 5 representa el método de Bolaños.

Para explicar la anterior, el primer valor (fila '1', columna '2') equivale a 4, de manera que el método 2 fue elegido 4 veces al compararse con el método 1 (**2 vs 1**). En contraste, al ver el valor de la fila '2' y la columna '1', equivalente a 21, el método 1 fue elegido 21 veces al compararse con el método 2 (**1 vs 2**). La suma de estos dos valores es 25, el número total de pruebas aplicadas. Nótese que el valor en la fila '1' y la columna '1' es 0, por no haber aplicado juicios de igualdad. En la Tabla 2 se puede ver la matriz de proporciones en la cual se expresa el número de oyentes que seleccionó cada estímulo sobre la totalidad de la muestra (n=25).

Tabla 2. Matriz de proporciones ordenada de los datos para el tema Symphony No. 5.

Matriz de Proporciones					
Symphony No 5					
	2	5	3	1	4
1	0,16	0,36	0,6	0	0,48
2	0	0,4	0,6	0,84	0,72
3	0,4	0,32	0	0,4	0,64
4	0,28	0,36	0,36	0,52	0
5	0,6	0	0,68	0,64	0,64
DE=0,2559					

Por ejemplo, en el caso de **1 vs 5** en donde se tiene una frecuencia observada de 9 juicios de preferencia, se sabe que la muestra es de n=25, la proporción de preferencia es de 0,36. Cada valor de la matriz de proporciones fue calculado mediante la Ecuación 26. Además, se calculó la desviación estándar de todos los datos obtenidos en este caso.

Ecuación 26

$$P = \frac{\text{Frecuencia de preferencia}}{\text{Numero total de oyentes}}$$

Después de hallar los valores de las proporciones, es necesario normalizar con su correspondiente puntuación normalizada Z, las cuales son obtenidas en las tablas de

distribución estándar normalizadas (ver Anexo A). Con los valores normalizados, se suman los datos por columnas verticales. Posteriormente se calcula el promedio de cada método de procesamiento dividiendo el resultado de la suma de cada columna entre el número de estímulos presentados (5 en este caso), los cuales se observan en la matriz de proporciones normalizadas. Finalmente se elabora la escala final transformando los promedios al restar el valor inferior, en este caso -0,2533. Estos datos se pueden ver en la Tabla 3, mientras que la Figura 29 ilustra la escala unidimensional de preferencia obtenida.

Tabla 3. Matriz de proporciones normalizadas Z y escala de preferencia (gris) para la canción Symphony No. 5.

Matriz de proporciones normalizadas					
Symphony No 5					
	2	5	3	1	4
1	-0,9945	-0,3585	0,2533		-0,0502
2		-0,2533	0,2533	0,9945	0,5828
3	-0,2533	-0,4677		-0,2533	0,3585
4	-0,5828	-0,3585	-0,3585	0,0502	
5	0,2533		0,4677	0,3585	0,3585
SUMA	-1,5773	-1,4380	0,6159	1,1497	1,2496
PROMEDIO	-0,3155	-0,2876	0,1232	0,2299	0,2499
DISTRIBUCIÓN Z	0,0000	0,0279	0,4386	0,5454	0,5654

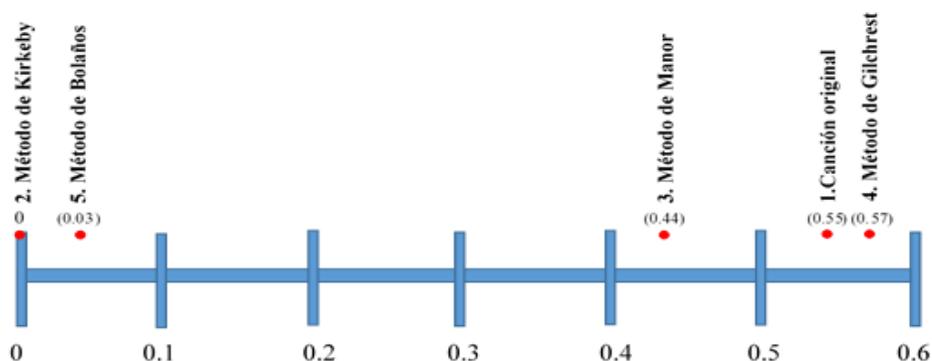


Figura 29. Resultado del escalamiento de preferencia para el tema Symphony No. 5.

Como se puede observar en la Figura 29, el primer valor de la escala de referencia es el 0, el cual equivale al valor del método 2 (Método de Kirkeby), indicando que el procesamiento de audio menos seleccionado por los oyentes según su preferencia fue dicho método para el tema Symphony No. 5. A partir de allí los diferentes métodos se clasifican según sus unidades de desviación estándar (Tabla 3, última fila). El método 4 (Método de Gilchrest) fue el de mayor según los oyentes encuestados. Los resultados de la preferencia para cada tema musical y las 3 en conjunto se analizan más adelante.

4.2 Desarrollo de la aplicación móvil

Teniendo los dos mejores métodos de simulación de crosstalk (métodos de Manor y de Gilchrest), basándose en las pruebas realizadas en la etapa anterior, el siguiente paso fue desarrollar una aplicación para dispositivos móviles, capaz de procesar y reproducir archivos de audio mediante los 2 métodos mencionados.

4.2.1 Plataforma de desarrollo

La primera opción de plataforma para el desarrollo del programa fue JUCE, un marco de aplicación de código abierto especializado en software de audio, como plugins o aplicaciones ‘stand alone’. Sin embargo, a la hora de intentar reproducir archivos de audio en un dispositivo Android Moto G4+ se generaban clicks digitales y el audio no se reproducía correctamente ni a una buena calidad. Tras este intento, la siguiente opción evaluada fue Unity, un motor de videojuegos multiplataforma. Aquí fue posible tanto reproducir como procesar archivos de audio. El procesamiento realizado no fue en tiempo real, sino que se tomó todo el arreglo de muestras del audio y se generó todo el arreglo de salida, para después reproducirlo. Al exportar el proyecto al dispositivo Android se reprodujo el audio sin ningún

tipo de inconveniente. Por este motivo se optó por desarrollar la aplicación en Unity, haciendo uso de códigos creados en el lenguaje de programación C#. El link para visualizar el proyecto de Unity de la aplicación desarrollada puede consultarse en el Anexo C.

4.2.2 Estructura del software

Se planteó un diagrama de flujo básico que debía seguir la aplicación, el cual está representado en la Figura 30.

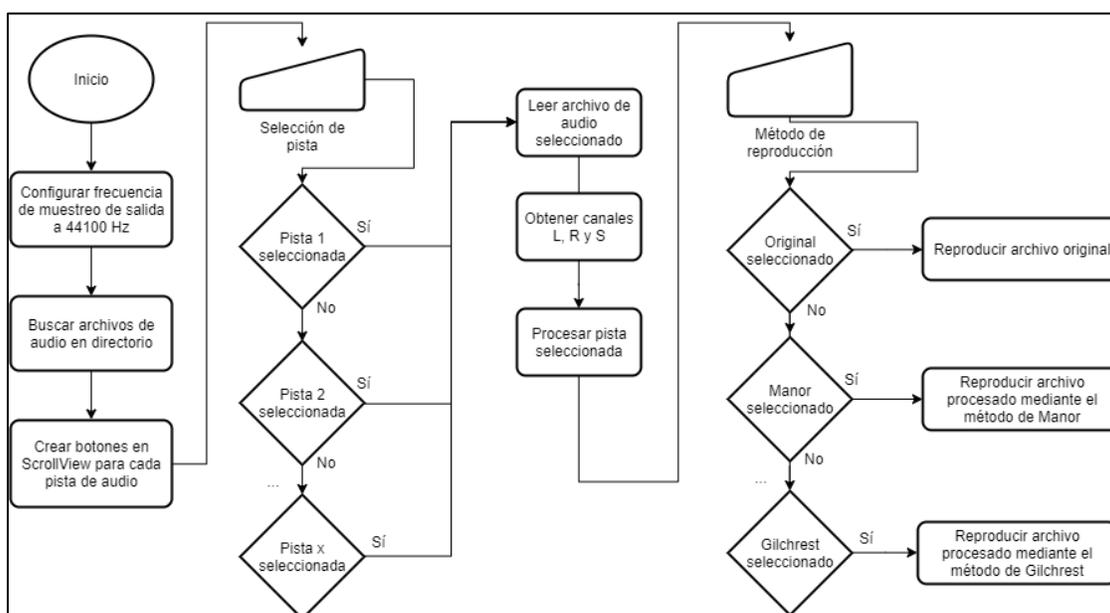


Figura 30. Diagrama de flujo de la aplicación desarrollada.

La aplicación desarrollada se divide en 2 grandes componentes, correspondientes a los GameObjects principales: 'AudioPlayer' y 'GUI'.

4.2.2.1 GameObject 'AudioPlayer'

Este objeto maneja todo lo relacionado al audio, es decir, cargar el archivo wav seleccionado, procesarlo mediante ambos métodos de simulación de crosstalk y reproducirlo.

Las clases creadas para este objeto fueron:

- **PlayClip.cs:** Es la clase central de la aplicación. Da las instrucciones de carga, procesamiento y reproducción del audio.
- **LoadAudioFiles.cs:** Carga el audio desde una dirección y notifica a la clase PlayClip que el audio está listo para ser procesado.
- **GetChannels.cs:** Se le asigna un AudioClip y obtiene los arreglos de muestras correspondientes a los canales izquierdo y derecho, así como la señal Side.
- **Filter.cs:** Filtra un archivo de audio contenido en un arreglo de tipo flotante mediante procesamiento digital. Sus entradas son el arreglo con las muestras de audio y los coeficientes de realimentación y retroalimentación del filtro.
- **ManorProcessor.cs:** Recibe un archivo de audio contenido en un arreglo y lo procesa mediante el método de Manor, haciendo uso de la clase *Filter.cs* con los coeficientes propios del método.
- **GilchrestProcessor.cs:** Recibe un archivo de audio contenido en un arreglo y lo procesa mediante el método de Gilchrest.

Además, se le asignó un componente 'AudioSource' a este GameObject con el fin de reproducir el 'AudioClip' seleccionado por el usuario, como ilustra la Figura 31.

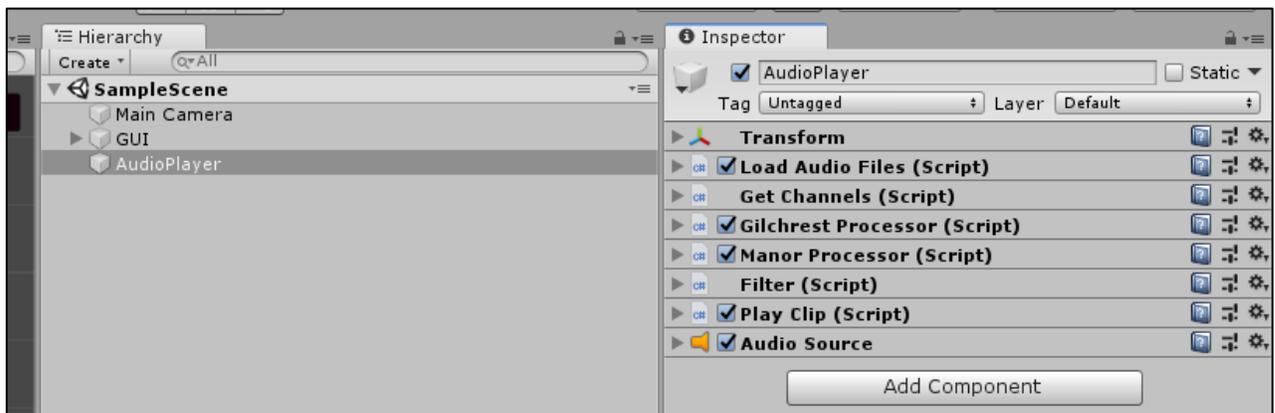


Figura 31. Clases implementadas en el GameObject 'AudioPlayer'.

4.2.2.2 GameObject 'GUI'

Los componentes de la interfaz gráfica de la aplicación son:

- **ScrollView:** Genera una lista de los archivos de audio encontrados. Permite que el usuario seleccione el archivo de audio que desee reproducir.
- **Slider:** Se actualiza constantemente para mostrar la posición actual de reproducción para el archivo seleccionado. Al desplazar su posición cambia el tiempo de reproducción del archivo seleccionado.
- **Botón Play/Pause:** Permite al usuario pausar y reanudar la reproducción.
- **Botón 'Original':** Reproduce el archivo de audio original.
- **Botón 'Manor':** Reproduce el archivo de audio procesado mediante el método de Manor.
- **Botón 'Gilchrest':** Reproduce el archivo de audio procesado mediante el método de Gilchrest.

En la Figura 32 se ilustra la interfaz gráfica de la aplicación desarrollada.

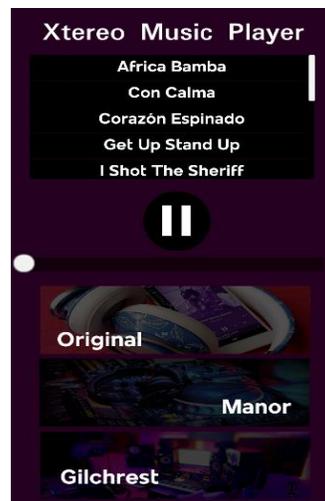


Figura 32. Interfaz gráfica de la aplicación desarrollada.

La Figura 33 ilustra el diagrama de bloques de la aplicación desarrollada, con los componentes de la interfaz gráfica en color azul y las funciones de cada clase implementada en color beige.

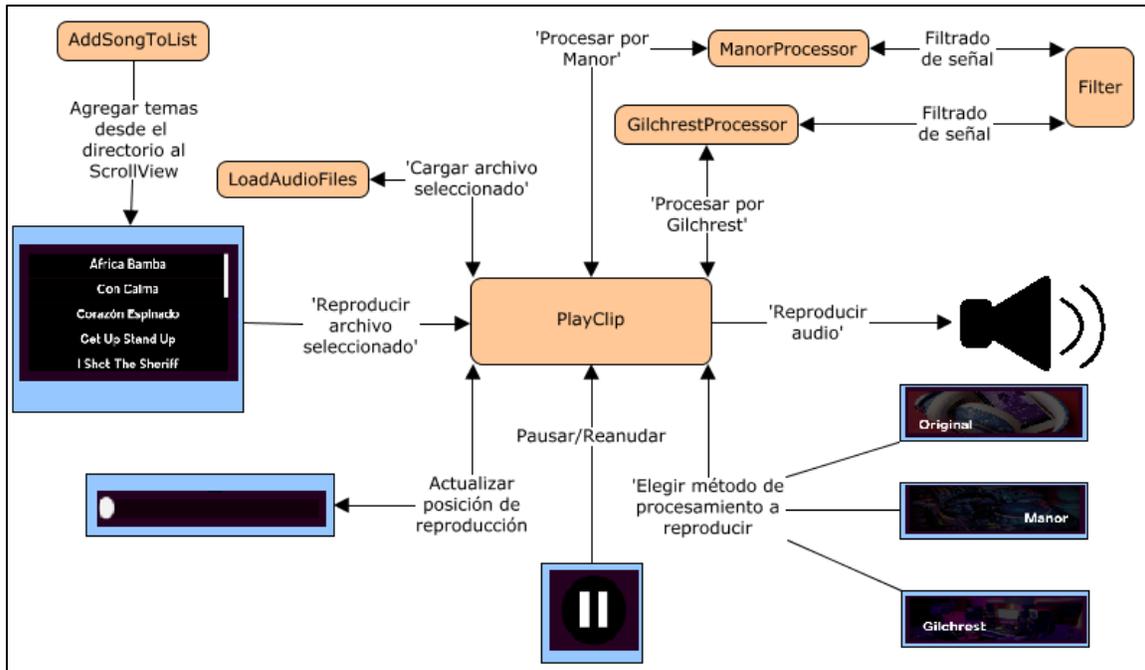


Figura 33. Diagrama de bloques para las clases implementadas en la aplicación.

