

RAE

1. **TIPO DE DOCUMENTO:** Trabajo de grado para optar por el título de INGENIERA DE SONIDO.
2. **TÍTULO:** DESARROLLO DE UN DISPOSITIVO POR MEDIO DE UN SISTEMA EMBEBIDO QUE IMPLEMENTE EL MEJOR MÉTODO OBTENIDO DEL ANÁLISIS COMPARATIVO ENTRE LPC, SPECTRUM Y BANCO DE FILTROS PARA MODULAR LA SEÑAL DE UN INSTRUMENTO MUSICAL CON LA ENVOLVENTE DE LA VOZ.
3. **AUTOR:** María Paula Claro Vargas.
4. **LUGAR:** Bogotá, D.C
5. **FECHA:** Mayo de 2015.
6. **PALABRAS CLAVE:** Banco de filtros, brillo, calidez, filtros digitales, inteligibilidad, instrumento musical, linear predictive coding, moduladora, Modulación AM, procesamiento digital de señales, portadora, señal de voz, señal de audio, sistema embebido, transformada rápida de Fourier.
7. **DESCRIPCIÓN DEL TRABAJO:** El objetivo principal de este proyecto es desarrollar un dispositivo por medio de un sistema embebido, que implemente el mejor método obtenido del análisis comparativo entre los métodos de reconstrucción de la voz LPC, Spectrum y Banco de filtros con el fin de modular en amplitud la señal de un instrumento musical con la señal de la voz y determinar si el sonido obtenido de esta modulación puede ser usado con fines creativos dentro de la producción o musical, audiovisual, entre otras.
8. **LÍNEAS DE INVESTIGACIÓN:** El proyecto pertenece a la línea de investigación de la USB: Tecnologías actuales y sociedad. La línea de investigación de la facultad y el núcleo problemático pertenecen a la rama de análisis y procesamiento de señales.
9. **METODOLOGÍA:** El enfoque de este proyecto de grado es Empírico-Analítico.
10. **CONCLUSIONES:** Se concluye que es posible llevar a cabo el desarrollo de un dispositivo en un sistema embebido que implemente un método elegido entre LPC, Spectrum y Banco de filtros. Eligiendo el que obtenga mejores resultados en mediciones de Latencia, consumo de memoria RAM y comparación de las señales de entrada y salida del algoritmo mediante un análisis espectral de las mismas. Además, la señal de la voz obtenida por el algoritmo elegido puede usarse como la señal moduladora en el proceso de modulación en amplitud entre la voz y un instrumento musical y el sonido obtenido de la modulación puede usarse dentro de procesos creativos en la música y en la ingeniería de sonido.

**Desarrollo de un dispositivo por medio de un sistema embebido, que
implemente el mejor método, obtenido del análisis comparativo entre LPC,
spectrum y banco de filtros para modular la señal de un instrumento musical
con la envolvente de la voz** ii

Tesis Presentada Para Obtener El Título De
Ingeniera de Sonido
Universidad de San Buenaventura, Bogotá

María Paula Claro Vargas
Mayo 2015

Agradecimientos

Gracias infinitas a mi familia, en especial a mi mamá y a mi papá por su amor, su paciencia y por permitirme estudiar. A mis tías Fabiola, Stella y Lucy por su apoyo, su cariño y por ser incondicionales conmigo.

Gracias a mis amigas por ser mi alegría en los momentos difíciles durante estos 5 años.

Gracias a mis profesores, en especial a mi tutor, el Ingeniero Miguel Sotaquirá por ser mi guía durante el desarrollo de mi proyecto de grado, al Ingeniero Félix Julián Gutiérrez por su colaboración en este proceso, y al Ingeniero Manuel Torres por ayudarme con el nacimiento de la idea que dio vida a este proyecto.

Finalmente, gracias a mis compañeros de universidad y a todas las personas que han hecho parte de este camino que me ha formado como profesional.

Tabla de contenido

v

Capítulo 1.....	1
Introducción.....	1
Capítulo 2.....	3
Planteamiento del problema.....	3
Antecedentes.....	3
Descripción y formulación del problema.....	4
Pregunta problema.....	4
Justificación.....	5
Objetivos.....	6
Objetivo general.....	6
Objetivos específicos.....	6
Alcances y limitaciones.....	6
Alcances.....	6
Limitaciones.....	7
Capítulo 3.....	8
Marco de referencia.....	8
Marco teórico.....	8
Criterios de adquisición.....	8
Linear Predictive Coding (LPC).....	10
Spectrum.....	11
Banco de filtros.....	11
Detección de envoltura.....	12
Modulación AM.....	13
Modulación AM con portadora suprimida o de doble banda lateral DSB.....	17
Sistema embebido.....	18
Memoria RAM.....	19
Procesamiento digital de señales.....	20
Marco legal o normativo.....	21
Aspectos legales del software libre o de código abierto (Open Source).....	21
Capítulo 4.....	23
Metodología.....	23
Enfoque de investigación.....	23
Línea de investigación de la universidad/ línea de investigación de la facultad/ núcleos problemáticos.....	23
Hipótesis.....	23
Variables.....	24
Capítulo 5.....	25
Desarrollo ingenieril.....	25
Elección del sistema embebido.....	25
Memoria RAM.....	27
Entradas y salidas de audio.....	27
Resolución en bits de la conversión A/D.....	28
Memoria de almacenamiento.....	29
Entorno de desarrollo.....	29
Sistema embebido elegido.....	30
Adquisición y procesamiento de señales en el sistema NI myRIO.....	30
Validación de la etapa de amplificación con el integrado AD620.....	34
Implementación y medición de los algoritmos (Banco de filtros, Spectrum y LPC).....	38
Procedimiento de medición de latencia de los algoritmos.....	39

Procedimiento de medición de consumo de memoria RAM de los algoritmos	40vi
Procedimiento para la comparación espectral entre la señal de entrada y salida a cada algoritmo....	40
Implementación y medición de Banco de filtros.....	43
Implementación y medición de Spectrum	53
Implementación y medición de Linear Predictive Coding (LPC)	58
Detección de envolvente.....	66
Proceso de modulación AM	67
Caracterización de la señal del sintetizador.	67
Caracterización de la señal de voz.	69
Modulación AM en LabVIEW.....	71
Aplicación de encuesta para valoración subjetiva del sonido obtenido.....	71
Descripción de los audios utilizados en las encuestas.....	73
Categorización de las variables de medición.	74
Descripción del formulario generado para la encuesta.....	74
Capítulo 6	76
Presentación y análisis de resultados	76
Análisis comparativo de los algoritmos para determinar su eficiencia	76
Presentación y análisis de resultados de encuesta realizada para valoración subjetiva del sonido obtenido	83
Resultado de encuesta realizada a la muestra de estudiantes de ingeniería de sonido que están cursando sexto semestre o superiores.....	83
Resultado de encuesta realizada a la muestra de estudiantes de ingeniería de sonido que están cursando quinto semestre e inferiores.	87
Resultado de encuesta realizada a muestra de músicos de profesión.	91
CONCLUSIONES.....	95
RECOMENDACIONES	99
REFERENCIAS	101