4.2.3 Funcionamiento de la aplicación

En primer lugar, la clase *PlayClip* configura la frecuencia de muestreo de salida a 44.1 kHz. La clase *AddSongToList* registra el directorio establecido mediante la función *GetFilesInDirectory()* y añade un botón al *GameObject ScrollView* por cada archivo de audio en formato .wav (44.1 kHz, 16 bits) encontrado con ayuda de la función *ShowFilesInScrollView()*. A cada botón se le asigna una función para ejecutar al ser presionado, la cual activa la función *PlayAudio()* de la clase *PlayClip* con la dirección y el nombre del archivo seleccionado para reproducción.

Al seleccionar una pista de audio en el ScrollView, la función *PlayAudio()* activa la función *LoadClip()* de la clase *LoadAudioFiles*. Aquí se obtiene un ‘AudioClip’ a partir del archivo seleccionado, el cual es asignado a la clase *PlayClip*. Después de esto, la clase *PlayClip* hace uso de las clases *ManorProcessor* y *GilchrestProcessor* para procesar el ‘AudioClip’ seleccionado mediante los métodos de simulación de crosstalk de Manor y Gilchrest. Cada método hace uso de la clase *GetChannels* para obtener los canales L y R (Gilchrest también obtiene la señal Side) mediante la función *SetChannels()*. Ésta separa el arreglo unidimensional estereofónico (datos de L y R alternados) y genera arreglos independientes para cada canal. El procesamiento de señal se realizó de la siguiente manera:

- *ManorProcessor*; Procesa los canales L y R mediante los filtros high-shelving y low-pass propios del método de Manor, haciendo uso de la función *FilterSignal()* de la clase *Filter*. Se calculó el arreglo de salida mediante la Ecuación 27 y Ecuación 28.

Ecuación 27

$$y_l(n) = hs_l(n) + lp_r(n - idt)$$

Ecuación 28

$$y_r(n) = hs_r(n) + lp_l(n - idt)$$

Donde:

- $y_i(n)$ es el canal i de salida;
- $hs_i(n)$ es el canal i de entrada procesado por el filtro high-shelving;
- $lp_i(n - idt)$ es el canal i de entrada procesado por el filtro low-pass y retrasado idt muestras. Para un retraso de 0.5ms a una frecuencia de muestreo de 44.1 kHz, $idt = 22$.

- *GilchrestProcessor*: Procesa los Canales L y R mediante el filtro low-pass. Además, procesa la señal Side mediante los filtros ‘Head-Model’ y high-pass. El filtrado se realiza mediante la función *FilterSignal()* de la clase *Filter* Se calcula el arreglo de salida mediante la Ecuación 29 y Ecuación 30:

Ecuación 29

$$y_l(n) = x_l(n) + hm(n) + hps(n - idtHP) + lp_r(n - idtLP)$$

Ecuación 30

$$y_r(n) = x_r(n) - hm(n) - hps(n - idtHP) + lp_l(n - idtLP)$$

Donde:

- $y_i(n)$ es el canal i de salida;
- $x_i(n)$ es el canal i de entrada;
- $hm(n)$ es la señal Side procesada por el filtro ‘Head-Model’;
- $hps(n - idtHP)$ es la señal Side procesada por el filtro high-pass y retrasada $idtHP$ muestras. Para un retraso de 15ms a una frecuencia de muestreo de 44.1 kHz, $idtHP = 661$;
- $lp_i(n - idtLP)$ es el canal i de entrada procesado por el filtro low-pass y retrasado $idtLP$ muestras. Para un retraso de 20ms a una frecuencia de muestreo de 44.1 kHz, $idtHP = 881$.

Los ‘AudioClips’ procesados son asignados a la clase *PlayClip*. La variable privada *selectedMethod* de esta clase varía en función de qué método de reproducción es seleccionado por los botones de la interfaz gráfica.

Una de las limitaciones de la aplicación fue el hecho de que no realizaba procesamiento en tiempo real, de manera que al elegir un tema de la lista se realizaba el procesamiento de todo el audio por cada uno de los métodos, lo cual añadía aproximadamente unos 20 segundos de espera para cargar los 3 AudioClips correspondientes al audio original y al de los dos métodos de simulación de crosstalk implementados. Al tener tantos AudioClips en simultáneo la aplicación presentaba problemas al reproducir archivos de audio de más de 4 minutos de duración, debido a la gran cantidad de memoria necesaria. Fue por esto que cada archivo WAV dentro del directorio fue limitado a 3 minutos.

4.3 Evaluaciones calidad sonora de la aplicación móvil

Para esta prueba se escogieron los mismos temas musicales de la primera etapa (Evaluación de Preferencia para Métodos de Simulación de Crosstalk) cada uno en formato .wav con frecuencia de muestreo de 44.1kHz y 16bits de profundidad. Se extrajo un fragmento de 3 minutos de cada tema. Además, a cada uno se le aplicó un Fade In y Fade Out de 0.5 segundos para suavizar su reproducción durante la prueba.

La prueba constaba de tres fases. La primera parte era introductoria, donde se le explicaba al sujeto el concepto de crosstalk y de los atributos a evaluar. Esta parte también se explicó a detalle en el consentimiento y la primera visualización del Google Form para la aplicación de la prueba (revisar Anexo B: Anexos de prueba de calidad sonora). La segunda fase estuvo destinada a la entrega del dispositivo móvil con la aplicación cargada (Xtereo Music Player) y los audífonos (Audio Technnica M40X). El participante podía familiarizarse con el uso de la aplicación al reproducir alguno de los archivos de audio disponibles. Por último, debía calificar a su criterio qué tan bueno o deficiente era cada atributo (impresión estereofónica, impresión espacial, timbre y calidad general) mediante la plataforma Google

Forms. El sujeto debía realizar los mismos pasos para cada una de las canciones a evaluar. Para esta prueba se usaron las mismas 3 piezas musicales de la prueba de preferencia previamente realizada.

En la Figura 34 se observa el diagrama de conexionado para la prueba de calidad sonora. Los audífonos Audio Technica M40X se conectaron a la salida de audio de un celular Huawei P10 Lite. El sujeto leía y diligenciaba el formato establecido en Google Forms para guardar sus resultados. Cada persona ajustó el nivel de ganancia de los audífonos según su comodidad.



Figura 34. Diagrama de conexionado de la prueba de calidad sonora.

Las pruebas de calidad sonora se realizaron en los estudios de grabación de la Universidad de San Buenaventura Sede Bogotá, los cuales presentan bajos niveles de ruido de fondo y un ambiente controlado, permitiendo a los participantes de la prueba mayor concentración. El banco de jurados que se tomó para la prueba fue de un total de 25 personas, acorde a lo estipulado por la recomendación ITU-R BS 1284-1 (International Telecommunication Union, 2003), los cuales fueron estudiantes cursando Octavo Semestre

o superiores de Ingeniería de Sonido. En promedio cada prueba tenía una duración de 10 a 20 minutos.

A diferencia de las primeras pruebas aplicadas con los 4 métodos de simulación de crosstalk, para esta segunda prueba se hicieron evaluaciones subjetivas de 4 atributos sonoros para cada canción reproducida por cada uno de los métodos implementados en la aplicación móvil (original, Manor y Gilchrest). Los atributos evaluados se tomaron de la recomendación ITU-R BS 1284-1 (International Telecommunication Union, 2003), los cuales también fueron tenidos en cuenta en las evaluaciones de calidad sonora realizadas por Bolaños (2016) para métodos de simulación de crosstalk en audífonos. Los atributos evaluados fueron:

- **Impresión estereofónica:** La imagen sonora tiene la distribución direccional correcta y apropiada de las fuentes sonoras.
- **Impresión espacial:** El evento sonoro tiene lugar en un entorno espacial apropiado.
- **Timbre:** Se genera una representación exacta de las diferentes características de la(s) fuente(s) sonora(s).
- **Calidad general:** La integridad de la imagen sonora y la interacción entre los diversos parámetros es adecuada.

4.4 Evaluaciones calidad sonora adicionales a oyentes expertos

Fueron realizadas pruebas complementarias a fin de obtener resultados más precisos. Para esto se realizó la misma evaluación de calidad sonora con 10 oyentes expertos que tuvieran 10 o más años de experiencia en las áreas de mezcla, producción, mastering, captura y demás. Las personas seleccionadas para esta prueba fueron profesores de Ingeniería de

Sonido de la Universidad San Buenaventura sede Bogotá enfocados a los temas anteriormente mencionados, e ingenieros de grabación y mezcla como lo indica la

Figura 35 y la Tabla 4.



Figura 35. Oyente experto realizando prueba de calidad sonora.

Tabla 4. Panel adicional de oyentes expertos.

Sujeto	Edad	Genero	Ocupación	Años de experiencia
Sujeto 1	55	Hombre	Profesor	30
Sujeto 2	38	Mujer	Profesor	15
Sujeto 3	33	Hombre	Profesor	10
Sujeto 4	38	Hombre	Profesor	15
Sujeto 5	38	Hombre	Profesor	20
Sujeto 6	37	Hombre	Profesor	20
Sujeto 7	28	Mujer	Productora Musical	10
Sujeto 8	42	Hombre	Ingeniero de Sonido	20
Sujeto 9	33	Hombre	Ingeniero de Sonido	13
Sujeto 10	27	Hombre	Ingeniero de Sonido	12

A partir de esto se obtuvieron los datos mostrados en los Resultados de pruebas de calidad sonora de la aplicación móvil realizadas a oyentes expertos.

5 Discusión y Análisis de Resultados

5.1 Resultados de las pruebas de preferencia de los métodos de simulación de crosstalk

A continuación, se presentan los resultados de las pruebas de preferencia aplicadas para cada canción seleccionada procesadas en los 4 diferentes métodos, y se mostrará los resultados estadísticos y escalados según su distribución Z.

5.1.1 Resultados de pruebas de preferencia para Symphony No. 5

De acuerdo con la Figura 29, y las tablas contenidas en el Anexo A. Se puede observar que para este tema el método de menor preferencia fue el método Kirkeby, ya que corresponde al menor valor de la escala proporcionada en la distribución Z; seguido se encuentra el método Bolaños de con (0.03) en el escalamiento, lo que indica que fue el segundo método menos seleccionado; posteriormente se encuentra el método de Manor (0.44), seguido de la canción original (0.55). Finalmente, el método de Gilchrest, con (0.57) fue el que obtuvo la mayor preferencia entre los oyentes para el tema Symphony No. 5.

5.1.2 Resultados de pruebas de preferencia para Money

En la Figura 36 se expone el escalamiento de los resultados de preferencia para el tema musical Money. De igual manera, en el Anexo A se denotan las tablas de datos correspondientes a la aplicación de la ley de Thurstone desde la Tabla 32 hasta la Tabla 34. Se puede observar que el método con menor preferencia fue el de Bolaños; posterior a este se encuentra el método de Kirkeby, el cual obtuvo un valor de (0.13) según la escala; en tercer lugar, el método de mayor frecuencia de selección, se encuentra el método de Manor con (0.86) en su puntaje de escala. Esto indica que los dos métodos de procesamiento de

mayor preferencia después de la canción original fueron los modelos presentados por Gilchrest y Manor respectivamente, y a pesar de que el método de Gilchrest obtuvo una preferencia por debajo de la canción original, esta diferencia fue de un valor mínimo (0.04) en comparación con las diferencias de los demás métodos respecto a el tema original, y esto ocurrió específicamente para este tema musical.

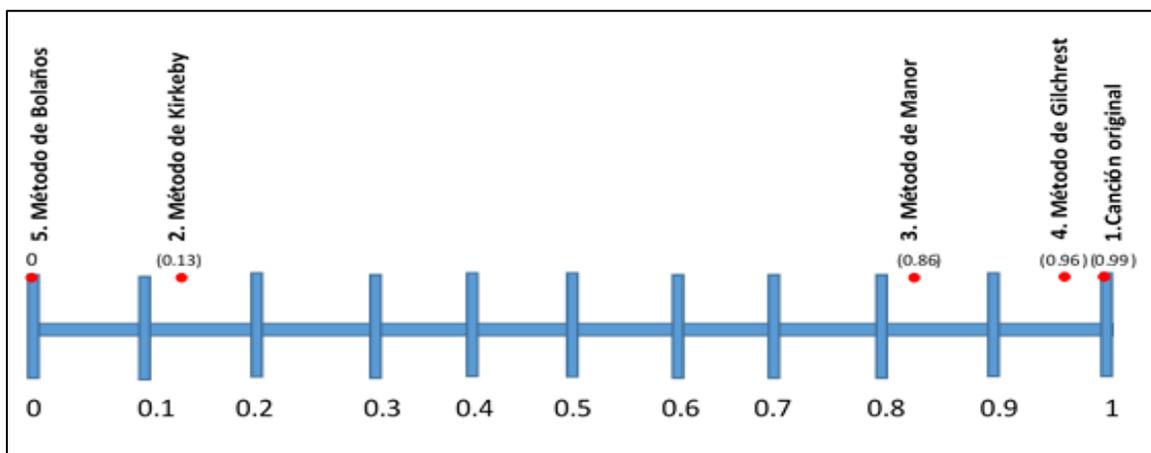


Figura 36. Resultado del escalamiento para el tema Money.

5.1.3 Resultados de pruebas de preferencia para la canción “Con Calma”

Los resultados obtenidos para las pruebas de preferencia realizadas para este tema musical (Genero urbano), se muestran en el Anexo A en las tablas se datos correspondientes a la aplicación de la ley de Thurstone desde la

Tabla 29 hasta la Tabla 31, y según sus valores de proporciones en distribución Z en las Tablas de distribución normal, mostrados en escaladamente en la Figura 37.

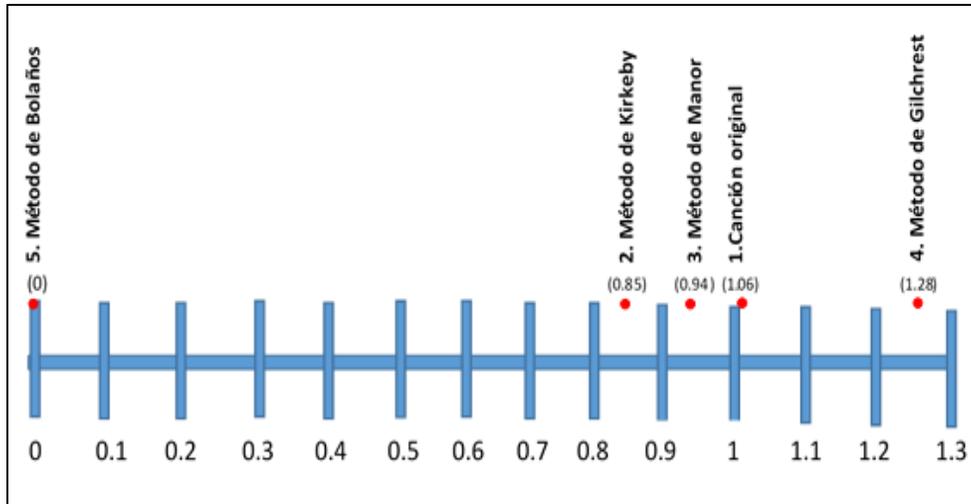


Figura 37. Resultado del escalamiento para el tema Con Calma.