Lista de tablas

Tabla 1. Características de los sistemas embebidos.....	26
Tabla 2. Especificaciones técnicas Sennheiser e835 y Shure SM 58.	31
Tabla 3. Especificaciones técnicas de 3 circuitos integrados.	32
Tabla 4. Voltaje de los armónicos de la señal. Generador PeakTeck 4025.	37
Tabla 5. Comparación de los filtros FIR e IIR.....	44
Tabla 6. Frecuencias de corte para el banco de filtros.....	48
Tabla 7. Frecuencias de corte para el banco de filtros con respuesta plana.....	49
Tabla 8. Resultados de las mediciones para Banco de filtros	51
Tabla 9. Error cuadrático medio de las señales de voz y de las envolventes.....	53
Tabla 10. Resultados de las mediciones para Spectrum	55
Tabla 11. Error cuadrático medio de las señales de voz y de las envolventes.....	57
Tabla 12. Error cuadrático medio entre señal filtrada y señal original.	63
Tabla 13. Resultado de las mediciones para LPC.....	64
Tabla 14. Error cuadrático medio de las señales de voz y de las envolventes.....	66
Tabla 15. Resultados de medición de latencia de cada algoritmo.	76
Tabla 16. Resultados de medición de consumo de memoria RAM.....	77
Tabla 17. Número de bloques necesarios en la implementación de cada algoritmo.	78
Tabla 18. Error cuadrático de las señales de entrada y salida para cada algoritmo.....	80
Tabla 19. Error cuadrático de las envolventes de las señales de voz de entrada y salida para cada algoritmo.	82

Lista de figuras

Figura 1. Método de detección de envolvente de ley cuadrática. Fuente: www.mathworks.com/help/dsp/examples/envelope-detection	13
Figura 2. Señal moduladora $m(t)$. Fernández, Marcos Martín (2002). Modulación en amplitud. [Figura]. Universidad de Valladolid	14
Figura 3. Señal modulada sin sobremodulación. Fernández, Marcos Martín (2002). Modulación en amplitud. [Figura]. Universidad de Valladolid	15
Figura 4. Señal modulada con sobremodulación. Fernández, Marcos Martín (2002). Modulación en amplitud. [Figura]. Universidad de Valladolid	15
Figura 5. Espectro de la señal moduladora. Fernández, Marcos Martín (2002). Modulación en amplitud. [Figura]. Universidad de Valladolid	16
Figura 6. Espectro de la señal modulada. Fernández, Marcos Martín (2002). Modulación en amplitud. [Figura]. Universidad de Valladolid	16
Figura 7. Señal modulada DSB. Fernández, Marcos Martín (2002). Modulación en amplitud. [Figura]. Universidad de Valladolid	17
Figura 8. Diagrama de conexiones a NI myRIO.....	34
Figura 9. Ganancia resultado de medición en modo diferencial, generador PeakTeck 4025.....	36
Figura 10. Ganancia resultado de medición en modo común, generador PeakTeck 4025.	38
Figura 11. Diagrama de bloques del algoritmo usado para medir latencia y tiempo de ejecución de un algoritmo.	39
Figura 12. La figura de arriba es la FFT de la señal de la voz. La figura de abajo corresponde a la FFT de la envolvente de la señal de la voz.	42
Figura 13. Diagrama de flujo del Banco de Filtros.....	44
Figura 14. La figura de la izquierda es la respuesta en frecuencia de filtro FIR pasabanda con ventana rectangular y la figura de la derecha es la respuesta en frecuencia del filtro FIR pasabanda con ventana Hamming.	46
Figura 15. Respuesta en frecuencia de filtro IIR para 20 coeficientes.	47
Figura 16. Respuesta en frecuencia de banco de filtros 1.....	49
Figura 17. Respuesta en frecuencia de banco de filtros 2.....	50
Figura 18. Diagrama de bloques del algoritmo Banco de Filtros.	50
Figura 19. La gráfica de la izquierda muestra la FFT de la señal de voz original (señal de entrada) y la gráfica de la derecha muestra la FFT de la señal de salida del banco de filtros.	51
Figura 20. La gráfica de la izquierda muestra la FFT de la envolvente de voz original (señal de entrada) y la gráfica de la derecha muestra la FFT de la envolvente de la señal de salida del banco de filtros.	52
Figura 21. Diagrama de flujo de Spectrum.....	54
Figura 22. Diagrama de bloques del algoritmo Spectrum.	55
Figura 23. La gráfica de la izquierda muestra la FFT de la señal de voz original (señal de entrada) y la gráfica de la derecha muestra la FFT de la señal de salida del método Spectrum.....	56
Figura 24. La gráfica de la izquierda muestra la FFT de la envolvente de voz original (señal de entrada) y la gráfica de la derecha muestra la FFT de la envolvente de la señal de salida del método Spectrum.	57
Figura 25. Diagrama de flujo de LPC.....	60
Figura 26. Señal original. Señal de entrada al algoritmo LPC.....	61

Figura 27. Señal original multiplicada por una ventana Hamming.	61ix
Figura 28. Salida del filtro FIR. Señal esperada.	62
Figura 29. Error de predicción obtenido.	63
Figura 30. Diagrama de bloques del algoritmo LPC.	64
Figura 31. La gráfica de la izquierda muestra la FFT de la señal de voz original (señal de entrada) y la gráfica de la derecha muestra la FFT de la señal de salida del método LPC.	65
Figura 32. La gráfica de la izquierda muestra la FFT de la envolvente de voz original (señal de entrada) y la gráfica de la derecha muestra la FFT de la envolvente de la señal de salida del método LPC.	65
Figura 33. Espectro en frecuencia de la nota Do y la nota Re, respectivamente.	69
Figura 34. Espectro en frecuencia de la nota Sol y la nota La, respectivamente.	69
Figura 35. FFT de señal de entrada y su envolvente. La señal corresponde a la vocal "A".	70
Figura 36. FFT de señal de entrada y su envolvente. La señal corresponde a la vocal "E".	70
Figura 37. Diagrama de bloques de modulación AM realizada entre la voz e instrumento.	71
Figura 38. Diagrama de flujos de grabación del sonido para las encuestas.	73
Figura 39. La gráfica superior izquierda muestra la FFT de la señal de voz original (señal de entrada), la gráfica superior derecha muestra la FFT de la señal de salida del banco de filtros, la gráfica inferior izquierda muestra la señal de salida del método Spectrum y la gráfica inferior derecha muestra la señal de salida del método LPC.	79
Figura 40. La gráfica superior izquierda muestra la FFT de la envolvente de voz original (señal de entrada), la gráfica superior derecha muestra la FFT de la envolvente de la señal de salida del banco de filtros, la gráfica inferior izquierda muestra la FFT de la envolvente de la señal de salida del método Spectrum y la gráfica inferior derecha muestra la FFT de la envolvente de la señal de salida del método LPC.	81
Figura 41. Gráfica de distribución de frecuencias.	83
Figura 42. Caracterización del Audio 1.	84
Figura 43. Gráfica de distribución de frecuencias.	85
Figura 44. Caracterización del Audio 2.	86
Figura 45. Gráfica de distribución de frecuencias.	87
Figura 46. Caracterización de la señal (Audio 1).	88
Figura 47. Gráfica de distribución de frecuencias.	89
Figura 48. Caracterización de la señal (Audio 2).	90
Figura 49. Gráfica de distribución de frecuencias.	91
Figura 50. Caracterización de la señal (Audio 1).	92
Figura 51. Gráfica de distribución de frecuencias.	93
Figura 52. Caracterización de la señal (Audio 2).	94

Capítulo 1

Introducción

Existen diferentes métodos que permiten reconstruir las señales de voz; tres de ellos son: Banco de filtros, Spectrum y LPC. El método de banco de filtros consiste en separar la señal de audio original con filtros pasa banda, en varias señales que reciben el nombre de “sub-banda” y cada una contiene la información de una banda de frecuencia, luego decima cada señal sub-banda para sumarlas, obteniendo una representación de la señal original. (Madrid, 2006); el método Spectrum implementa FFT y adquiere valores numéricos de la onda acústica a intervalos discretos, pero suficientemente próximos en el tiempo, que permiten analizar y recomponer posteriormente una señal (Zolzer, 2002). Por otro lado, el método de predicción lineal o LPC, no codifica la forma de onda en sí, sino las características del modelo de producción de la voz. LPC es un modelo matemático que se aproxima a los modelos biológicos y acústicos reales, por ende, no es una representación precisa. Sin embargo, a través de este método es posible obtener una adecuada representación de la señal de la voz (Sinha, 2010). Con los métodos mencionados se pretende reconstruir la señal de la voz y usarla como señal moduladora en un proceso de modulación AM, con la finalidad de obtener un nuevo sonido en donde la señal generada por un instrumento musical se vea modificada en amplitud a partir de la voz humana.

El alcance general de este proyecto es determinar cuál de estos métodos puede ser el apropiado de acuerdo a ciertas características (Consumo de memoria RAM, latencia y comparación espectral de las señales de entrada y salida a cada método, las cuales serán descritas a profundidad más adelante), para el desarrollo de un dispositivo que obtenga un nuevo sonido a partir de la modulación AM entre la voz y un instrumento musical, como ya fue mencionado. El

resultado es una señal de instrumento que varía en amplitud con la variación de la señal de la voz; luego se pretende definir si el sonido obtenido resulta agradable en la medida que pueda ser utilizado en la producción de una canción o el diseño sonoro de un videojuego o una producción audiovisual; esto será definido mediante una encuesta realizada a una población de músicos y estudiantes de ingeniería de sonido.

Los resultados más relevantes obtenidos en este proyecto mostraron que el método con mejores resultados en la caracterización de consumo de memoria RAM, latencia y comparación espectral es el banco de filtros y que el sonido obtenido de la modulación AM entre la voz y el instrumento ha resultado agradable a la población encuestada. El desarrollo de las mediciones y las encuestas será descrito detalladamente a lo largo de este documento que busca contribuir a la comprensión y estudio de dispositivos pensados para el procesamiento de señales de audio a nivel universitario.

Capítulo 2

Planteamiento del problema

Antecedentes

El documento Design and applications of embedded systems for speech processing (Jhing-Fa Wang, Po-Chun Lin y Bo-Weu Chense, 2012), se centra en las técnicas de procesamiento del habla que implican la verificación y recuperación de las formantes del habla humana. Menciona también varios problemas del diseño de hardware, incluyendo algoritmos con sus correspondientes soluciones mejoradas a los sistemas integrados. En este proyecto los autores hicieron uso del modelo de los coeficientes Cepstrales de las frecuencias de Mel (MFCC). El documento se centra en el diseño de chips que llevan a cabo la extracción de las características del habla. Por último, se considera una tarjeta computacional FPGA (Field Programmable Gate Array) como caso de estudio para realizar en diseño. Aunque se hace una caracterización del modelo de MFCC (Mel Frequency Cepstral Coefficient) no se compara este método con ningún otro para determinar si existe otra opción que mejore la velocidad con la que se realiza el procesamiento de voz. Finalmente, se llega a la conclusión de que un diseño adecuado de hardware puede llevar a la mejora del sistema.

En “A mixed excitation LPC (Linear Predictive Coding) vocoder model for low bit rate speech coding” (Cree A. V, 2008). LPC se presenta como una de las maneras más eficientes de codificar el habla humana, sin embargo el sonido sintético y los ruidos tonales que se generan mediante el proceso pueden ser molestos, y estos problemas aumentan si hay ruido de fondo presente en la entrada de la voz. En este trabajo se presenta un nuevo modelo mixto de excitación LPC, que conserva la tasa baja de bits, pero añade parámetros de señal de excitación libre de

manera que el sintetizador pueda imitar características más naturales de la voz. A pesar de los logros obtenidos, en este trabajo tampoco se establece una caracterización del modelo LPC en cuanto a velocidad de procesamiento, ni es mencionado el hardware implementado (McCree, 2008). En los antecedentes encontrados, no existe una caracterización de los diferentes modelos de codificación existentes, aunque en todos se pretenda llegar a una optimización en el procesamiento de la voz.

Descripción y formulación del problema

Debido a que no se ha encontrado una caracterización de los métodos usados para reconstruir la señal de la voz, se pretende hacer dicha caracterización y determinar cuál sería el método ideal para realizar tareas en donde se requiera optimizar recursos durante la adquisición de la señal de voz, de manera que el procesamiento en el sistema embebido sea más rápido, pero sin perder la calidad de la señal. En este caso, la tarea a realizar es la modulación en amplitud de un instrumento musical teniendo la voz humana como señal moduladora.

Para conocer el método que obtiene la envolvente de la voz con mayor semejanza a la original y menor cantidad de tiempo de procesamiento, es necesario hacer un análisis comparativo de cada uno, por medio de mediciones de latencia de los algoritmos, su consumo de memoria RAM y una comparación espectral entre las envolventes obtenidas y la envolvente original. Se espera obtener una reconstrucción de la señal de voz diferente para cada método de manera que sea posible determinar las diferencias entre ellas.

Pregunta problema

¿Cuál es el método de reconstrucción de voz ideal para llevar a cabo la implementación de un dispositivo en un sistema embebido, que halle la envolvente de la voz y con esta module la

amplitud de la señal de un instrumento musical, con la finalidad de obtener un nuevo sonido que pueda ser usado en la producción musical o audiovisual?

Justificación

Este proyecto busca aportar una caracterización, hasta el momento desconocida, de algunos métodos o algoritmos que son utilizados para reconstruir la señal de la voz y posteriormente hallar su envolvente. Con esta caracterización se elegirá el método que de mejores resultados en mediciones de tiempo de procesamiento en el sistema embebido y comparaciones espectrales. El sistema embebido junto con el método elegido conformará un sistema “instrumento/procesador” que usará la envolvente de la voz para modular la amplitud de una señal de un instrumento musical con la finalidad de obtener un nuevo sonido a partir de la voz.

La caracterización determinará cuál es el tiempo de latencia, el consumo de memoria RAM y la similitud de cada envolvente con la original, para los tres métodos, LPC, Spectrum y Banco de Filtros y definirá si el sonido obtenido con el método que resulte más eficiente, realmente puede tener una acogida en una población de personas dedicadas a la producción musical, dependiendo de si el sonido obtenido por el sistema realmente resulta inteligible y natural. Además aporta la posibilidad de comprender la forma de desarrollar un sistema que sirva para la manipulación de señales sonoras con la finalidad de obtener sonidos que puedan ser utilizados en procesos creativos dentro de la producción musical.

Objetivos

Objetivo general

- Desarrollar un dispositivo por medio de un sistema embebido, que implemente el mejor método obtenido del análisis comparativo entre LPC, Spectrum y Banco de filtros con el fin de hallar la envolvente de la voz y modular con ésta la señal de un instrumento musical.

Objetivos específicos

- Implementar los métodos de análisis de voz y realizar una caracterización para determinar el tiempo de latencia de cada algoritmo, número de bloques de código necesarios en la implementación de cada método, consumo de memoria RAM y similitud de las envolventes obtenidas con la envolvente original mediante una comparación espectral.

- Realizar la modulación de un instrumento musical mediante la envolvente de la voz, implementando en un sistema embebido el método que presente mejores resultados en los análisis de tiempo de latencia, consumo de memoria RAM, número de bloques y comparación de la envolvente obtenida con la envolvente original.

- Realizar una encuesta que permita medir el grado de aceptación del sistema en una población de músicos de la ciudad de Bogotá, mediante la valoración subjetiva de la inteligibilidad de la palabra, calidez y brillo del sonido y finalmente, si cada una de esas características le resultan agradables al individuo.