Se indica que el método con menos frecuencia de selección de preferencia, al igual que en el caso anterior, fue el Método de Bolaños, con un valor de proporción normalizada promedio de 0 en este caso, siendo el valor de escala de referencia. Posteriormente, está ubicado el segundo método con menos selección de preferencia, el método de Kirkeby con un valor de escala de (0.85), lo cual indica que la diferencia respecto al método de Bolaños es significativa en este tema musical en específico. Luego se encuentra el método de Manor con una diferencia de (0.09) respecto a su anterior. La segunda opción más seleccionada según la evaluación de preferencia realizada a los oyentes fue la canción original. Por último, se observa que el método de Gilchrest, fue la opción que tuvo una mayor frecuencia de selección por parte de los oyentes, por encima de la señal original, con una diferencia de 0.22 según el escalamiento.

Para esta canción en particular, se observa que la diferencia en valor de preferencia según el escalamiento normalizado Z promedio, entre el método Bolaños y su posterior (Método de Kirkeby) es de 0.85 que corresponde a la mayor diferencia entre métodos según

los demás escalamientos de las demás canciones, lo que indica que para este género musical el método de Bolaños no es agradable subjetivamente para la escucha a través de audífonos.

5.1.4 Análisis del consolidado de datos para las pruebas de preferencia

En la Tabla 37 (Ver Resultados y aplicación de la ley Thurstone para prueba de preferencia en ANEXO A) se muestran los resultados de la matriz normalizadas Z para el consolidado de datos de los diferentes métodos respecto a cada canción, teniendo en cuenta la frecuencia de selección de preferencia para los 25 oyentes encuestados. De igual manera se puede observar que el método de procesamiento que obtuvo menor preferencia fue el método de Bolaños, mientras que el método de Kirkeby fue el de mayor preferencia, superando al audio original.

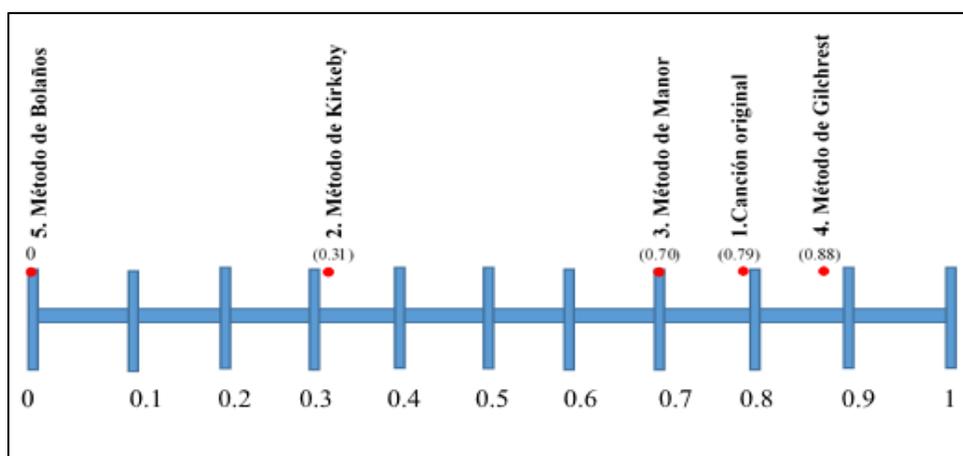


Figura 38. Escala de preferencia obtenida de los datos consolidados para las tres canciones evaluadas.

En el escalamiento de preferencia para los temas Symphony No.5 (género clásico) y Con Calma (género urbano) se puede observar que el método de procesamiento de mayor preferencia fue el método de Gilchrest. Por otro lado, en la Figura 36 se observa que para el

tema musical Money (género rock), la versión original fue la de mayor preferencia. Así mismo, en cada una de estas escalas, el segundo método de procesamiento preferido por los oyentes (sin tener en cuenta el audio original) fue el método de Manor.

Para las canciones Money (Figura 36) y Con Calma (Figura 37) el método de menor preferencia fue el método de Bolaños, mientras que para la canción Symphony No. 5 (Figura 29) el método menos preferido fue el de Kirkeby. se evidencia que el género musical puede afectar la preferencia del método de procesamiento por parte de los oyentes de prueba.

5.2 Resultados de las pruebas de calidad sonora de la aplicación móvil

Teniendo en cuenta que para esta etapa solo se implementaron los dos métodos de mayor preferencia según el Análisis del consolidado de datos para las pruebas de preferencia, se hizo uso del software 'IBM SPSS Statistics', en el cual se analizaron los resultados de las pruebas subjetivas de la aplicación móvil. Se emplearon los estadísticos descriptivos para observar las medias obtenidas de cada atributo en función del tema y el método evaluados. Posteriormente, se aplicó la prueba H de Kruskal-Wallis para comprobar si las diferencias obtenidas en el análisis son estadísticamente significativas.

5.2.1 Resultados de pruebas de calidad sonora de la aplicación móvil realizadas a estudiantes

Evaluando los años de experiencia en el campo del audio de los 25 estudiantes entrevistados, se obtuvo una media de 4.96 años. A continuación, se presentan los resultados de las evaluaciones de calidad sonora de la aplicación 'Xtereo Music Player' para 25 estudiantes de Ingeniería de Sonido.

Se analizó la distribución de las calificaciones para cada atributo en función del método de procesamiento y el tema evaluado mediante la prueba de Shapiro-Wilk, obteniendo los resultados de la Tabla 5. Teniendo en cuenta los valores de significancia, los cuales son menores a 0.05 en todos los casos, se concluye que los datos obtenidos para la evaluación de calidad sonora de la aplicación móvil realizada a estudiantes no obedecen a una distribución normal, de manera que es factible la aplicación de la prueba H de Kruskal-Wallis para su análisis estadístico.

Tabla 5. Prueba de Shapiro-Wilk para los datos obtenidos con estudiantes.

Atributo	Normalidad por Tema				Normalidad por Método			
	Tema	Shapiro-Wilk			Método	Shapiro-Wilk		
		Estadístico	gl	Sig.		Estadístico	gl	Sig.
Estéreo	Symphony No. 5	0,791	75	0,000	Original	0,794	75	0,000
	Money	0,780	75	0,000	Manor	0,795	75	0,000
	Con Calma	0,837	75	0,000	Gilcrest	0,826	75	0,000
Espacialidad	Symphony No. 5	0,832	75	0,000	Original	0,802	75	0,000
	Money	0,818	75	0,000	Manor	0,822	75	0,000
	Con Calma	0,851	75	0,000	Gilcrest	0,824	75	0,000
Timbre	Symphony No. 5	0,811	75	0,000	Original	0,763	75	0,000
	Money	0,792	75	0,000	Manor	0,773	75	0,000
	Con Calma	0,845	75	0,000	Gilcrest	0,846	75	0,000
Calidad	Symphony No. 5	0,825	75	0,000	Original	0,771	75	0,000
	Money	0,749	75	0,000	Manor	0,797	75	0,000
	Con Calma	0,796	75	0,000	Gilcrest	0,862	75	0,000

5.2.1.1 Resultados de pruebas de calidad sonora de la aplicación móvil realizadas a estudiantes con el tema musical Symphony No. 5

Los estadísticos descriptivos para este tema se presentan en la Tabla 6. En negrilla se resalta la media de mayor valor para cada atributo sonoro evaluado. Al comparar las medias de cada atributo, la impresiones estereofónica y espacial obtuvieron un mayor valor en método de Gilcrest, el timbre tuvo un mayor valor en la versión original y la calidad general

fue mayor en el método de Manor. Sin embargo, las desviaciones estándar fueron mayores que las diferencias entre los valores de las medias, lo que supone que las diferencias de las medias no fueron significativas.

Tabla 6. Estadísticos descriptivos de las evaluaciones de calidad sonora con estudiantes para el tema Symphony No. 5.

Estadísticos descriptivos con estudiantes para el tema Symphony No. 5					
Método de reproducción	Atributo	Mínimo	Máximo	Media	Desviación
Original	Estéreo	3,00	5,00	4,16	0,6245
	Espacialidad	3,00	5,00	4,04	0,7348
	Timbre	3,00	5,00	4,24	0,6633
	Calidad	3,00	5,00	3,84	0,7461
Manor	Estéreo	3,00	5,00	4,20	0,8165
	Espacialidad	2,00	5,00	4,08	0,7594
	Timbre	3,00	5,00	4,08	0,6403
	Calidad	3,00	5,00	3,96	0,7895
Gilchrest	Estéreo	3,00	5,00	4,32	0,7483
	Espacialidad	2,00	5,00	4,16	0,9866
	Timbre	2,00	5,00	3,92	0,7594
	Calidad	2,00	5,00	3,88	0,9713

Con el fin de verificar si existe diferencia estadísticamente significativa entre los métodos de procesamiento según el banco de jurados, se realizó un análisis mediante la prueba H de Kruskal-Wallis, evaluando cada atributo en función del método de procesamiento. Los resultados para el tema Symphony No. 5 representados en la Tabla 7 evidencian que no hubo diferencias significativas entre cada método, ya que ningún indicador de significancia asintótica fue menor a 0.05. La impresión estereofónica y la impresión espacial obtuvieron un mayor rango promedio en el método de Gilchrest, seguido por el método de Manor y el audio original. El timbre tuvo un mayor rango promedio en el audio original, seguido por los métodos de Manor y de Gilchrest respectivamente. La calidad

general tuvo un mayor rango promedio en el método de Manor, seguido por el método de Gilchrest dejando en último lugar al audio original.

Tabla 7. Estadísticos de Kruskal-Wallis de las evaluaciones de calidad sonora con estudiantes para el tema Symphony No. 5z.

Estadísticos de Kruskal-Wallis con estudiantes para el tema Symphony No. 5				
Atributo	Método	Rango promedio	H de Kruskal-Wallis	Significancia
Estéreo	Original	35,48	0,8726	0,6464
	Manor	37,74		
	Gilchrest	40,78		
Espacialidad	Original	35,96	0,6721	0,7146
	Manor	37,44		
	Gilchrest	40,60		
Timbre	Original	42,42	2,3642	0,3066
	Manor	37,62		
	Gilchrest	33,96		
Calidad	Original	36,66	0,2616	0,8774
	Manor	39,60		
	Gilchrest	37,74		

5.2.1.2 Resultados de pruebas de calidad sonora de la aplicación móvil realizadas a estudiantes con el tema musical Money.

En la Tabla 8 se observan los estadísticos descriptivos obtenidos para el tema Money. El audio original tuvo la mayor media de impresión estereofónica y el método de Manor obtuvo los mayores puntajes de espacialidad (junto con el método de Gilchrest), timbre y calidad general. Sin embargo, las desviaciones estándar fueron mayores a las diferencias entre las medias.

Tabla 8. Estadísticos descriptivos de las evaluaciones de calidad sonora con estudiantes para el tema Money.

Estadísticos descriptivos con estudiantes para el tema Money					
Método de reproducción	Atributo	Mínimo	Máximo	Media	Desviación
Original	Estéreo	3,00	5,00	4,40	0,57735
	Espacialidad	3,00	5,00	4,08	0,70238
	Timbre	2,00	5,00	4,28	0,67823
	Calidad	1,00	5,00	4,00	0,81650
Manor	Estéreo	3,00	5,00	4,24	0,66332
	Espacialidad	3,00	5,00	4,20	0,70711
	Timbre	3,00	5,00	4,40	0,57735
	Calidad	3,00	5,00	4,28	0,61373
Gilchrest	Estéreo	3,00	5,00	4,16	0,74610
	Espacialidad	2,00	5,00	4,20	0,86603
	Timbre	3,00	5,00	3,96	0,73485
	Calidad	3,00	5,00	4,08	0,64031

La Tabla 9 presenta los estadísticos de Kruskal-Wallis para el tema Money. Aquí, al igual que en el tema Symphony No. 5 tampoco hubo diferencias significativas entre cada método. La impresión estereofónica tuvo un mayor rango promedio en el audio original, seguido por los métodos de Manor y de Gilchrest respectivamente. La impresión espacial tuvo un mayor rango promedio en el método de Gilchrest, seguido por el método de Gilchrest y el audio original. La impresión tímbrica tuvo un mayor rango promedio en el método de Manor, seguido por el audio original y el método de Gilchrest, pero con la diferencia que su significancia asintota es de 0.064, siendo un valor muy cercano de aprobación de significancia (0.05). Por último, la calidad general tuvo un mayor rango promedio para el método de Manor, seguido por el método de Gilchrest y el audio original.

Tabla 9. Estadísticos de Kruskal-Wallis de las evaluaciones de calidad sonora con estudiantes para el tema Money.

Estadísticos de Kruskal-Wallis con estudiantes para el tema Money				
Atributo	Método	Rango promedio	H de Kruskal-Wallis	Significancia
Estéreo	Original	41,60	1,3555	0,5078
	Manor	37,12		
	Gilcrest	35,28		
Espacialidad	Original	35,34	0,6990	0,7050
	Manor	38,72		
	Gilcrest	39,94		
Timbre	Original	40,24	5,4824	0,0645
	Manor	43,06		
	Gilcrest	30,70		
Calidad	Original	35,68	1,8827	0,3901
	Manor	42,22		
	Gilcrest	36,10		

5.2.1.3 Resultados de pruebas de calidad sonora de la aplicación móvil realizadas a estudiantes con el tema Con Calma.

La Tabla 10 muestra los estadísticos descriptivos para el tema Con Calma. El método de Manor obtuvo los valores promedio más altos en los atributos de impresión estereofónica, impresión espacial y calidad general. El audio original obtuvo el valor más alto para el atributo de timbre.

Tabla 10. Estadísticos descriptivos de las evaluaciones de calidad sonora con estudiantes para el tema Con Calma.

Estadísticos descriptivos con estudiantes para el tema Con Calma					
Método de reproducción	Atributo	Mínimo	Máximo	Media	Desviación
Original	Estéreo	3,00	5,00	3,88	0,72572
	Espacialidad	3,00	5,00	3,76	0,59722
	Timbre	2,00	5,00	4,12	0,72572
	Calidad	3,00	5,00	3,92	0,49329
Manor	Estéreo	3,00	5,00	4,16	0,68799
	Espacialidad	3,00	5,00	3,88	0,66583
	Timbre	3,00	5,00	4,04	0,61101
	Calidad	3,00	5,00	4,12	0,60000
Gilchrest	Estéreo	2,00	5,00	3,68	0,85245
	Espacialidad	2,00	5,00	3,76	1,01160
	Timbre	2,00	5,00	3,44	0,76811
	Calidad	2,00	5,00	3,60	0,86603

La Tabla 11 presenta los estadísticos de Kruskal-Wallis para el tema Con Calma, en el cual se obtuvieron diferencias significativas entre los métodos de reproducción para los atributos de timbre y calidad general. La impresión estereofónica tuvo un mayor rango promedio en el método de Manor, seguido por el audio original y el método de Gilchrest. La impresión espacial tuvo un mayor rango promedio en el método de Manor, seguido por el método de Gilchrest y el audio original. La impresión tímbrica tuvo un mayor rango promedio en el audio original, seguido por los métodos de Manor y de Gilchrest respectivamente. La calidad general tuvo un mayor rango promedio en el método de Manor, seguido por el audio original y el método de Gilchrest.

Tabla 11. Estadísticos de Kruskal-Wallis de las evaluaciones de calidad sonora con estudiantes para el tema Con Calma.

Estadísticos de Kruskal-Wallis con estudiantes para el tema Con Calma				
Atributo	Método	Rango promedio	H de Kruskal-Wallis	Significancia
Estéreo	Original	37,14	4,7807	0,0916
	Manor	44,68		
	Gilcrest	32,18		
Espacialidad	Original	36,56	0,3216	0,8515
	Manor	39,74		
	Gilcrest	37,70		
Timbre	Original	45,34	13,3901	0,0012
	Manor	42,56		
	Gilcrest	26,10		
Calidad	Original	38,44	6,0325	0,0490
	Manor	44,30		
	Gilcrest	31,26		

5.2.1.4 Análisis del consolidado de las pruebas de calidad sonora por la aplicación móvil realizadas a estudiantes

En la Tabla 12 se exponen los datos estadísticos descriptivos de las evaluaciones subjetivas de calidad sonora de los 3 temas musicales evaluados. Los resultados de las medias para los atributos evaluados por los estudiantes indicaron que el método de Manor presenta mayor calificación promedio en los atributos de impresión estereofónica, espacialidad y calidad general, mientras que el audio original fue el que obtuvo mayor valor medio en el atributo de timbre. Sin embargo, por tener valores de desviación estándar mayores a las diferencias de las medias entre los métodos, no es posible con estos descriptivos determinar si el método de Manor mejora de manera significativa los atributos sonoros evaluados.

Tabla 12. Estadísticos descriptivos de las evaluaciones de calidad sonora con estudiantes para los tres temas evaluados.

Estadísticos descriptivos con estudiantes para los tres temas evaluados					
Método de reproducción	Atributo	Mínimo	Máximo	Media	Desviación
Original	Estéreo	3,00	5,00	4,15	0,67169
	Espacialidad	3,00	5,00	3,96	0,68655
	Timbre	2,00	5,00	4,21	0,68366
	Calidad	1,00	5,00	3,92	0,69282
Manor	Estéreo	3,00	5,00	4,20	0,71660
	Espacialidad	2,00	5,00	4,05	0,71458
	Timbre	3,00	5,00	4,17	0,62327
	Calidad	3,00	5,00	4,12	0,67704
Gilchrest	Estéreo	2,00	5,00	4,05	0,82024
	Espacialidad	2,00	5,00	4,04	0,96479
	Timbre	2,00	5,00	3,77	0,78108
	Calidad	2,00	5,00	3,85	0,84938

Para la totalidad de los datos obtenidos, presentados en la Tabla 13, solo hubo diferencias significativas en el atributo de timbre, teniendo un valor de significancia asintótica de 0.0002. En cuanto a los rangos promedio, tanto la impresión estereofónica como la calidad general tuvieron mayor valor en el método Manor; en impresión espacial el método con mayor valor fue el método de Gilchrest; el audio original tuvo un rango promedio mayor para el atributo de timbre.

Tabla 13. Estadísticos de Kruskal-Wallis de las evaluaciones de calidad sonora con estudiantes para los tres temas evaluados.

Estadísticos de Kruskal-Wallis con estudiantes para los tres temas evaluados				
Atributo	Método	Rango promedio	H de Kruskal-Wallis	Significancia
Estéreo	Original	113,19	1,1444	0,5643
	Manor	118,17		
	Gilchrest	107,64		
Espacialidad	Original	106,67	1,2956	0,5232
	Manor	114,82		
	Gilchrest	117,51		
Timbre	Original	126,93	16,8912	0,0002
	Manor	121,80		
	Gilchrest	90,27		
Calidad	Original	109,37	4,4423	0,1085
	Manor	124,52		
	Gilchrest	105,11		

Teniendo en cuenta los atributos para los cuales se obtuvo una diferencia significativa entre los métodos de reproducción (impresión tímbrica y calidad general para el tema ‘Con Calma’, e impresión tímbrica en la totalidad de los datos), el método de Gilchrest fue el que tuvo menor calificación por parte de los evaluadores. Esto contrasta con las observaciones de las pruebas de preferencia realizadas en la primera etapa del proyecto, cuyos resultados mostraron que el método de Gilchrest fue el que obtuvo mayor preferencia por encima de los demás. En la prueba de preferencia los audios duraban 15 segundos y se mostraban por pares, permitiendo solo una repetición. Por otro lado, la prueba de calidad sonora, el estudiante que evaluaba la aplicación podía escuchar hasta 2 minutos de la canción, repitiendo las veces que deseara y escuchando en tiempo real el cambio entre cada uno de los 3 métodos de reproducción implementados. A partir de esto, es posible evidenciar que la diferencia de los resultados se debe a que, es más simple que los estudiantes denoten una calificación mediante una prueba de preferencia (Método de pares comparados) a que escalen aspectos específicos de calidad sonora. Según Otto (1999) y los parámetros de calidad sonora descritos por Koehl

y Paquier (2013), las pruebas de escala requieren un banco de jurados con mayor experiencia en temas afines.

5.2.2 Resultados de pruebas de calidad sonora de la aplicación móvil realizadas a oyentes expertos

Considerando los resultados obtenidos con el banco de jurado de estudiantes (Sección 5.2.1), se decidió consultar un nuevo banco de jurados conformado por 10 expertos para realizar pruebas complementarias, dicho banco fue constituido por profesores de la carrera de ingeniería de Sonido en la Universidad San Buenaventura sede Bogotá y por ingenieros de audio dedicados al área de captura, mezcla, mastering, etc. con 10 años o más de experiencia. Esto se llevó a cabo con el fin de obtener resultados significativos. El promedio de años de experiencia para este panel fue de 16.5 años. En la Tabla 14 se encuentra la prueba de normalidad de los datos obtenidos. Se halló que los datos obtenidos con oyentes expertos tampoco obedecen a una distribución normal.

Tabla 14. Prueba de Shapiro-Wilk para los datos obtenidos con oyentes expertos.

Atributo	Normalidad por Tema				Normalidad por Método			
	Tema	Shapiro-Wilk			Método	Shapiro-Wilk		
		Estadístico	gl	Sig.		Estadístico	gl	Sig.
Estereo	Symphony No. 5	0,878	30	0,003	Original	0,772	30	0,000
	Money	0,791	30	0,000	Manor	0,792	30	0,000
	Con Calma	0,851	30	0,001	Gilcrest	0,845	30	0,000
Espacialidad	Symphony No. 5	0,859	30	0,001	Original	0,707	30	0,000
	Money	0,791	30	0,000	Manor	0,720	30	0,000
	Con Calma	0,851	30	0,001	Gilcrest	0,841	30	0,000
Timbre	Symphony No. 5	0,901	30	0,009	Original	0,800	30	0,000
	Money	0,902	30	0,010	Manor	0,882	30	0,003
	Con Calma	0,864	30	0,001	Gilcrest	0,903	30	0,010
Calidad	Symphony No. 5	0,824	30	0,000	Original	0,820	30	0,000
	Money	0,859	30	0,001	Manor	0,774	30	0,000
	Con Calma	0,899	30	0,008	Gilcrest	0,902	30	0,009

5.2.2.1 Resultados de pruebas de calidad sonora realizadas a oyentes expertos para Symphony No. 5.

La Tabla 15 presenta los estadísticos descriptivos para el tema Symphony No. 5 con oyentes expertos. Aquí los valores de media en todos los atributos de calificación son superiores en el audio original respecto a los demás métodos. En el método de Gilchrest el atributo de espacialidad obtuvo la misma calificación que el audio original.

Tabla 15. Estadísticos descriptivos de las evaluaciones de calidad sonora con oyentes expertos para el tema Symphony No 5.

Estadísticos descriptivos con expertos para el tema Symphony No 5					
Método de reproducción	Atributo	Mínimo	Máximo	Media	Desviación
Original	Estéreo	3,00	5,00	4,10	0,7379
	Espacialidad	3,00	5,00	4,00	0,8165
	Timbre	3,00	5,00	4,10	0,7379
	Calidad	3,00	5,00	4,10	0,7379
Manor	Estéreo	2,00	4,00	3,20	0,6325
	Espacialidad	3,00	4,00	3,50	0,5270
	Timbre	2,00	5,00	3,40	0,8433
	Calidad	3,00	4,00	3,60	0,5164
Gilchrest	Estéreo	2,00	5,00	3,80	1,2293
	Espacialidad	2,00	5,00	4,00	1,0541
	Timbre	1,00	5,00	3,10	1,1972
	Calidad	1,00	5,00	3,40	1,2649

La Tabla 16 contiene los estadísticos de Kruskal-Wallis para este tema. Para ninguno de los atributos evaluados se obtuvieron diferencias estadísticamente significativas, ya que ninguno tuvo una significancia asintótica menor a 0.05. Se evidencia que, en la impresión estereofónica, timbre y calidad general su rango promedio de mayor valor se encuentra en el audio original, seguido por el método de Gilchrest y el método de Manor. Por otro lado, la impresión espacial obtuvo mayor calificación con el método de Gilchrest, seguido por el audio original y el método de Manor.

Tabla 16. Estadísticos de Kruskal-Wallis de las evaluaciones de calidad sonora con oyentes expertos para el tema Symphony No. 5.

Estadísticos de Kruskal-Wallis con expertos para el tema Symphony No.5				
Atributo	Método	Rango promedio	H de Kruskal-Wallis	Significancia
Estéreo	Original	19,05	5,007	0,082
	Manor	10,85		
	Gilchrest	16,60		
Espacialidad	Original	17,05	2,684	0,261
	Manor	12,00		
	Gilchrest	17,45		
Timbre	Original	20,30	5,065	0,079
	Manor	13,85		
	Gilchrest	12,35		
Calidad	Original	18,85	2,745	0,253
	Manor	13,40		
	Gilchrest	14,25		

5.2.2.2 Resultados de pruebas de calidad sonora realizadas a oyentes expertos para la canción Money.

En la Tabla 17 se presentan los estadísticos descriptivos para el tema Money. Se destaca los mayores valores de medias en el audio original para todos los atributos. En relación con el método de Manor y de Gilchrest sus medias son idénticas entre cada atributo, lo cual sugiere que los panelistas no percibieron diferencias entre estos métodos.

Tabla 17. Estadísticos descriptivos de las evaluaciones de calidad sonora con oyentes expertos para el tema Money.

Estadísticos descriptivos con expertos para el tema Money					
Método de reproducción	Atributo	Mínimo	Máximo	Media	Desviación
Original	Estéreo	4,00	5,00	4,50	0,5270
	Espacialidad	4,00	5,00	4,10	0,3162
	Timbre	3,00	5,00	4,30	0,6749
	Calidad	3,00	5,00	4,40	0,6992
Manor	Estéreo	3,00	4,00	3,70	0,4830
	Espacialidad	3,00	4,00	3,70	0,4830
	Timbre	2,00	5,00	3,30	0,9487
	Calidad	3,00	5,00	3,60	0,6992
Gilchrest	Estéreo	3,00	4,00	3,70	0,4830
	Espacialidad	3,00	4,00	3,70	0,4830
	Timbre	2,00	5,00	3,30	0,9487
	Calidad	3,00	5,00	3,60	0,6992

Lo anterior se corrobora con los estadísticos de Kruskal-Wallis que se presentan en la Tabla 18. El rango promedio de mayor valor se haya en la señal original para cada atributo. En cuanto a la significancia asintótica se encuentra que hay diferencias significativas en los atributos de timbre y calidad general. En los atributos de impresión estereofónica e impresión espacial el rango de menor valor se halló en el método de Manor; caso contrario ocurre con el timbre y la calidad general donde el método Gilchrest obtuvo el rango de menor valor.

Tabla 18. Estadísticos de Kruskal-Wallis de las evaluaciones de calidad sonora con oyentes expertos para el tema Money.

Estadísticos de Kruskal-Wallis con expertos para el tema Money				
Atributo	Método	Rango promedio	Kruskal-Wallis	Significancia
Estéreo	Original	20,25	5,682	0,058
	Manor	11,65		
	Gilcrest	14,60		
Espacialidad	Original	18,10	1,696	0,428
	Manor	13,70		
	Gilcrest	14,70		
Timbre	Original	22,30	10,775	0,005
	Manor	14,10		
	Gilcrest	10,10		
Calidad	Original	20,80	5,986	0,050
	Manor	13,20		
	Gilcrest	12,50		

5.2.2.3 Resultados de pruebas de calidad sonora realizadas a oyentes expertos para Con Calma

En cuanto a las medias obtenidas en el tema Con Calma, las cuales se pueden ver en la Tabla 19, los atributos de impresión estereofónica, impresión espacial y timbre presentan el mayor valor en la señal original. Para el caso de la calidad general existe un valor idéntico de media entre la señal original y el método de Manor.

Tabla 19. Estadísticos descriptivos de las evaluaciones de calidad sonora con oyentes expertos para el tema Con Calma.

Estadísticos descriptivos con expertos para el tema Con Calma					
Método de reproducción	Atributo	Mínimo	Máximo	Media	Desviación
Original	Estéreo	3,00	5,00	4,10	0,5676
	Espacialidad	2,00	4,00	3,80	0,6325
	Timbre	3,00	5,00	4,10	0,7379
	Calidad	2,00	5,00	3,90	0,8756
Manor	Estéreo	3,00	5,00	3,70	0,6749
	Espacialidad	3,00	5,00	3,60	0,6992
	Timbre	2,00	5,00	3,90	0,8756
	Calidad	3,00	5,00	3,90	0,7379
Gilchrest	Estéreo	1,00	5,00	3,20	1,6193
	Espacialidad	1,00	5,00	3,00	1,6330
	Timbre	1,00	4,00	2,70	1,1595
	Calidad	1,00	5,00	2,80	1,1353

En la Tabla 20, donde se presentan los estadísticos de Kruskal-Wallis, se evidencia que el rango promedio de mayor valor se obtuvo con la señal original en todos los atributos. El método de Gilchrest obtuvo los valores de rango promedio más bajos. Se hallaron diferencias estadísticamente significativas en los atributos de timbre y calidad general.

Tabla 20. Estadísticos de Kruskal-Wallis de las evaluaciones de calidad sonora con oyentes expertos para el tema Con Calma.

Estadísticos de Kruskal-Wallis con expertos para el tema Con Calma				
Atributo	Método	Rango promedio	Kruskal-Wallis	Significancia
Estéreo	Original	18,50	2,046	0,359
	Manor	14,50		
	Gilchrest	13,50		
Espacialidad	Original	18,40	2,321	0,313
	Manor	15,35		
	Gilchrest	12,75		
Timbre	Original	19,35	8,819	0,012
	Manor	17,95		
	Gilchrest	9,20		
Calidad	Original	18,60	7,158	0,028
	Manor	18,20		
	Gilchrest	9,70		

5.2.2.4 Análisis del consolidado de las pruebas de calidad sonora da la aplicación móvil realizadas a oyentes expertos

La Tabla 21 contiene los estadísticos descriptivos para los tres temas de la prueba. Aquí se tuvieron valores más altos con el audio original para cada uno de los atributos evaluados. La impresión estereofónica obtuvo el mismo valor de media en los métodos de Manor y Gilchrest. Además, este último método tiene los valores más bajos para los atributos de impresión espacial, timbre y calidad general.