Alcances y limitaciones

Alcances

- Se realizará una caracterización de tres métodos que se utilizan para hallar la envolvente de la voz.

- El dispositivo planteado logrará la modulación de la amplitud de una señal de instrumento musical, en tiempo real.
- El dispositivo final implementará el método que obtenga mejores resultados en el análisis comparativo que determina el tiempo de procesamiento de los métodos y la semejanza de la envolvente obtenida con la original.
- Se realizará una valoración subjetiva y cualitativa del grado de aceptación del sistema en una población de músicos.

Limitaciones

- No se establecerá relación entre otros métodos de análisis de la envolvente de la voz diferentes a LPC, Banco de filtros y Spectrum.
- Las entradas y salidas de audio y demás limitaciones del sistema embebido que va a ser utilizado en el desarrollo del dispositivo.

Capítulo 3

Marco de referencia

Marco teórico

Criterios de adquisición.

Un sistema de adquisición de señales permite convertir una señal proveniente de un fenómeno físico en una señal eléctrica, susceptible de ser procesada posteriormente. Para llevar a cabo esta adquisición, primero es necesario un transductor y un acondicionamiento de la señal. El transductor es el elemento encargado de convertir un tipo de energía en otro, por ejemplo un micrófono, el cual convierte energía acústica en una señal eléctrica; sus parámetros básicos son margen de medida, sensibilidad, resolución y exactitud. Luego de esto, la señal debe ser acondicionada para adaptarla a la entrada del convertidor A/D (Análogo/Digital), en esta etapa se adapta el rango de tensión de salida del transductor al rango de entrada del convertidor o se acopla la impedancia de salida de uno con la impedancia del otro. Después es necesaria una conversión analógica-digital, es decir que la señal pasa de ser continua en cuanto a amplitud y en el tiempo a ser discreta, en donde la amplitud y el tiempo alcanzan un rango discreto de valores. Sin embargo, muchos de estos sistemas cuentan con entradas digitales además de entradas analógicas. Pallás Areni (1993) afirma:

La conversión de señales analógicas en señales digitales se realiza en dos etapas: primero se cuantifica la señal y luego se codifica. La cuantificación consiste en representar la amplitud de la señal mediante un número finito de valores distintos en instantes determinados; si el convertidor es de n bits, hay 2^n valores posibles; por otro lado, la codificación es la representación del valor asignado a la señal, mediante combinaciones de

símbolos que se representan habitualmente con dos niveles de tensión. La cuantificación y codificación las realiza el convertidor analógico-digital. (p.5)

Una vez se han adquirido los datos se visualizan, analizan y almacenan en un ordenador. Todos estos procesos se pueden desarrollar mediante lenguajes de programación. Finalmente, viene la etapa de salida, cuya función es obtener señales analógicas a partir del código de salida que da el procesador. Sin embargo ni los dispositivos de visualización, ni los medios de transmisión son directamente compatibles con las señales obtenidas en la etapa de salida, y por ello hay que acondicionarlas. En particular, los niveles de tensión y corriente, y la potencia disponible, suelen ser insuficientes (Pallás Areny, 1993).

La primera parte del hardware de adquisición de una señal de voz es el micrófono, que es acondicionado para funcionar como entrada al circuito de pre-amplificación y filtrado. Idealmente debe poseer una respuesta en frecuencia adecuada dentro del rango de frecuencias de 100 Hz a 4000 Hz, para asegurar un correcto funcionamiento para la voz. Luego, es necesario tener un dispositivo de captura de datos que puede ser una tarjeta de adquisición de señal de tipo clásico o un sistema de procesamiento de señal con etapa de entrada/salida analógica incorporada. El dispositivo debe contar con conversores A/D Y D/A y entradas y salidas de audio (Pallás Areny, 1993).

La información adquirida, directamente o tras el procesamiento de las señales, se emplea a veces para actuar sobre el proceso o sistema de donde provienen; además, es frecuente presentarla, almacenarla o ambas cosas a la vez. Dichas funciones se suelen realizar directamente con las señales digitales, sin convertirlas de nuevo en analógicas. Para llevar a cabo un pre procesamiento de la señal de voz que permita hacer una representación apta de ella con la información obtenida de un muestro previo es necesario un software que permita el desarrollo de

algoritmos que realizan una adquisición de datos, conversión de datos al formato interno, pre enfatización del vector de muestras, división de muestras en tramas adecuadas para un análisis de la señal y extracción de características de la misma (Pallás Areny, 1993).

Linear Predictive Coding (LPC).

Se define como un método digital para la codificación de una señal analógica en la que un valor particular, se conoce previamente por una función lineal de los valores pasados de la señal; este se basa en una aproximación matemática del tracto vocal representada por un tubo de diámetro variable. El método LPC, realiza una suma lineal de muestras tempranas, usando un filtro digital. Los filtros intentan predecir valores futuros de la señal de entrada, basados en señales pasadas. Luego se divide la señal de voz en bloques y se determina la señal de entrada y a partir de eso el método determina los coeficientes del filtro que será utilizado para reproducir el bloque actual del discurso. Esta información es cuantificada y transmitida. En la decodificación, LPC reconstruye el filtro basado en los coeficientes recibidos. La información adicional de la señal de voz original es utilizada por el decodificador para determinar la señal de entrada o de excitación que se envía (Bradbury, 2000).

LPC tiene dos componentes principales: el análisis y la codificación, y la síntesis y la decodificación. La parte de análisis implica la separación de la señal de voz en segmentos y su descomposición en segmentos o bloques, esta descomposición se hace para conocer si es el segmento es sordo o no y a partir de ahí, saber que filtro será necesario para modelar el segmento actual. El análisis se realiza normalmente por un remitente que responde a estas preguntas y por lo general transmite las respuestas en un receptor. El receptor realiza la síntesis de LPC mediante el uso de las respuestas recibidas para construir un filtro que cuando se proporciona la fuente de

entrada correcta será capaz de reproducir con precisión la señal de voz original. Bradbury (2000) afirma: “En esencia, la síntesis LPC intenta imitar la producción del habla humana” (p.19).

Spectrum.

Método que implementa la transformada rápida de Fourier y adquiere valores numéricos de la onda acústica a intervalos discretos, pero suficientemente próximos en el tiempo, que permiten analizar y recomponer posteriormente una señal.

Transformada y Anti-transformada de Fourier

En términos matemáticos, “la transformada de Fourier es un operador (F) que aplicado a una función temporal $g(t)$ la convierte en otra función de la frecuencia $G(f)$ que aporta la misma información que la primera” (Martí, 1987, p.235):

$$F[g(t)] = G(f) \qquad \text{Ecuación 1}$$

“La función $G(f)$ informa de todas las periodicidades de la función $g(t)$. Así mismo, existe también el operador anti-transformada (F^{-1}) que aplicado a una función frecuencial $G(f)$ permite recomponer su función temporal original $g(t)$ ” (Martí,1987, p.236):

$$F^{-1}[G(f)] = g(t) \qquad \text{Ecuación 2}$$

El algoritmo llamado transformada rápida de Fourier, conocido por las iniciales FFT, ha facilitado la rapidez de este cálculo para hacerlo en tiempo real, de manera que sea posible visualizar la forma de la función frecuencial a medida que se produce la evolución del correspondiente parámetro (Martí, 1987).

Banco de filtros.