Tabla 21. Estadísticos descriptivos de las evaluaciones de calidad sonora con oyentes expertos para los tres temas.

Estadísticos descriptivos con expertos para los tres temas evaluados					
Método de reproducción	Atributo	Mínimo	Máximo	Media	Desviación
Original	Estéreo	3,00	5,00	4,23	0,6261
	Espacialidad	2,00	5,00	3,97	0,6149
	Timbre	3,00	5,00	4,17	0,6989
	Calidad	2,00	5,00	4,13	0,7761
Manor	Estéreo	2,00	5,00	3,53	0,6288
	Espacialidad	3,00	5,00	3,60	0,5632
	Timbre	2,00	5,00	3,53	0,8996
	Calidad	3,00	5,00	3,70	0,6513
Gilchrest	Estéreo	1,00	5,00	3,53	1,4559
	Espacialidad	1,00	5,00	3,47	1,4794
	Timbre	1,00	5,00	2,83	1,1167
	Calidad	1,00	5,00	3,13	1,2794

Los rangos promedio de mayor en los estadísticos de Kruskal-Wallis mostrados Tabla 22 se encontraron en la señal original. Para los atributos de impresión estereofónica, timbre y calidad general se hallaron diferencias estadísticamente significativas. El rango promedio de la impresión estereofónica fue idéntico en los métodos de Manor y Gilchrest. Este último método obtuvo los menores valores en los tres atributos restantes

Tabla 22. Estadísticos de Kruskal-Wallis de las evaluaciones de calidad sonora con oyentes expertos para las tres canciones.

Estadísticos de Kruskal-Wallis con expertos para los tres temas evaluados				
Atributo	Método	Rango promedio	Kruskal-Wallis	Significancia
Estéreo	Original	56,55	10,296	0,006
	Manor	36,30		
	Gilchrest	43,65		
Espacialidad	Original	52,52	4,051	0,132
	Manor	40,17		
	Gilchrest	43,82		
Timbre	Original	61,08	22,396	0,000
	Manor	44,80		
	Gilchrest	30,62		
Calidad	Original	57,35	12,288	0,002
	Manor	43,95		
	Gilchrest	35,20		

Por lo anterior, en los atributos con los cuales se encontraron diferencias significativas, el método de Gilchrest tiene los valores más bajos de rango promedio, lo cual significa que no mejora la calidad de los atributos con respecto al audio original o el método de Manor. Al igual que con las evaluaciones de calidad sonora de la aplicación realizadas a estudiantes, se haya una contradicción al comparar los resultados con lo obtenido en las pruebas de preferencia de la primera etapa del proyecto, en las cuales se encontró que el método de Gilchrest mejoraba la calidad con respecto a los demás. Debido a que no se realizaron pruebas de preferencia en un panel de oyentes expertos, no es posible analizar apropiadamente las diferencias entre el tipo de prueba aplicada.

5.2.3 Comparación de pruebas de calidad sonora entre estudiantes y oyentes expertos.

A partir de los datos obtenidos de estadísticos descriptivos en las secciones 755.2.1 y 5.2.2, que corresponden a las pruebas de calidad sonora resultantes para estudiantes y oyentes

expertos. Se realizaron las diferencias entre medias y desviación estándar, estos datos se presentan en la Tabla 23.

Tabla 23. Comparación de medias y desviación entre estudiantes y oyentes expertos

Método de reproducción	Atributo	Diferencia entre medias	Diferencia entre desviaciones
Original	Estéreo	0,08	0,046
	Espacialidad	0,01	0,072
	Timbre	0,04	0,015
	Calidad	0,21	0,083
Manor	Estéreo	0,67	0,088
	Espacialidad	0,45	0,151
	Timbre	0,64	0,276
	Calidad	0,42	0,026
Gilchrest	Estéreo	0,52	0,636
	Espacialidad	0,57	0,515
	Timbre	0,94	0,336
	Calidad	0,72	0,430

Los valores de media con menor diferencia para todos los atributos corresponden a el audio original. De igual manera las diferencia entre desviaciones para los atributos de impresión estereofónica, espacialidad y timbre, obtuvieron su valor más bajo en el audio original. Para el atributo de calidad general del audio la desviación estándar de menor valor se obtuvo en el método de Manor y la diferencia entre medias fue de 0,42.

Respecto a los estadísticos de Kruskal-Wallis, se realizó la comparación de datos presentada en la Tabla 24.

Tabla 24. Comparación de estadísticos de Kruskal-Wallis entre datos consolidados de estudiantes y oyentes expertos

Comparación de estadísticos de Kruskal-Wallis entre datos consolidados de estudiantes y oyentes expertos					
Atributo	Método	Rango promedio expertos	Rango promedio Estudiantes	Significancia en Expertos	Significancia en Estudiantes
Estéreo	Original	56,55	113,19	0,006	0,5643
	Manor	36,3	118,17		
	Gilchrest	43,65	107,64		
Espacialidad	Original	52,52	106,67	0,132	0,5232
	Manor	40,17	114,82		
	Gilchrest	43,82	117,51		
Timbre	Original	61,08	126,93	0	0,0002
	Manor	44,8	121,8		
	Gilchrest	30,62	90,27		
Calidad	Original	57,35	109,37	0,002	0,1085
	Manor	43,95	124,52		
	Gilchrest	35,2	105,11		

Se observa que para el atributo de impresión estereofónica el dato significativo corresponde a los valores recolectados de oyentes expertos, Para este atributo el audio original obtuvo mayor rango promedio (56,55). Observando el atributo de espacialidad, ninguno de los dos bancos de jurados indicó resultados significativos. En este atributo el audio original obtuvo el mayor rango promedio con oyentes expertos, mientras que, para estudiantes, el método de Gilchrest obtuvo el valor. Por otro lado, en el atributo de timbre se presentó que en ambos bancos de jurados los resultados fueron significativos, y concuerdan en que el audio original fue el de mayor calificación. En cuanto a la calidad general del audio, se presenta que los resultados significativos fueron los obtenidos por el banco de oyentes expertos, en los cuales el audio original fue el de mayor rango promedio. Por parte del jurado de estudiantes, el método Manor obtuvo el mayor valor de rango promedio para este atributo, aunque no hubo diferencia significativa entre los métodos.

6 Conclusiones y recomendaciones

6.1 Conclusiones

La aplicación móvil desarrollada en Unity permitió la lectura, procesamiento y reproducción de archivos de audio en formato WAV presentes en la memoria interna del dispositivo. Fue posible implementar en la aplicación los dos métodos de simulación de crosstalk de mayor preferencia en la primera etapa del proyecto.

Observando los estadísticos descriptivos del consolidado de datos obtenidos con estudiantes, la calidad general del audio procesado mediante el método de Manor fue mayor que en los demás (Tabla 12). Sin embargo, estas diferencias no fueron estadísticamente significativas. Para oyentes de mayor experiencia, el audio original fue el de mayor calidad general en cada uno de los atributos evaluados. Los resultados obtenidos en las evaluaciones de calidad sonora de la aplicación desarrollada indican que los métodos de simulación de crosstalk implementados no mejoraron la calidad sonora del audio original de una manera significativa.

Las pruebas de preferencia por pares comparados realizadas a estudiantes de Ingeniería de Sonido en la primera etapa del proyecto (Figura 38) mostraron que el método de Gilchrest logró mayor preferencia que el audio original (0.88 en la escala de preferencia), de manera que dicho método mejoró la calidad sonora. El método de Manor fue el segundo mejor método de simulación de crosstalk (0.70 en la escala), sin mejorar la calidad del audio original (0.79 en la escala). Los métodos de Kirkeby y Bolaños fueron los de menor preferencia (0.31 y 0 respectivamente).

Para las evaluaciones de calidad sonora mediante pares comparados, se obtuvo una mayor preferencia con el método de Gilchrest, seguido por el audio original y los métodos de Manor, Kirkeby y Bolaños respectivamente. Con esto se concluye que el método de Gilchrest logró mejorar la calidad sonora del audio original. Sin embargo, en uno de los temas (Money) el audio original fue el de mayor preferencia (0.99 en la escala de preferencia), seguido del método de Gilchrest (0.96 en la escala).

Realizando evaluaciones de calidad sonora de pistas de audio reproducidas mediante la aplicación desarrollada, con un panel de jurados con experiencia promedio de 4.9 años (Tabla 13), se obtuvieron diferencias estadísticamente significativas entre los tres métodos de reproducción implementados para el atributo de timbre ($s = 0.002$), siendo el audio original el de mayor valor de rango promedio y, por ende, mejor calificación. Al realizar la misma prueba con un panel de mayor experiencia, esta vez con un promedio de 16.5 años de experiencia (Tabla 14), se obtuvieron diferencias estadísticamente significativas entre los métodos de reproducción implementados para los atributos de impresión estereofónica ($s = 0.006$), timbre ($s = 0$) y calidad general ($s = 0.002$), prevaleciendo en cada uno el audio original como el de mayor valoración.

El único caso en el cual el audio procesado fue mejor valorado que el audio original fue en el tema Con Calma mediante el método de Manor, según el panel de jurados compuesto por estudiantes (Tabla 11). Se obtuvieron diferencias estadísticamente significativas entre los tres métodos de reproducción ($s = 0.049$), siendo Nearfield Crosstalk el de mayor rango promedio según la prueba H de Kruskal-Wallis.

Para ambos bancos de jurados se obtuvieron diferencias estadísticamente significativas entre los métodos de procesamiento implementados al evaluar el atributo de

timbre (con estudiantes, $s = 0.0002$; con expertos, $s = 0$), concluyendo en ambos análisis que la simulación de crosstalk en los métodos implementados afecta de manera negativa este atributo. Esto tiene relación con lo mencionado por Kirkeby (2002), quien afirma que el contenido en frecuencia que añaden los métodos de simulación de crosstalk implementados es lo que genera su poca preferencia. Además, se asemeja a los resultados obtenidos por Lorho (2002), en los cuales ningún método de simulación de crosstalk evaluado logró mejorar la calidad del audio original.

Según los estudiantes encuestados, los métodos de simulación de crosstalk implementados mejoran los atributos de impresión estereofónica (Manor) e impresión espacial (Gilchrest) aunque no de manera estadísticamente significativa (Tabla 13). Analizando el consolidado de datos, el método de Manor obtuvo el mayor valor de media para la impresión estereofónica, mientras que el de Gilchrest logró lo mismo para el atributo de impresión espacial.

La calidad sonora de las pistas de audio procesadas por los métodos de simulación de crosstalk implementados varía en función de su género musical. Con las pruebas de preferencia por pares comparados, el método de Gilchrest logró mejorar la calidad sonora de los temas Symphony No. 5 (música clásica) y Con Calma (música urbana), mientras que para el tema Money (rock) predominó el audio original. Por otro lado, el método de Manor logró mayores calificaciones que los otros métodos en cuanto a calidad general, con el banco de estudiantes evaluando el tema Con Calma. En los otros dos temas, ningún método de simulación de crosstalk mejoró la calidad general.

Se pudo observar que, realizando dos tipos diferentes de pruebas de calidad sonora (preferencia por pares comparados y evaluación de atributos sonoros mediante escala

semántica) a un banco de jurados con la misma experiencia, se obtuvieron resultados diferentes en cada una. Esto sugiere que la metodología de evaluación de calidad sonora influye en los resultados que se obtienen de esta. Es necesario determinar de manera correcta el tiempo de presentación de cada muestra de audio y la manera de evaluar su calidad sonora.

La experiencia del banco de jurados encuestado para las pruebas de calidad sonora influyó en la significancia estadística de sus resultados. Al encuestar 25 estudiantes de Ingeniería de Sonido con una experiencia promedio de 4.96 años solo se obtuvieron dos resultados significativos con el tema musical urbano y uno con el consolidado de datos, mientras que con diez oyentes expertos, con un promedio de 16.5 años de experiencia, se obtuvieron dos resultados significativos para el tema de rock, dos para el tema clásico y tres con el consolidado total de datos. El atributo evaluado por los expertos con mayor diferencias entre cada método fue el timbre. Esto pudo darse por los fenómenos de coloración mencionados por Lorho (2002) que se dan con la simulación del crosstalk. Estos cambios en la magnitud o la fase del espectro en frecuencia requieren de una mayor experiencia en el audio para poder ser detectados.

Los resultados de este proyecto contrastan con los trabajos de Manor (2015) y Gilchrest (2016), quienes afirman que sus métodos mejoran la calidad del audio en audífonos. Una posible causa de esto es el tipo de filtros que se implementan. En este proyecto se eligieron filtros de primer orden para el método de Manor y de segundo orden para el método de Gilchrest. Sin embargo, ningún autor especifica el orden ni el factor de calidad de los filtros que usaron, por lo que los resultados pueden variar en función de los filtros que se implementen.

En resumen, los métodos implementados en la aplicación móvil lograron mejorar aspectos de estereofonía y espacialidad para ciertos temas, así como mejorar la calidad de uno de ellos, a pesar de no tener diferencias estadísticamente significativas con el audio original. El tipo de prueba es un factor que puede variar los resultados obtenidos.

6.2 Recomendaciones a futuro

Es necesario profundizar en la influencia que el dispositivo de reproducción tiene sobre los resultados, ya que se realizaron pruebas mediante una sola referencia de interfaz de audio, audífonos y dispositivo móvil. Es posible que para dispositivos diferentes se obtengan otros resultados, ya que en las evaluaciones de sistemas de simulación de crosstalk no se menciona el uso de dispositivos móviles.

Se recomienda realizar más estudios acerca de los tipos de filtros usados en métodos de simulación de crosstalk. Según la bibliografía, métodos como earGoggles y Nearfield Crosstalk mejoran la calidad sonora. Sin embargo, al no haber claridad en cuanto al tipo al orden, el factor de calidad de los filtros usados o si son FIR o IIR, es difícil replicar dichos resultados.

Para simulación de crosstalk en dispositivos móviles, es necesario desarrollar un aplicativo capaz de procesar el audio proveniente de plataformas de streaming, debido al crecimiento que han tenido en los últimos años. En este proyecto se reproducían archivos de audio desde la memoria del celular. Sin embargo, las plataformas de streaming musical son las más usadas para escuchar música en estos dispositivos, de manera que, para permitir la implementación de los métodos de simulación de crosstalk, se requiere que el aplicativo procese el audio del streaming y en tiempo real.

Las diferencias significativas de los resultados obtenidos dependen de los años de experiencia del banco de jurados. Por lo tanto, se recomienda recurrir a oyentes más expertos a la hora de realizar evaluaciones de métodos de simulación de crosstalk. Al tener un panel relativamente más joven, les es más difícil discernir las diferencias entre los métodos, lo que genera que no se obtengan diferencias significativas entre los métodos de procesamiento. Dichas diferencias son percibidas en mayor grado por oyentes expertos en las áreas de producción musical y audio.