Un banco de filtros es una colección de filtros digitales con una entrada común o una salida común. El cálculo del espectro de una señal de entrada variable en el tiempo, puede ser

interpretado como un banco paralelo de N filtros con respuestas al impulso y transformadas de Fourier dadas por:

$$hk(n) = h(n)WN^{-nk}, \quad k = 0, 1, \dots, N - 1$$

Ecuación 3

$$Hk(e^{-i\omega t}) = H(e^{i(\omega t - \omega t)}) \quad \omega t = \frac{2\pi}{N}k$$

Ecuación 4

Cada señal pasa banda $y_k(n)$ se obtiene filtrando la señal de entrada $x(n)$ con el filtro pasa banda $h_k(n)$ correspondiente.

El banco de filtros se encarga de reconstruir la señal perfectamente o casi perfectamente. La reconstrucción de la señal se puede evaluar mediante el examen de las distorsiones en el banco de filtros. Hay tres distorsiones fundamentales en los bancos de filtros: Aliasing, distorsión de magnitud y distorsión de fase. Si un banco de filtros está libre de estas distorsiones, este hará una reconstrucción perfecta (PR). Si estas distorsiones son pequeñas el banco será casi perfecto (NPR). Si el banco de filtros hace una reconstrucción perfecta, la señal de salida será una versión escalada y retardada de la señal de entrada (Zolzer, 2002).

Detección de envolvente.

La envolvente de una señal es equivalente a su contorno. Si se tomaran todos los picos de la señal en el dominio del tiempo y se unieran con una línea, se obtendría la envolvente. Para este proyecto se utiliza un método de detección de envolvente que recibe el nombre "Detector de ley cuadrática". Este método asume la señal de entrada como una señal modulada y la envolvente como una señal moduladora de baja frecuencia y que se puede obtener con un filtro pasa-bajo. Este método⁵ implica elevar al cuadrado la señal de entrada para luego filtrarla; la mitad de la energía de la señal se concentra en las frecuencias altas y la otra mitad se desplaza hacia abajo,

hacia DC, a continuación se reduce la resolución de la señal para disminuir la frecuencia de muestreo siempre y cuando la señal no tenga frecuencias altas que puedan causar solapamiento; de ser así, se debe aplicar un filtro paso bajo antes de disminuir la resolución de la señal. Finalmente se pasa la señal por un filtro pasa-bajos de fase mínima para eliminar la energía de alta frecuencia y obtener la envolvente de la señal.

Para mantener la escala correcta, se debe amplificar la señal por un factor de dos pues se tiene solo la mitad inferior de la energía de la señal y a la salida del filtro de fase mínima se halla la raíz cuadrada de la señal para revertir la distorsión de escala causada por elevarla al cuadrado. A continuación se muestra el proceso de detección de ley cuadrática.

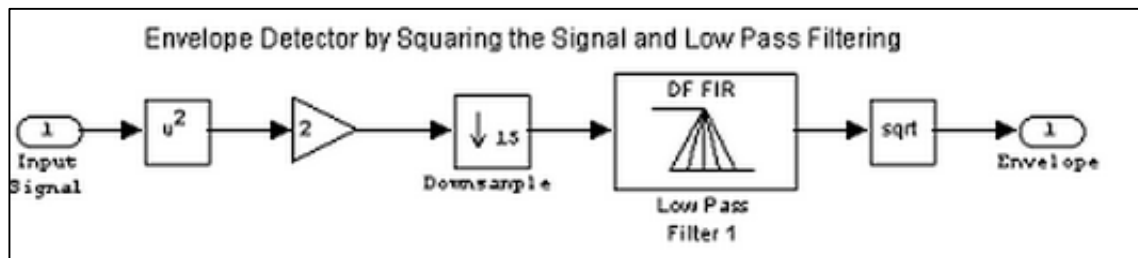


Figura 1. Método de detección de envolvente de ley cuadrática.
Fuente: www.mathworks.com/help/dsp/examples/envelope-detection.

Modulación AM.

En la modulación AM la amplitud de una señal portadora varía de acuerdo a la amplitud de una señal moduladora. En este proyecto, la señal portadora es la señal de un instrumento musical, más específicamente, un sintetizador, y la señal moduladora es la voz humana. De acuerdo a esto, la finalidad es que la amplitud de la señal del sintetizador varíe de acuerdo a la amplitud de la señal de la voz.

En una descripción matemática dada por Fernández (2002) se define que “una señal portadora puede ser una señal sinusoidal dada por la siguiente ecuación, donde A_c es la amplitud, f_c es la frecuencia y se asume por conveniencia que la fase es cero” (p.3).

$$c(t) = A_c \cos(2\pi f_c t) \quad \text{Ecuación 5}$$

De acuerdo con lo anterior y siendo $m(t)$ la señal banda base o señal moduladora que contiene la información, la señal $c(t)$ es independiente de $m(t)$. Teniendo esto en cuenta, la modulación AM, o modulación de amplitud, se define como el proceso en el cual la amplitud de la portadora $c(t)$ varía en torno a un valor medio de forma lineal con la señal de banda base $m(t)$ según la siguiente ecuación, donde K_a es una constante denominada sensibilidad en amplitud del modulador (Fernández, 2002).

$$s(t) = A_c [1 + K_a m(t)] \cos(2\pi f_c t) \quad \text{Ecuación 6}$$

Si se supone que A_c es igual a la unidad y $m(t)$ es una señal sinusoidal, como la de la siguiente imagen,

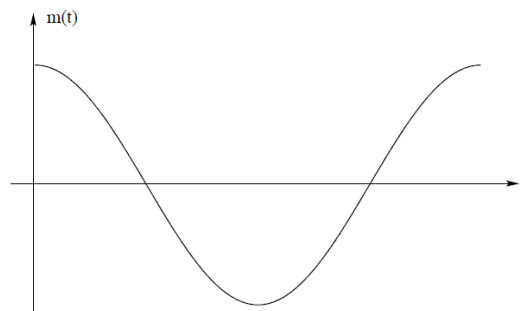


Figura 2. Señal moduladora $m(t)$.

Fernández, Marcos Martín (2002). *Modulación en amplitud*. [Figura]. Universidad de Valladolid

Se pueden dar dos casos:

- Si $|K_a m(t)| < 1$ se tiene la siguiente señal:

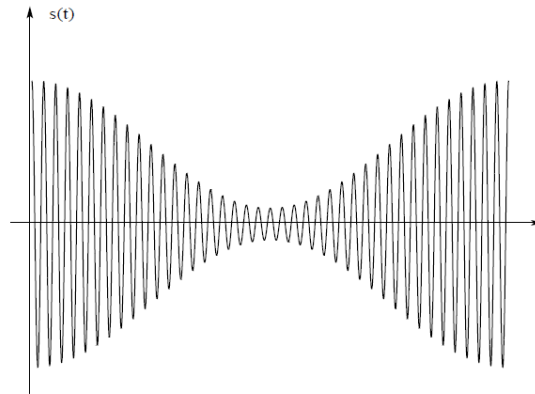


Figura 3. Señal modulada sin sobremodulación.

Fernández, Marcos Martín (2002). Modulación en amplitud. [Figura]. Universidad de Valladolid

- Si $|Ka m(t)| > 1$ se tiene la siguiente señal:

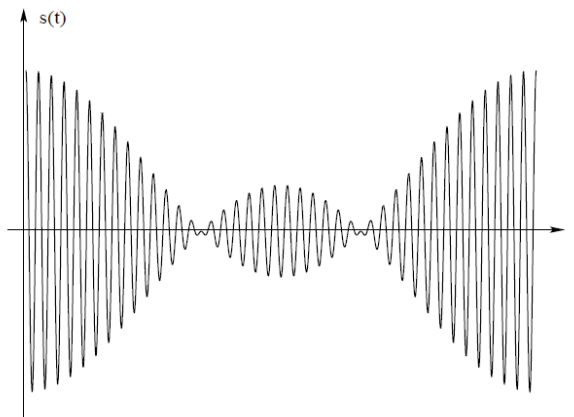


Figura 4. Señal modulada con sobremodulación.