7 Referencias

- Angel Gutierrez, J. C. (2012). La correcta utilización de los promedios. *Revista Universidad Eafit*, 86.
- Arévalo, C., & Rodríguez, L. (2014). *Desarrollo de un stand-alone Software para simular el comportamiento de una cabeza binaural*. Bogotá D.C, Colombia: Universidad de San Buenaventura.
- Austerberry, D. (2005). Streaming. En *The Technology of Video and Audio Streaming*. Taylor & Francis.
- Barsainya, R., Rawat, T. K., & Mahendra, R. (2016). A new realization of wave digital filters using GIC and fractional bilinear transform. *Engineering Science and Technology, an International Journal* 19, 429-437.
- Bartlett, B. (1991). *Stereo Microphone Techniques*. Focal Press.
- Bauer, B. B. (1965). Improving Headphone Listening Comfort. *Journal of the Audio Engineering Society*, 13(4), 300-302.
- Billboard. (2019). *Hot Latin Songs*. Obtenido de <https://www.billboard.com/charts/latin-songs>
- Blauert, J. (1997). *Spatial Hearing: The Psychophysics of Human Sound Localization*. MIT Press.
- Blauert, J., & Jekosch, U. (2012). A Layer Model of Sound Quality. *Journal of the Audio Engineering Society*.
- Bolaños, J., Pilkki, V., & Mäkivirta, A. (2016). Headphone stereo enhancement using equalized binaural responses to preserve headphone sound quality. *AES Conference on Headphone Technology*. Aalborg, Dinamarca.
- Campbell, D., Jones, E., & Glavin, M. (2009). Audio quality assessment techniques—A review, and recent developments. *Signal Processing* 89, 1489-1500.
- Ceballos Sierra, J. (2007). *Microsoft C#. Lenguaje y Aplicaciones. Segunda Edición*. Madrid: RA-MA.
- Chu, E. (2008). *Discrete and Continuous Fourier Transforms: Analysis, Applications and Fast Algorithms*. CRC Press.
- Cooper, B. (2008). *Beethoven*. Oxford: Oxford University Press.
- Crawford-Emery, R., & Lee, H. (2014). The Subjective Effect of BRIR Length on Perceived Headphone Sound Externalization and Tonal Coloration. *AES 136th Convention Paper 9044*, 1-9.
- Deloitte. (2018). *Consumo móvil en Colombia*. Obtenido de <https://www2.deloitte.com/content/dam/Deloitte/co/Documents/technology-media-telecommunications/Consumo%20movil%202018.pdf>
- Departamento Administrativo Nacional de Estadística. (2016). *Encuesta de Consumo Cultural - ECC 2016*. Obtenido de https://www.dane.gov.co/files/investigaciones/eccultural/bole_ecc_2016.pdf

- Deutsch, T., Reimes, J., Fiebig, A., & Oehler, M. (2019). Relevance of stimuli and test environment in the perceptual evaluation of car audio systems. *Applied Acoustics* 156, 404-415.
- Dickins, G., McGrath, D., McKeag, A., Cartwright, R., & Reilly, A. (2009). *USA Patente nº US 7,536,021 B2*.
- Dicovskiy Riobóo, L. M. (2008). *Estadística Básica*. Esteli, Nicaragua: UNI Norte.
- Fastl, H., & Zwicker, E. (2007). Binaural Hearing. En *Psychoacoustics: Facts and Models* (pág. 293). Springer.
- Fleischmann, F., Silzle, A., & Plogsties, J. (2011). Headphone Equalization - Measurement, Design and Psychoacoustic Evaluation. En A. Heuberger, G. Elst, & R. Hanke, *Microelectronic Systems. Circuits, Systems and Applications* (págs. 301-312). Berlin: Springer.
- Flórez, M., & Gómez, J. D. (2016). *Algoritmo de Reproducción de Fuentes de Sonido Virtuales en un Espacio Virtual Tridimensional a través de un Sistema Binaural*. Bogotá D.C, Colombia: Pontificia Universidad Javeriana.
- Geig, M. (2013). *Unity Game Development in 24 Hours, Sams Teach Yourself*. Sams Publishing.
- Gelfand, S. (1981). *Hearing: An Introduction to Psychological and Physiological Acoustics*. Butterworth-Heinemann.
- Gilcrest, K. (2016). Spatial Post-Processing of Hard Panned Music for Headphone Reproduction. *AES 140th Convention*. Paris, Francia.
- Graf, U. (2000). *Applied Laplace Transforms and z-Transforms for Scientists and Engineers: A Computational Approach using a Mathematica Package*. Bienne: Birkhäuser Verlag.
- Hernández, M., & Mosquera, C. (2015). *Evaluación de la Percepción Binaural de un Espacio Tridimensional Definido Implementando un Plugin VST*. Bogotá D.C, Colombia: Universidad de San Buenaventura.
- Hernández, R., Fernández, C., & Baptista, P. (2014). *Metodología de la investigación*. McGraw-Hill Education.
- Hertz, B. (1981). 100 Years with Stereo - The Beginning. *Journal of the Audio Engineering Society*, Vol. 29, No. 5, 369-372.
- International Federation of the Phonographic Industry. (2018). *Global Music Report 2018*. Obtenido de <https://www.ifpi.org/downloads/GMR2018.pdf>
- International Telecommunication Union. (2003). *Recommendation ITU-R BS.1284-1: General methods for the subjective assessment of sound quality*.
- Kang, J. (2006). *Urban Sound Environment*. CRC Press.
- Kirkeby, O. (2002). A Balanced Stereo Widening Network for Headphones. *AES 22nd International Conference on Virtual, Synthetic and Entertainment Audio*.

- Koehl, V., & Paquier, M. (2013). A comparative study on different assessment procedures applied to loudspeaker sound quality. *Applied Acoustics* 74, 1448-1457.
- Kuo, S. M., Lee, B. H., & Tian, W. (2006). *Real-Time Digital Signal Processing*. West Sussex: John Wiley & Sons, Ltda.
- Landa, N. (2013). *Unity, Diseño y Programación de Videojuegos*. Revista Users.
- Lian, Y., & Lim, Y. C. (1993). Linear-Phase Digital Audio Tone Control Using Multiplication-Free FIR Filter. *Journal of the Audio Engineering Society*, 791-794.
- Liebetrau, J., Nowak, J., Sporer, T., Krause, M., Rekkitt, M., & Schneider, S. (2013). Paired Comparison as a Method for Measuring Emotions. *AES 135th Convention*, 1-7.
- Lorho, G., Isherwood, D., Zacharov, N., & Huopaniemi, J. (2002). Round Robin Subjective Evaluation of Stereo Enhancement Systems for Headphones. *AES 22nd International Conference on Virtual, Synthetic and Entertainment Audio*.
- Manor, E., Martens, W. L., & Cabrera, D. A. (2012). Preferred Spatial Post-Processing of Popular Stereophonic Music for Headphone Reproduction. *AES 133th Convention*. San Francisco, CA, USA.
- Manor, E., Martens, W. L., Marui, A., & Cabrera, D. (2015). Nearfield Crosstalk Increases Listener Preferences for Headphone-Reproduced Stereophonic Imagery. *Journal of the Audio Engineering Society*, 324-335.
- Merimaa, J. (2009). Modification of HRTF Filters to Reduce Timbral Effects in Binaural Synthesis. *AES 127th Convention* (págs. 1-14). New York: AES.
- Morosini, E. (Septiembre de 2012). *Escalamiento psicologico de Thrustone*. Obtenido de <http://www.slidesshare.net/EnriqueMorosini>
- Nacach, S. (2014). The Duplex Panner: Comparative Testing and Applications of an Enhanced Stereo Panning Technique for Headphone-Reproduced Commercial Music. *AES 137th Convention*. Los Angeles, CA, USA.
- Oppenheim, A. V., Willsky, A. S., & Nawab, S. H. (1998). *Señales y sistemas*. Pearson Educación.
- Ortega, I., Orjuela, A., & Cortes, D. (2008). *Diseño e Implementación de Prototipo Funcional de Reverberador por Convolución en Tiempo Real*. Bogotá D.C, Colombia: Universidad de San Buenaventura.
- Otto, N. (1999). *Guidelines for Jury Evaluations of Automotive Audio*. Michigan: The Engineering Society for Advancing Mobility Land Sea and Air Space.
- Park, D., Park, S., Kim, W., Rhiu, I., & Yun, M. H. (2019). A comparative study on subjective feeling of engine acceleration sound by automobile types. *International Journal of Industrial Ergonomics*, 1-7.

- Pedrosa, I., Juarros-Basterretxea, J., Robles-Fernández, A., Basteiro, J., & García-Cueto, E. (2015). Pruebas de bondad de ajuste en distribuciones simétricas, ¿qué estadístico utilizar? *Universitas Psychologica Vol. 14, No. 1*, 245-254.
- Reising, R. (2005). *Speak to Me: The Legacy of Pink Floyd's The Dark Side of the Moon*. Ashgate Publishing.
- Rey, D. J., & Gil, L. X. (2013). *Valoración Subjetiva de las Preferencias Acústico-Musicales del Bambuco y Pasillo Colombiano*. Bogotá D.C: Tesis de Grado, Ingeniería de Sonido, Universidad de San Buenaventura.
- Rodriguez Ojeda, L. (2007). *PROBABILIDAD Y ESTADÍSTICA BÁSICA PARA INGENIEROS*. Guayaquil, Ecuador: ICM ESPOL.
- Rumsey, F., & McCormick, T. (2009). *Sound and Recording*. Taylor & Francis.
- Schiff, J. L. (2013). *The Laplace Transform: Theory and Applications*. Springer Science & Business Media.
- Shirley, B., Kendrick, P., & Claire, C. (2007). The Effect of Stereo Crosstalk on Intelligibility: Comparison of a Phantom Stereo Image and a Central Loudspeaker Source. *Journal of the Audio Engineering Society*, 852-863.
- Statista. (2017a). *Usage frequency of headphones in the United States*. Obtenido de <https://www.statista.com/statistics/283620/us-consumer-purchase-plans-smartphone-accessories-2010/>
- Statista. (2017b). *Purposes headphones are used for in the United States*. Obtenido de <https://www.statista.com/statistics/696862/uses-of-headphones-in-the-us/>
- Surhone, M., L., T. T., & Marseken, S. F. (2010). *Shapiro-Wilk Test*. VDM Publishing.
- Swart, D., Bekker, A., & Bienert, J. (2018). The subjective dimensions of sound quality of standard production electric vehicles. *Applied Acoustics* 129, 354-364.
- Tan, L. (2008). *Fundamentals of Analog and Digital Signal Processing*. AuthorHouse.
- Theile, G. (1991). On the Naturalness of Two-Channel Stereo Sound. *Journal of the Audio Engineering Society*, Vol. 39, No. 10, 761-767.
- Tyril, M., & Pedersen, J. A. (2001). Digital Filters for Low-Frequency Equalization. *Journal of the Audio Engineering Society*, Vol. 49, No 1/2, 36-43.
- UIT-R, R. B. (1997). Métodos para la evaluación subjetiva de la calidad de sonido.
- Walpole, R., Myers, R., Myers, S., & Keying, Y. (2012). *Probabilidad y estadística para ingeniería y ciencias Novena edición*. Mexico: PEARSON.
- Willert, V. (2006). A Probabilistic Model for Binaural Sound Localization. *IEEE Transactions on Systems, Man and Cybernetics Part B: Cybernetics*, Vol. 36, No. 5, 13.
- Zölzer, U. (2002). *DAFX - Digital Audio Effects*. John Wiley & Sons, LTD.

8 Anexos

8.1 Anexo A

8.1.1 Tablas de distribución normal Z

normal	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,50000	0,50399	0,50798	0,51197	0,51595	0,51994	0,52392	0,52790	0,53188	0,53586
0,1	0,53983	0,54380	0,54776	0,55172	0,55567	0,55962	0,56356	0,56749	0,57142	0,57535
0,2	0,57926	0,58317	0,58706	0,59095	0,59483	0,59871	0,60257	0,60642	0,61026	0,61409
0,3	0,61791	0,62172	0,62552	0,62930	0,63307	0,63683	0,64058	0,64431	0,64803	0,65173
0,4	0,65542	0,65910	0,66276	0,66640	0,67003	0,67364	0,67724	0,68082	0,68439	0,68793
0,5	0,69146	0,69497	0,69847	0,70194	0,70540	0,70884	0,71226	0,71566	0,71904	0,72240
0,6	0,72575	0,72907	0,73237	0,73565	0,73891	0,74215	0,74537	0,74857	0,75175	0,75490
0,7	0,75804	0,76115	0,76424	0,76730	0,77035	0,77337	0,77637	0,77935	0,78230	0,78524
0,8	0,78814	0,79103	0,79389	0,79673	0,79955	0,80234	0,80511	0,80785	0,81057	0,81327
0,9	0,81594	0,81859	0,82121	0,82381	0,82639	0,82894	0,83147	0,83398	0,83646	0,83891
1,0	0,84134	0,84375	0,84614	0,84849	0,85083	0,85314	0,85543	0,85769	0,85993	0,86214
1,1	0,86433	0,86650	0,86864	0,87076	0,87286	0,87493	0,87698	0,87900	0,88100	0,88298
1,2	0,88493	0,88686	0,88877	0,89065	0,89251	0,89435	0,89617	0,89796	0,89973	0,90147
1,3	0,90320	0,90490	0,90658	0,90824	0,90988	0,91149	0,91308	0,91466	0,91621	0,91774
1,4	0,91924	0,92073	0,92220	0,92364	0,92507	0,92647	0,92785	0,92922	0,93056	0,93189
1,5	0,93319	0,93448	0,93574	0,93699	0,93822	0,93943	0,94062	0,94179	0,94295	0,94408
1,6	0,94520	0,94630	0,94738	0,94845	0,94950	0,95053	0,95154	0,95254	0,95352	0,95449
1,7	0,95543	0,95637	0,95728	0,95818	0,95907	0,95994	0,96080	0,96164	0,96246	0,96327
1,8	0,96407	0,96485	0,96562	0,96638	0,96712	0,96784	0,96856	0,96926	0,96995	0,97062
1,9	0,97128	0,97193	0,97257	0,97320	0,97381	0,97441	0,97500	0,97558	0,97615	0,97670
2,0	0,97725	0,97778	0,97831	0,97882	0,97932	0,97982	0,98030	0,98077	0,98124	0,98169
2,1	0,98214	0,98257	0,98300	0,98341	0,98382	0,98422	0,98461	0,98500	0,98537	0,98574
2,2	0,98610	0,98645	0,98679	0,98713	0,98745	0,98778	0,98809	0,98840	0,98870	0,98899
2,3	0,98928	0,98956	0,98983	0,99010	0,99036	0,99061	0,99086	0,99111	0,99134	0,99158
2,4	0,99180	0,99202	0,99224	0,99245	0,99266	0,99286	0,99305	0,99324	0,99343	0,99361
2,5	0,99379	0,99396	0,99413	0,99430	0,99446	0,99461	0,99477	0,99492	0,99506	0,99520
2,6	0,99534	0,99547	0,99560	0,99573	0,99585	0,99598	0,99609	0,99621	0,99632	0,99643
2,7	0,99653	0,99664	0,99674	0,99683	0,99693	0,99702	0,99711	0,99720	0,99728	0,99736
2,8	0,99744	0,99752	0,99760	0,99767	0,99774	0,99781	0,99788	0,99795	0,99801	0,99807
2,9	0,99813	0,99819	0,99825	0,99831	0,99836	0,99841	0,99846	0,99851	0,99856	0,99861
3,0	0,99865	0,99869	0,99874	0,99878	0,99882	0,99886	0,99889	0,99893	0,99896	0,99900
3,1	0,99903	0,99906	0,99910	0,99913	0,99916	0,99918	0,99921	0,99924	0,99926	0,99929
3,2	0,99931	0,99934	0,99936	0,99938	0,99940	0,99942	0,99944	0,99946	0,99948	0,99950
3,3	0,99952	0,99953	0,99955	0,99957	0,99958	0,99960	0,99961	0,99962	0,99964	0,99965
3,4	0,99966	0,99968	0,99969	0,99970	0,99971	0,99972	0,99973	0,99974	0,99975	0,99976
3,5	0,99977	0,99978	0,99978	0,99979	0,99980	0,99981	0,99981	0,99982	0,99983	0,99983
3,6	0,99984	0,99985	0,99985	0,99986	0,99986	0,99987	0,99987	0,99988	0,99988	0,99989
3,7	0,99989	0,99990	0,99990	0,99990	0,99991	0,99991	0,99992	0,99992	0,99992	0,99992
3,8	0,99993	0,99993	0,99993	0,99994	0,99994	0,99994	0,99994	0,99995	0,99995	0,99995
3,9	0,99995	0,99995	0,99996	0,99996	0,99996	0,99996	0,99996	0,99996	0,99997	0,99997
4,0	0,99997	0,99997	0,99997	0,99997	0,99997	0,99997	0,99998	0,99998	0,99998	0,99998