Fernández, Marcos Martín (2002). Modulación en amplitud. [Figura]. Universidad de Valladolid

Para que la envolvente de la señal modulada siga la forma de la señal banda base $m(t)$ se debe satisfacer que $|Ka m(t)| < 1$. Esto asegura que $1 + Ka m(t)$ es siempre positivo y es posible expresar la envolvente de la señal $s(t)$ como $Ac [1 + Ka m(t)]$. Cuando $|Ka m(t)| > 1$ debido a que Ka es demasiado grande, la señal AM se sobremodula, entonces la fase de la señal AM se invierte siempre que $1 + Ka m(t)$ cambia de signo, lo que da lugar a una distorsión en la envolvente (Fernández, 2002).

En segundo lugar, Fernández (2002) establece que: “Se debe satisfacer que la frecuencia de la portadora f_c sea mucho mayor que la componente frecuencial superior de $m(t)$. Si esto no se satisface, la envolvente no seguirá la señal moduladora” (p. 5). Por ende, si se supone W como el ancho de banda de $m(t)$, se debe cumplir que:

$$f_c \gg W$$

Ecuación 7

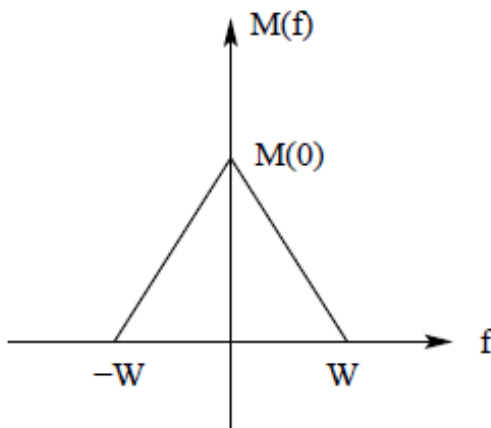


Figura 5. Espectro de la señal moduladora.
 Fernández, Marcos Martín (2002). Modulación en amplitud. [Figura]. Universidad de Valladolid

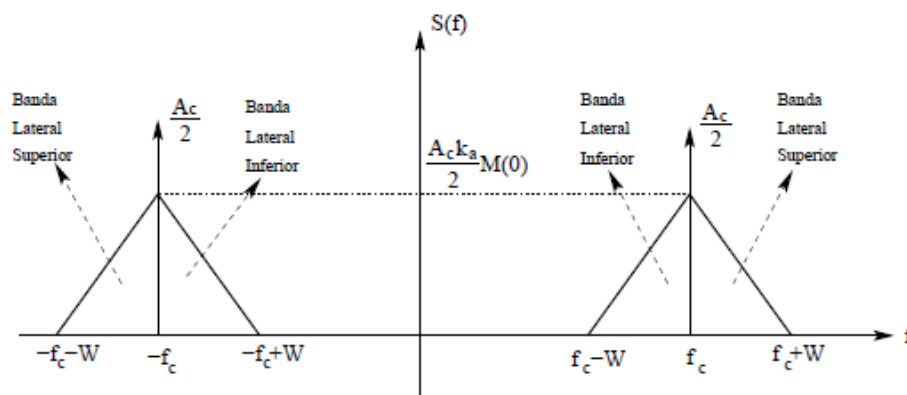


Figura 6. Espectro de la señal modulada.
 Fernández, Marcos Martín (2002). Modulación en amplitud. [Figura]. Universidad de Valladolid

De la Figura 6 se destaca que para frecuencias positivas la parte del espectro por encima de f_c y para frecuencias negativas la parte del espectro por debajo de $-f_c$ se denomina banda lateral superior y para frecuencias positivas la parte del espectro por debajo de f_c y para

frecuencias negativas la parte del espectro por encima de $-fc$ se denomina banda lateral inferior. La condición $fc > W$ asegura que las bandas laterales inferiores no se solapen. También se destaca que para frecuencias positivas, la componen frecuencial superior es $fc+W$ y la inferior $fc-W$. La diferencia entre ambas define el ancho de banda de transmisión de la señal AM (Fernández, 2002).

Modulación AM con portadora suprimida o de doble banda lateral DSB.

Como fue mencionado antes, la señal portadora $c(t)$ es completamente independiente de la información de la señal $m(t)$, por lo tanto transmitir la portadora significa un desperdicio de potencia (Fernandez, 2002). La componente portadora de la señal modulada se puede suprimir, dando lugar a una modulación doble banda lateral con portadora suprimida. Entonces suprimiendo la portadora se tiene una señal que será proporcional al producto de la portadora por la señal banda base según la siguiente ecuación:

$$s(t) = c(t) * m(t) = Ac \cos(2\pi fct) m(t) \quad \text{Ecuación 8}$$

La señal así modulada presenta un cambio de fase siempre que la señal $m(t)$ cruce por cero. Ahora a diferencia del caso AM, para DSB la envolvente de la señal no sigue a la señal moduladora $m(t)$ (Fernandez, 2002).

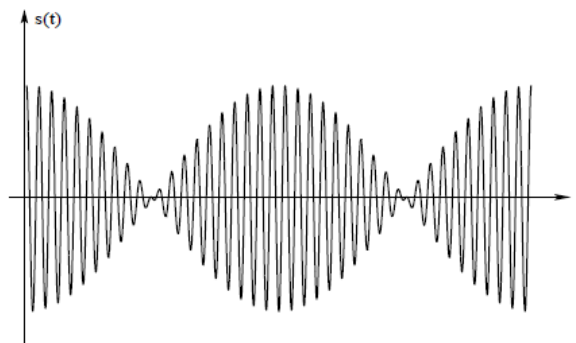


Figura 7. Señal modulada DSB.

Sistema embebido.

Un sistema embebido es un sistema electrónico que funciona como un ordenador que es diseñado para realizar tareas específicas, entre ellas, el procesamiento de señales de audio mediante su captura a través de transductores y conversores que llevan la señal al dominio digital y una vez allí, esta puede ser procesada por uno o varios algoritmos que generen diferentes resultados. Implementa un procesador digital o varios procesadores digitales, dependiendo de su aplicación. Estos procesadores pueden ser FPGA, Microprocesadores, Micro-controladores o DSP (Digital signal processing). El tipo de procesador es una característica fundamental de cualquier sistema embebido junto con el software de programación (Lipiansky, 2012).

Un sistema embebido tiene un sistema informático integrado, pero a pesar de eso no se considera un computador. Existen muchos ejemplos de sistemas embebidos, estos incluyen teléfonos celulares, reproductores de música, GPS, etc. No hay una definición clara para sistema embebido, que tan grandes o rápidos son o donde residen. Sin embargo, todos estos sistemas tienen características en común. Según Lipiansky (2012): “Un sistema informático y su hardware asociado son una parte integral del sistema, sin necesidad del que el sistema embebido sea precisamente un sistema informático, como un ordenador personal de propósito general” (p.15). Todos los sistemas de cómputo tienen cuatro grandes bloques: El cerebro de la máquina, llamado CPU; los dispositivos de memoria que contienen las instrucciones del programa, los datos de entrada y los datos generados por las aplicaciones; los dispositivos de entrada; y los dispositivos de salida. Estos permiten al sistema ser conectados con el mundo exterior. Por otro lado, hay un bus de direcciones, un bus de datos y un bus de control para interconectar los bloques principales.

El hardware embebido consta de todos los componentes electrónicos, como circuitos integrados, inductores, condensadores, resistencias, etc. El hardware también comprende componentes mecánicos o electromecánicos tales como la presión, la temperatura y sensores de flujo de aire; relés; actuadores; pantallas y luces de advertencia. Además, incluye el chasis del sistema cuando es aplicable, o una placa de circuito impreso (Lipiansky, 2012).

Memoria RAM.

Una de las partes más importantes de un ordenador es la memoria RAM. Para que un programa pueda ser ejecutado, antes tiene que estar cargado en memoria RAM, utilizándose también esta memoria para guardar los datos que utiliza el programa. Dos características importantes de una memoria RAM son su capacidad y su velocidad de acceso. Los chips de memoria RAM necesitan un tiempo mínimo para identificar exactamente la información que se pretende leer, tiempo de acceso, y otro tiempo para transvasar esa información al lugar de destino, tiempo de carga. Pues bien, el tiempo total requerido para efectuar toda la operación se le llama tiempo de un ciclo de memoria, y es el resultado de sumar los dos tiempo anteriores (Caballar, 1994).

$\text{Ciclo de memoria} = \text{tiempo de acceso} + \text{tiempo de carga}.$

Es importante definir que hay diferentes tipos de memoria RAM, entre ellos la memoria caché que son unos chips de memoria RAM, llamados SRAM o RAM estáticos, que son mucho más rápidos (y mucho más caros) que los chips DRAM o RAM dinámicos, que son los que se usan generalmente para formar lo que se llama habitualmente memoria RAM. Estos chips se colocan entre el microprocesador y los chips DRAM, de forma que cuando el microprocesador necesita información de la memoria la coge de los chips SRAM sin necesidad de tiempos de espera (Caballar, 1994). Los sistemas embebidos tenidos en cuenta para la elección del sistema

final que será utilizado en el desarrollo del dispositivo, cuentan con estos dos tipos de memoria RAM.

Procesamiento digital de señales.

En el procesamiento digital de señales se tiene un sistema digital programable que permite flexibilidad a la hora de reconfigurar las operaciones de procesado digital de señales sin más que cambiar el programa. Un sistema digital permite un mejor control de los requisitos de precisión. Tales requisitos, a su vez, resultan en la especificación de requisitos en la precisión del conversor A/D y del procesador digital de señales, en términos de longitud de palabra, aritmética de coma flotante frente a coma fija y factores similares. Las señales digitales se almacenan fácilmente en soporte magnético sin deterioro o pérdida en la fidelidad de la señal, aparte de la introducida en la conversión A/D. Como consecuencia, las señales se hacen transportables y pueden procesarse en tiempo no real en un laboratorio remoto. El método de procesado digital de señales también posibilita la implementación de algoritmos de procesado de señal más sofisticados (Proakis, 1998).

Como consecuencia de estas ventajas, el procesado digital de señales se ha aplicado a sistemas prácticos que cubren un amplio rango de disciplinas, por ejemplo, la aplicación de técnicas de procesado digital de señales de voz. Sin embargo, el procesamiento digital tiene sus limitaciones, como la velocidad de operación de los conversores A/D. Las señales con anchos de banda extremadamente grandes, precisan conversores A/D con una velocidad de muestro alta y procesadores digitales de señales rápidos. Así, existen señales analógicas con grandes anchos de banda para las que la solución mediante procesado digital de señales se encuentra más allá del “estado del arte” del hardware digital (Proakis, 1998).

Latencia

Se denomina latencia a una suma de retardos temporales en un sistema. En un dispositivo de procesamiento de señales, como es el caso, que tiene una memoria RAM y un procesador, la latencia puede ser el tiempo en que el procesador tarda en adquirir los datos y ejecutar las instrucciones que se encuentran en la memoria. Todos estos retardos temporales que pueden tener que ver con la cantidad de datos, los buffers de almacenamiento, etc. Afectan el tiempo de ejecución de un algoritmo. En este orden de ideas, se asume q la latencia y el tiempo de ejecución están directamente relacionados, y para conocer los retardos temporales generados por cada uno de los métodos, se mide su tiempo de ejecución.

Marco legal o normativo**Aspectos legales del software libre o de código abierto (Open Source)**

Según Rios (2004): “De manera ordinaria y mundialmente aceptada, se tiene que los programas de Computación o de ordenador, también denominados Soporte lógico (Software), son creaciones protegidas por la propiedad intelectual, y de manera más precisa por la disciplina del derecho de autor, en los mismos términos de las obras del dominio literario” (p.2).

También afirma Rios (2004): “En 1985 nace la FREE SOFTWARE FOUNDATION (FSF) como una organización sin ánimo de lucro para el desarrollo del software libre basado principalmente en el uso de EMACS y en el proyecto GNU. Los recursos de la FSF se basan en la recepción de donaciones y contribuciones, pero principalmente por los dineros que se logran por la venta de copias de software libre y todos los demás servicios relacionados” (p.4).

“Cuando se utiliza la expresión Software Libre como traducción de la palabra original en idioma Ingles “Free Software”, nótese que estamos denotando que nos encontramos frente a un software libre en cuanto a su utilización, y cuyo código fuente es abierto (Open Sorce), y no

necesariamente a software que sea en estricto sentido gratuito o no comercial. Un software libre puede perfectamente ser diseñado para fines comerciales y ello no desnaturaliza su propósito.

Las prerrogativas y libertades que conlleva la expresión “Software Libre”, hacen alusión a las facultades que tiene el usuario de este tipo de herramientas, para usarlas, estudiarlas, modificarlas, mejorarlas, adaptarlas, rediseñarlas y redistribuirlas; con la única condición de que luego de realizar cualquiera de las anteriores conductas, no introduzca ninguna restricción al producto así obtenido, adicionado, modificado, etc” (Rios,R. Wilson, 2004, p.6).

Capítulo 4

Metodología

Enfoque de investigación

El enfoque de este proyecto de grado es empírico analítico, ya que se pretende caracterizar los métodos implementados en la reconstrucción de la señal de la voz y el posterior hallazgo de su envolvente, con la finalidad de probar que el desarrollo de un dispositivo en un embebido, que permita obtener la envolvente de la voz para usarla en la modulación de un instrumento musical es posible a través del método que obtiene la envolvente con mayor semejanza a la original y menor cantidad de tiempo de procesamiento. Además, se pretende caracterizar el sonido obtenido con esta modulación entre la voz y el instrumento de manera subjetiva mediante encuestas realizadas a una población de músicos.

Línea de investigación de la universidad/ línea de investigación de la facultad/ núcleos problemáticos

El proyecto pertenece a la línea de investigación de la universidad: Tecnologías actuales y sociedad, ya que se pretende construir un sistema que puede ser útil para una población de músicos y personas que trabajen en producción musical. La línea de investigación de la facultad y el núcleo problemático pertenecen a la rama de análisis y procesamiento de señales.

Hipótesis

La implementación de un dispositivo en un sistema embebido, que permita obtener la envolvente de la voz para usarla en la modulación de un instrumento musical, es posible a través del método que reconstruya la señal de la voz y obtenga su envolvente con mayor semejanza a la original y menor cantidad de tiempo de procesamiento, siendo este determinado mediante un

análisis comparativo de los métodos LPC, Spectrum y banco de filtros y el sonido obtenido mediante esta modulación puede alcanzar un alto grado de aceptación en una población de músicos e ingenieros de sonido, de manera que estos determinen que usarían el sonido obtenido durante un proceso creativo dentro de su profesión.