normal	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
-4,0	0,00003	0,00003	0,00003	0,00003	0,00003	0,00003	0,00002	0,00002	0,00002	0,00002
-3,9	0,00005	0,00005	0,00004	0,00004	0,00004	0,00004	0,00004	0,00004	0,00003	0,00003
-3,8	0,00007	0,00007	0,00007	0,00006	0,00006	0,00006	0,00006	0,00005	0,00005	0,00005
-3,7	0,00011	0,00010	0,00010	0,00010	0,00009	0,00009	0,00008	0,00008	0,00008	0,00008
-3,6	0,00016	0,00015	0,00015	0,00014	0,00014	0,00013	0,00013	0,00012	0,00012	0,00011
-3,5	0,00023	0,00022	0,00022	0,00021	0,00020	0,00019	0,00019	0,00018	0,00017	0,00017
-3,4	0,00034	0,00032	0,00031	0,00030	0,00029	0,00028	0,00027	0,00026	0,00025	0,00024
-3,3	0,00048	0,00047	0,00045	0,00043	0,00042	0,00040	0,00039	0,00038	0,00036	0,00035
-3,2	0,00069	0,00066	0,00064	0,00062	0,00060	0,00058	0,00056	0,00054	0,00052	0,00050
-3,1	0,00097	0,00094	0,00090	0,00087	0,00084	0,00082	0,00079	0,00076	0,00074	0,00071
-3,0	0,00135	0,00131	0,00126	0,00122	0,00118	0,00114	0,00111	0,00107	0,00104	0,00100
-2,9	0,00187	0,00181	0,00175	0,00169	0,00164	0,00159	0,00154	0,00149	0,00144	0,00139
-2,8	0,00256	0,00248	0,00240	0,00233	0,00226	0,00219	0,00212	0,00205	0,00199	0,00193
-2,7	0,00347	0,00336	0,00326	0,00317	0,00307	0,00298	0,00289	0,00280	0,00272	0,00264
-2,6	0,00466	0,00453	0,00440	0,00427	0,00415	0,00402	0,00391	0,00379	0,00368	0,00357
-2,5	0,00621	0,00604	0,00587	0,00570	0,00554	0,00539	0,00523	0,00508	0,00494	0,00480
-2,4	0,00820	0,00798	0,00776	0,00755	0,00734	0,00714	0,00695	0,00676	0,00657	0,00639
-2,3	0,01072	0,01044	0,01017	0,00990	0,00964	0,00939	0,00914	0,00889	0,00866	0,00842
-2,2	0,01390	0,01355	0,01321	0,01287	0,01255	0,01222	0,01191	0,01160	0,01130	0,01101
-2,1	0,01786	0,01743	0,01700	0,01659	0,01618	0,01578	0,01539	0,01500	0,01463	0,01426
-2,0	0,02275	0,02222	0,02169	0,02118	0,02068	0,02018	0,01970	0,01923	0,01876	0,01831
-1,9	0,02872	0,02807	0,02743	0,02680	0,02619	0,02559	0,02500	0,02442	0,02385	0,02330
-1,8	0,03593	0,03515	0,03438	0,03362	0,03288	0,03216	0,03144	0,03074	0,03005	0,02938
-1,7	0,04457	0,04363	0,04272	0,04182	0,04093	0,04006	0,03920	0,03836	0,03754	0,03673
-1,6	0,05480	0,05370	0,05262	0,05155	0,05050	0,04947	0,04846	0,04746	0,04648	0,04551
-1,5	0,06681	0,06552	0,06426	0,06301	0,06178	0,06057	0,05938	0,05821	0,05705	0,05592
-1,4	0,08076	0,07927	0,07780	0,07636	0,07493	0,07353	0,07215	0,07078	0,06944	0,06811
-1,3	0,09680	0,09510	0,09342	0,09176	0,09012	0,08851	0,08692	0,08534	0,08379	0,08226
-1,2	0,11507	0,11314	0,11123	0,10935	0,10749	0,10565	0,10383	0,10204	0,10027	0,09853
-1,1	0,13567	0,13350	0,13136	0,12924	0,12714	0,12507	0,12302	0,12100	0,11900	0,11702
-1,0	0,15866	0,15625	0,15386	0,15151	0,14917	0,14686	0,14457	0,14231	0,14007	0,13786
-0,9	0,18406	0,18141	0,17879	0,17619	0,17361	0,17106	0,16853	0,16602	0,16354	0,16109
-0,8	0,21186	0,20897	0,20611	0,20327	0,20045	0,19766	0,19489	0,19215	0,18943	0,18673
-0,7	0,24196	0,23885	0,23576	0,23270	0,22965	0,22663	0,22363	0,22065	0,21770	0,21476
-0,6	0,27425	0,27093	0,26763	0,26435	0,26109	0,25785	0,25463	0,25143	0,24825	0,24510
-0,5	0,30854	0,30503	0,30153	0,29806	0,29460	0,29116	0,28774	0,28434	0,28096	0,27760
-0,4	0,34458	0,34090	0,33724	0,33360	0,32997	0,32636	0,32276	0,31918	0,31561	0,31207
-0,3	0,38209	0,37828	0,37448	0,37070	0,36693	0,36317	0,35942	0,35569	0,35197	0,34827
-0,2	0,42074	0,41683	0,41294	0,40905	0,40517	0,40129	0,39743	0,39358	0,38974	0,38591
-0,1	0,46017	0,45620	0,45224	0,44828	0,44433	0,44038	0,43644	0,43251	0,42858	0,42465
0,0	0,50000	0,49601	0,49202	0,48803	0,48405	0,48006	0,47608	0,47210	0,46812	0,46414

8.1.2 Protocolo para pruebas de preferencia

Tabla 25. Tabla de protocolo de aplicación de pruebas de preferencia

Protocolo de aplicación de pruebas de preferencia	
Cada participante será citado a un hora especifica teniendo en cuenta la duración de la prueba y las posibles variables	
Duración	Actividad
5 min	Se tendrán 5 minutos de tolerancia después de la hora acordada para esperar al participante. En este tiempo se ubicará en la posición para realizar la prueba
1 - 3 min	Se le explicara al participante en que consiste la prueba, el objetivo de esta y las tareas que debe ejecutar
3 - 5 min	El participante leerá el documento de consentimiento detalladamente y decidirá participar o no. Entre tanto, un responsable comprobara el estado de los equipos a usar dentro de la prueba para que estén en óptimas condiciones
1-2 min	Se le explicara al participante el funcionamiento de la interfaz gráfica y este se relacionará con los controles, eligiendo el nivel de ganancia en los audífonos que crea pertinente según el audio de ejemplo
30 - 35 min	La prueba contiene un total de 20 presentaciones de audio de 15 segundos por tema musical organizadas en pares aleatorios, son tres temas musicales y el participante podrá repetir cada muestra máximo una vez. Por lo cual se estima un tiempo de 30 minutos sin contar las posibles variables, esto dependerá de las condiciones propias de cada participante
	Se hará uso de dos estaciones para realizar la prueba, es decir se podrán realizar dos pruebas simultaneas, una en el live room y otra en el control room del estudio hibrido de la universidad San Buenaventura.
1 min	al finalizar la prueba se le agradecerá al participante por la colaboración y tiempo invertido
40 - 45min	Total de tiempo

8.1.3 Consentimiento de la prueba de preferencia por medio de pares comparados

Consentimiento Informado para Participantes de pruebas subjetivas

El propósito de este documento de consentimiento es proveer a los participantes de este proceso de evaluación una clara explicación de la naturaleza de la misma, así como de su rol en ella como participantes.

La presente investigación es desarrollada por los estudiantes Jhon Barreto, Mateo Bejarano, Yonathan Layton y es conducida por el profesor Luis Fernando Hermida de la Universidad San Buenaventura sede Bogotá. El objetivo de esta prueba es evaluar pistas de audio procesadas por 4 métodos de simulación de crosstalk diferentes. La metodología usada en esta prueba se fundamenta en evaluaciones de calidad sonora de carácter subjetivo por pares comparados.

Si usted accede a participar en este estudio, se le pedirá usar los audífonos en todo momento durante la prueba, para lo cual deberá ajustar la ganancia a un nivel que resulte cómodo para usted. Durante la prueba usted escuchará pistas de audio de 15 segundos, las cuales constan de 4 audios procesados digitalmente, así como el audio original para cada tema musical. **Su tarea será seleccionar la pista que sea de su preferencia general.** Esto se repetirá para 3 canciones, cada una de un género distinto. Podrá escuchar cada par un máximo de 2 veces. El tiempo estimado para la prueba es de aproximadamente 30 minutos.

La participación en este estudio es estrictamente voluntaria. La información que se recoja será confidencial y solo se usará con fines académicos. Sus respuestas serán codificadas usando el número de prueba que usted respondió y, por lo tanto, serán anónimas.

Si tiene alguna duda sobre este proyecto no dude en preguntar a uno de los responsables en cualquier momento durante su participación en él. Así mismo, puede retirarse del proyecto en cualquier momento sin que eso lo perjudique en ninguna forma.

Desde ya le agradecemos su participación.

Acepto participar voluntariamente en esta prueba, conducida por los estudiantes Jhon Barreto, Mateo Bejarano, Yonathan Layton y conducida por el profesor Luis Fernando Hermida de la Universidad San Buenaventura sede Bogotá. He sido informado (a) que el objetivo de esta prueba es realizar una escucha crítica para seleccionar una de las opciones de la pista que sea de mi preferencia.

Reconozco que la información que yo provea en el proceso de evaluación subjetiva es estrictamente confidencial y solo será usada con fines académicos. He sido informado(a) que puedo hacer preguntas sobre el proyecto en cualquier momento y que puedo retirarme del mismo cuando así lo decida, sin que esto acarree perjuicio alguno para mi persona. De tener preguntas sobre mi participación en este estudio, puedo contactar a Jhon Fredy Barreto al teléfono (+57) 3156146414.

Entiendo que puedo pedir información sobre los resultados de este estudio cuando éste se encuentre concluido. Para esto, puedo contactar a Jhon Fredy Barreto al teléfono anteriormente mencionado.

Firma del Participante

Firma del responsable

Fecha

8.1.4 Resultados y aplicación de la ley Thurstone para prueba de preferencia

A continuación, se presentan las tablas de resultados obtenidos por el método de pares comparados en (Resultados de las pruebas de preferencia de los métodos de simulación de crosstalk), usando la ley de juicio comparativo de Thurstone.

Tabla 26. Matriz de frecuencia observada de los diferentes métodos de procesamiento para el tema Symphony No 5.

Symphony No 5					
	5	2	3	1	4
1	9	4	15	0	12
2	10	0	15	21	18
3	8	10	0	10	16
4	9	7	9	13	0
5	0	15	17	16	16
Total	36	36	56	60	62

Tabla 27. Matriz de proporciones ordenada de los datos para el tema Symphony No. 5.

Matriz de Proporciones					
Symphony No 5					
	5	2	3	1	4
1	0,36	0,16	0,6	0	0,48
2	0,4	0	0,6	0,84	0,72
3	0,32	0,4	0	0,4	0,64
4	0,36	0,28	0,36	0,52	0
5	0	0,6	0,68	0,64	0,64
Desviación estándar	0,2559				

Tabla 28. Matriz de proporciones normalizadas Z para tema Symphony No. 5.

Matriz de proporciones normalizadas					
Symphony No 5					
	2	5	3	1	4
1	-0,9945	-0,3585	0,2533		-0,0502
2		-0,2533	0,2533	0,9945	0,5828
3	-0,2533	-0,4677		-0,2533	0,3585
4	-0,5828	-0,3585	-0,3585	0,0502	
5	0,2533		0,4677	0,3585	0,3585
SUMA	-1,5773	-1,4380	0,6159	1,1497	1,2496
PROMEDIO	-0,3155	-0,2876	0,1232	0,2299	0,2499
DISTRIBUCIÓN Z	0,0000	0,0279	0,4386	0,5454	0,5654

Tabla 29. Matriz de frecuencia observada de los diferentes métodos de procesamiento para el tema Con Calma.

Con Calma					
	5	2	3	1	4
1	2	16	9	0	15
2	6	0	15	9	19
3	5	10	0	16	14
4	3	6	11	10	0
5	0	19	20	23	22
Total	16	51	55	58	70

Tabla 30. Matriz de proporciones ordenada de los datos para Con Calma.

Matriz de Proporciones					
Con Calma					
	5	2	3	1	4
1	0,08	0,64	0,36	0	0,6
2	0,24	0	0,6	0,36	0,76
3	0,2	0,4	0	0,64	0,56
4	0,12	0,24	0,44	0,4	0
5	0	0,76	0,8	0,92	0,88

Desviación estándar	0,3040
---------------------	--------

Tabla 31. Matriz de proporciones normalizadas Z para el tema Con Calma.

Matriz de proporciones normalizadas					
Con Calma					
	5	2	3	1	4
1	-1,4051	0,3585	-0,3585	0,0000	0,2533
2	-0,7063	0,0000	0,2533	-0,3585	0,7063
3	-0,8416	-0,2533	0,0000	0,3585	0,1510
4	-1,1750	-0,7063	-0,1510	-0,2533	0,0000
5	0,0000	0,7063	0,8416	1,4051	1,1750
SUMA	-4,1280	0,1051	0,5855	1,1517	2,2856
PROMEDIO	-0,8256	0,0210	0,1171	0,2303	0,4571
DISTRIBUCIÓN Z	0,0000	0,8466	0,9427	1,0559	1,2827

Tabla 32. Matriz de frecuencia observada de los diferentes métodos de procesamiento para el tema Money.

Money					
	5	2	3	4	1
1	5	3	13	12	0
2	8	0	21	18	22
3	8	4	0	14	12
4	3	7	11	0	13
5	0	17	17	22	20
Total	24	31	62	66	67

Tabla 33. Matriz de proporciones ordenada de los datos para el tema Money.

Matriz de Proporciones					
Money					
	5	2	3	4	1
1	0,2	0,12	0,52	0,48	0
2	0,32	0	0,84	0,72	0,88
3	0,32	0,16	0	0,56	0,48
4	0,12	0,28	0,44	0	0,52
5	0	0,68	0,68	0,88	0,8
Desviación estándar	0,3044				

Tabla 34. Matriz de proporciones normalizadas Z para el tema Money.

Matriz de proporciones normalizadas					
Money					
	5	2	3	4	1
1	-0,8416	-1,1750	0,0502	-0,0502	0,0000
2	-0,4677	0,0000	0,9945	0,5828	1,1750
3	-0,4677	-0,9945	0,0000	0,1510	-0,0502
4	-1,1750	-0,5828	-0,1510	0,0000	0,0502
5	0,0000	0,4677	0,4677	1,1750	0,8416
SUMA	-2,9520	-2,2846	1,3613	1,8586	2,0166
PROMEDIO	-0,5904	-0,4569	0,2723	0,3717	0,4033
DISTRIBUCIÓN Z	0,0000	0,1335	0,8627	0,9621	0,9937

Tabla 35. Matriz de frecuencia observada de los diferentes métodos de procesamiento para las tres canciones

Total de datos consolidados					
	5	2	3	1	4
1	16	23	37	0	39
2	24	0	51	52	55
3	21	24	0	38	44
4	15	20	31	36	0
5	0	51	54	59	60
Total	76	118	173	185	198

Tabla 36. Matriz de proporciones ordenada de los datos consolidados para los tres temas.

Matriz de Proporciones					
Total de datos consolidados					
	5	2	3	1	4
1	0,2133	0,3067	0,4933	0,0000	0,5200
2	0,3200	0,0000	0,6800	0,6933	0,7333
3	0,2800	0,3200	0,0000	0,5067	0,5867
4	0,2000	0,2667	0,4133	0,4800	0,0000
5	0,0000	0,6800	0,7200	0,7867	0,8000
Desviación estándar	0,2716				

Tabla 37. Matriz de proporciones normalizadas Z para los tres temas.

Total					
	5	2	3	1	4
1	-0,7949	-0,5053	-0,0167	0,0000	0,0502
2	-0,4677	0,0000	0,4677	0,5053	0,6229
3	-0,5828	-0,4677	0,0000	0,0167	0,2190
4	-0,8416	-0,6229	-0,2190	-0,0502	0,0000
5	0,0000	0,4677	0,5828	0,7949	0,8416
SUMA	-2,6871	-1,1282	0,8149	1,2668	1,7337
PROMEDIO	-0,5374	-0,2256	0,1630	0,2534	0,3467
DSTRIBUCIÓN Z	0,0000	0,3118	0,7004	0,7908	0,8841

8.2 Anexo B: Anexos de prueba de calidad sonora de la aplicación móvil.

8.2.1 Consentimiento informado para participantes de prueba de calidad sonora

Consentimiento Informado para Participantes de pruebas subjetivas finales de calidad sonora

El propósito de este documento de consentimiento es proveer a los participantes de este proceso de evaluación una clara explicación de la naturaleza de la misma, así como de su rol en ella como participantes. La presente investigación es desarrollada por los estudiantes Jhon Barreto, Mateo Bejarano y Yonathan Layton, es conducida por el profesor Luis Fernando Hermida de la Universidad San Buenaventura sede Bogotá. El objetivo de esta prueba es evaluar pistas de audio procesadas por 2 métodos de simulación de crosstalk diferentes, mediante una aplicación en dispositivos móviles. La metodología usada en esta prueba se fundamenta en evaluaciones de calidad sonora de carácter subjetivo por escala de diferencial semántica. Si usted accede a participar en este estudio, se le pedirá usar los audífonos en todo momento durante la prueba, para lo cual deberá ajustar la ganancia a un nivel que resulte cómodo para usted. Durante la prueba usted escuchará diferentes canciones de diferentes géneros, en las cuales podrá elegir alguno de los dos métodos de procesamiento o el audio original en tiempo real. **Su tarea será seleccionar para cada método, según cada pregunta un valor en la escala según su juicio.** El tiempo estimado para la prueba es de aproximadamente 30 minutos.

La participación en este estudio es estrictamente voluntaria. La información que se recoja será confidencial y solo se usará con fines académicos. Sus respuestas serán codificadas usando el número de prueba que usted respondió y, por lo tanto, serán anónimas.

Si tiene alguna duda sobre este proyecto no dude en preguntar a uno de los responsables en cualquier momento durante su participación en él. Así mismo, puede retirarse del proyecto en cualquier momento sin que eso lo perjudique en ninguna forma.

Desde ya le agradecemos su participación.

Acepto participar voluntariamente en esta prueba, conducida por los estudiantes Jhon Barreto, Mateo Bejarano, Yonathan Layton y conducida por el profesor Luis Fernando Hermida de la Universidad San Buenaventura sede Bogotá. He sido informado (a) que el objetivo de esta prueba es realizar una escucha crítica para seleccionar una de las opciones de la pista que sea de mi preferencia.

Reconozco que la información que yo provea en el proceso de evaluación subjetiva es estrictamente confidencial y solo será usada con fines académicos. He sido informado(a) que puedo hacer preguntas sobre el proyecto en cualquier momento y que puedo retirarme del mismo cuando así lo decida, sin que esto acarree perjuicio alguno para mi persona. De tener preguntas sobre mi participación en este estudio, puedo contactar a Jhon Fredy Barreto al teléfono (+57) 3156146414.

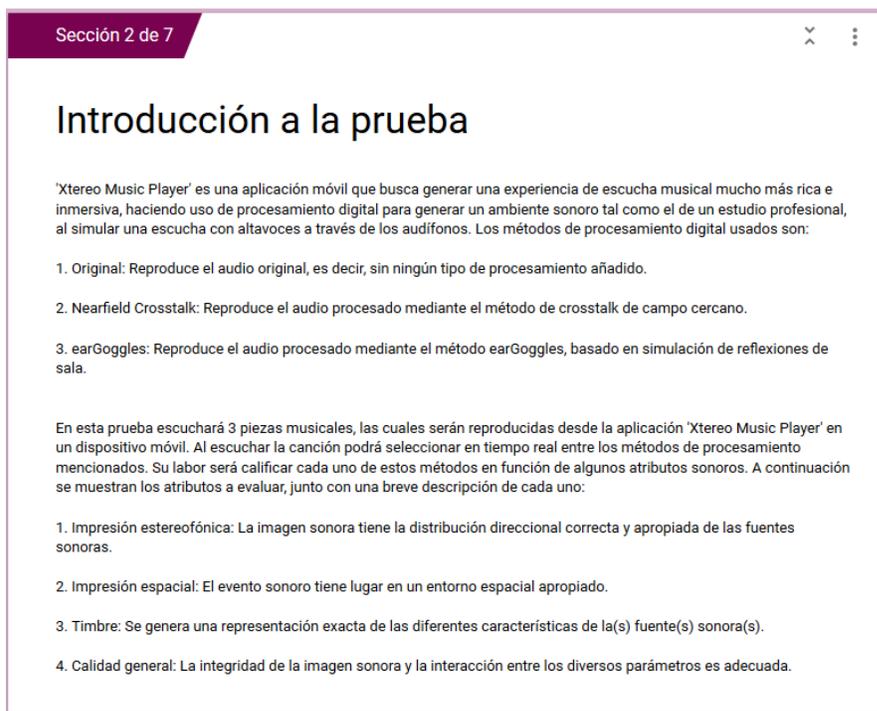
Entiendo que puedo pedir información sobre los resultados de este estudio cuando éste se encuentre concluido. Para esto, puedo contactar a Jhon Fredy Barreto al teléfono anteriormente mencionado.

8.2.2 Protocolo de aplicación para pruebas de calidad sonora

Tabla 38. Procedimiento para aplicación de pruebas de calidad sonora de la aplicación móvil.

Cada participante será citado a un hora específica teniendo en cuenta la duración de la prueba la disponibilidad del mismo y las posibles variables	
Duración	Actividad
5 min	Se tendrán 5 minutos de tolerancia después de la hora acordada para esperar al participante. En este tiempo se ubicará en la posición para la prueba
1 - 3 min	Se le explicara al participante en que consiste la prueba, el objetivo de la misma y las tareas que debe ejecutar
3 - 5 min	El participante leerá el documento de consentimiento detalladamente y decidirá participar o no. Entre tanto, un responsable comprobara el estado de los equipos a usar dentro de la prueba para que estén en óptimas condiciones
1-2 min	Se le explicara al participante el funcionamiento de la interfaz gráfica y este se relacionará con los controles, eligiendo el nivel de ganancia en los audífonos que crea pertinente según el usuario
6 - 15 min	La prueba contiene un total de 3 temas musicales de 2 minutos como máximo por tema musical, el participante podrá repetir cada canción máximo una vez, teniendo en cuenta que puede cambiar entre los dos diferentes métodos de procesamiento en tiempo real las veces que crea necesario Por lo cual se estima un tiempo de 15 minutos sin contar las posibles variables, esto dependerá de las condiciones propias de cada participante
	Se hará uso de tres estaciones para realizar la prueba, es decir se podrán realizar dos pruebas simultaneas, una en el live room y otra en el control room del estudio hibrido de la universidad San Buenaventura.
1 min	al finalizar la prueba se le agradecerá al participante por la colaboración y tiempo invertido
17 -31min	Total de tiempo

8.2.3 Imágenes de Google Form para prueba de calidad sonora.



Sección 2 de 7

Introducción a la prueba

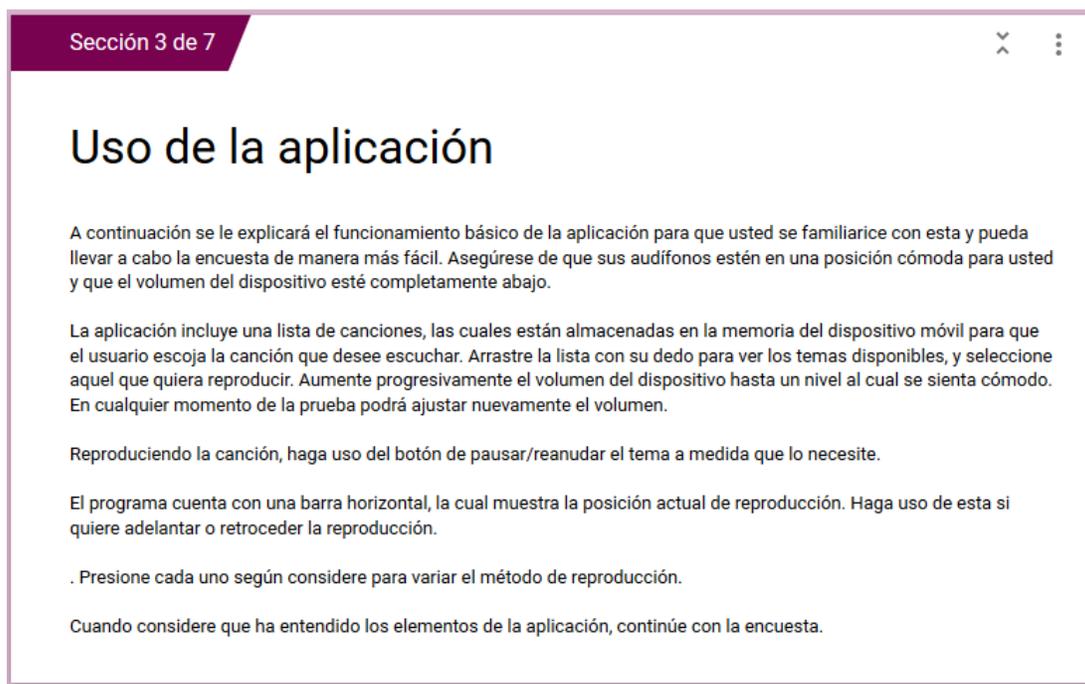
'Xtereo Music Player' es una aplicación móvil que busca generar una experiencia de escucha musical mucho más rica e inmersiva, haciendo uso de procesamiento digital para generar un ambiente sonoro tal como el de un estudio profesional, al simular una escucha con altavoces a través de los audífonos. Los métodos de procesamiento digital usados son:

1. Original: Reproduce el audio original, es decir, sin ningún tipo de procesamiento añadido.
2. Nearfield Crosstalk: Reproduce el audio procesado mediante el método de crosstalk de campo cercano.
3. earGoggles: Reproduce el audio procesado mediante el método earGoggles, basado en simulación de reflexiones de sala.

En esta prueba escuchará 3 piezas musicales, las cuales serán reproducidas desde la aplicación 'Xtereo Music Player' en un dispositivo móvil. Al escuchar la canción podrá seleccionar en tiempo real entre los métodos de procesamiento mencionados. Su labor será calificar cada uno de estos métodos en función de algunos atributos sonoros. A continuación se muestran los atributos a evaluar, junto con una breve descripción de cada uno:

1. Impresión estereofónica: La imagen sonora tiene la distribución direccional correcta y apropiada de las fuentes sonoras.
2. Impresión espacial: El evento sonoro tiene lugar en un entorno espacial apropiado.
3. Timbre: Se genera una representación exacta de las diferentes características de la(s) fuente(s) sonora(s).
4. Calidad general: La integridad de la imagen sonora y la interacción entre los diversos parámetros es adecuada.

Figura 39. Instrucciones en formulario Google Form para las pruebas de evaluación de calidad sonora.



Sección 3 de 7

Uso de la aplicación

A continuación se le explicará el funcionamiento básico de la aplicación para que usted se familiarice con esta y pueda llevar a cabo la encuesta de manera más fácil. Asegúrese de que sus audífonos estén en una posición cómoda para usted y que el volumen del dispositivo esté completamente abajo.

La aplicación incluye una lista de canciones, las cuales están almacenadas en la memoria del dispositivo móvil para que el usuario escoja la canción que desee escuchar. Arrastre la lista con su dedo para ver los temas disponibles, y seleccione aquel que quiera reproducir. Aumente progresivamente el volumen del dispositivo hasta un nivel al cual se sienta cómodo. En cualquier momento de la prueba podrá ajustar nuevamente el volumen.

Reproduciendo la canción, haga uso del botón de pausar/reanudar el tema a medida que lo necesite.

El programa cuenta con una barra horizontal, la cual muestra la posición actual de reproducción. Haga uso de esta si quiere adelantar o retroceder la reproducción.

. Presione cada uno según considere para variar el método de reproducción.

Cuando considere que ha entendido los elementos de la aplicación, continúe con la encuesta.

Figura 40. Instrucciones para uso de la aplicación móvil en el formulario de Google Form.

Sección 5 de 7

Tema 2: Pink Floyd - Money (1973)

A continuación deberá calificar los atributos sonoros del tema 'Money'. Por favor selecciónela en la aplicación y escúchela. Puede escuchar cada método de procesamiento a medida que va respondiendo la encuesta, o si prefiere puede escucharlos en su totalidad para después responder.

Procesamiento 'Original'

La preguntas que vienen a continuación buscan evaluar el audio original. Escuche detalladamente la pista para responder.

¿Cómo califica la impresión estereofónica (la imagen sonora tiene la distribución direccional correcta y apropiada de las fuentes sonoras)? *

- 5. Muy Buena
- 4. Buena
- 3. Regular
- 2. Deficiente
- 1. Muy Deficiente

Figura 41. Ejemplo de evaluación de calidad sonora en formulario Google Form.

8.3 Anexo C

Link para visualización y descarga del Unity Project de la aplicación 'Xtereo Music Player'.

<https://github.com/ambejarano/XtereoMusicPlayer>