Variables

- Métodos de medición para realizar el análisis comparativo de los métodos utilizados para hallar la envolvente de la voz.
- Algoritmos de cada método (LPC, Spectrum, Banco de filtros) implementados en el software de programación que será compilado con el sistema embebido.
- Variables para la caracterización subjetiva (mediante las encuestas) del sonido obtenido con la modulación AM: La inteligibilidad de la palabra que se refiere al correcto entendimiento del mensaje que se está transmitiendo, la calidez que se refiere a una presencia predominante de frecuencias bajas y el brillo que se refiere a una presencia predominante de frecuencias altas. Y por último el agrado que genera el sonido sobre el individuo encuestado el cual será determinado preguntando al individuo si usaría el sonido durante algún proceso creativo musical o de diseño sonoro.

Capítulo 5

Desarrollo ingenieril

Antes de realizar la implementación de los algoritmos de reconstrucción de la señal de voz, se determina cual debe ser el sistema embebido para llevar a cabo el desarrollo del dispositivo; esto con la finalidad de elegir el software en donde los algoritmos serán implementados; por lo tanto, en la primera parte del Desarrollo ingenieril se describen las características que se tuvieron en cuenta para elegir el sistema embebido y como se realiza la adquisición de las señales de voz y de instrumento. Luego de esto, se describe la implementación de los algoritmos Banco de Filtros, Spectrum y LPC y se mide el consumo de memoria RAM, latencia y se hace una comparación espectral de sus señales de salida y de entrada.

Elección del sistema embebido

Para llevar a cabo la elección del sistema embebido, se tuvieron en cuenta las siguientes características, considerando tres sistemas de fácil adquisición y bajo costo:

- Memoria RAM.
- Entradas y salidas de audio del embebido.
- Resolución (en bits) de la conversión A/D de las entradas análogas.
- Memoria de almacenamiento del sistema embebido.
- Software para compilar el sistema embebido.

Tabla 1. Características de los sistemas embebidos.

	Raspberry pi B+ de Raspberry	NI myRIO de National instruments	Pandaboard de Texas Instruments
Memoria RAM	512 MB LPDDR de memoria RAM	512 MB DDR3 de memoria RAM	1 GB DDR2 de memoria RAM.
Entradas y salidas de audio	Cuenta con una salida de audio de conector de 3.5mm y un puerto HDMI. Tiene entradas y salidas de video (Conector MIPI CSI y conector RCA)	Tiene una entrada de audio estéreo de 3.5mm. Salidas y entradas análogas MSP y una salida de audio estéreo.	La versión ES Rev B1 cuenta con salidas HDMI v1.3, DVI-D y entrada de audio 3.5mm.
Resolución en bits de la conversión AD de las entradas análogas	La resolución en bits de la conversión AD depende del conversor que se ubique antes de la las entradas digitales del embebido para recibir la señal.	La resolución del convertidor análogo digital presente en las entradas análogas es de 16 bits.	La resolución en bits de la conversión AD depende del conversor que se ubique antes de la las entradas digitales del embebido para recibir la señal.
Memoria de almacenamiento del sistema embebido	El almacenamiento de este sistema embebido de maneja mediante una memoria SD que se puede introducir en el hardware.	El sistema embebido tiene puertos USB 2.0 que se puede usar para almacenamiento de datos.	Micro-SD: admite todo tipo de tarjetas, hasta 64 GB (es donde se almacenará el SO y las aplicaciones).
Software para compilar el sistema embebido	Usa sistemas operativos basados en Linux. Uno de estos sistemas operativos es Raspbian, que es gratuito y se encuentra en la página oficial de Raspberry Pi.	NI myRIO está fabricado para ser compilado con LabVIEW 2012-2013.	PandaBoard funciona con Linux, ya sea con distribuciones tradicionales o con Android.

Memoria RAM.

Para empezar, la memoria es la parte del sistema embebido en la cual se almacenan datos, programas, instrucciones, etc. Y los procesadores son los encargados de escribir y leer esa información. DDR2 hace parte de la familia de memorias RAM y funcionan con 4 bits por ciclo a diferencia de las DDR1 que funcionan con 2 bits por ciclo; a su vez la memoria DDR3 funciona con 8 bits por ciclo. Los voltajes de entrada y salida son 1.8V para DDR2 y 1.5V para DDR3. DDR2 y DDR3 son más rápidas y generan menos consumo de energía que DDR1. Por otro lado, LPDDR es una forma modificada de las memorias DDR SDRAM cuya finalidad es reducir el consumo de energía, por eso su voltaje I/O es 1.8V a diferencia de las DDR1 cuyo voltaje es 2.5V, sin embargo, funciona con 2 bits por ciclo (Lipiansky, Ed, 2012). Según la descripción anterior, la memoria DDR3 por funcionar con 8 bits por ciclo y con voltaje de 1.5V, es más rápida y consume menos energía respecto a las memorias DDR2 y LPDDR.

Entradas y salidas de audio.

Raspberry Pi cuenta con una entrada de audio de 3.5 mm, un puerto HDMI y pines GPIO que pueden ser configurados como entradas y salidas digitales; para utilizar estos pines como entradas o salidas de audio, es necesario utilizar un conversor análogo/digital entre la señal de audio y los pines GPIO. Pandaboard, por otro lado, cuenta con una entrada y una salida de audio de 3.5mm y una salida de audio HDMI al igual que Raspberry Pi; cuenta con conectores JTAG. NI myRIO tiene una entrada y una salida de audio de 3.5mm 4 entradas análogas MXP y 2 salidas análogas MXP, las entradas MXP reciben un voltaje de entrada de 0 a 5V. Estas entradas permiten usar una señal de audio sin necesidad de utilizar un conversor AD externo, pues este ya viene integrado en el NI myRIO, pero si es necesario sumar un voltaje offset a la señal de audio pues esta tiene valores de voltaje positivos y negativos y estas entradas solo reciben señales positivas hasta 5V.

Resolución en bits de la conversión A/D.

En el proceso de conversión A/D de la señal de voz, que considera que la voz tiene componentes frecuenciales hasta 4 kHz y de acuerdo con el Teorema de Nyquist, la frecuencia de muestreo mínima requerida para realizar una captura digital de calidad, debe ser igual al doble de la frecuencia de audio de la señal analógica que se pretenda digitalizar y grabar, por lo tanto la frecuencia de muestro mínima debe ser 8.1 kHz; en el proceso de digitalización de la voz se usa una tasa de 8 bits por muestra durante la cuantización de la señal y esto supone una velocidad de transmisión de 64 kb/s (Jeremy, Bradbury, 2000). Debido a que la calidad del audio digital depende directamente de los sistemas de conversión, para los procesos de muestreo y cuantización de la señal, se requería un sistema embebido capaz de digitalizar la señal a 8100 Hz como frecuencia de muestro y por lo menos 8 bits en el proceso de cuantización. Sin embargo, el rango audible en frecuencias del ser humano va de 20 Hz a 20 kHz, siendo necesaria una frecuencia de muestro de 44100 Hz, para asegurar una mejor calidad en la digitalización de la señal, con cuantificación de 16 bits que equivalen a 65536 pasos o niveles de señal diferentes, para obtener mejor resolución. Cada uno de los sistemas embebidos mencionados cuentan con una entrada de audio de 3.5", y el proyecto requiere dos entradas de audio (Una para la señal de voz y otra para el instrumento), de modo que es necesario ingresar la segunda señal por una de las entradas alternativas. Raspberry Pi cuenta con pines GPIO que pueden ser configurados como entradas y salidas digitales, lo cual haría necesario realizar un conversor A/D adicional al embebido para ingresar la señal análoga. Por otro lado, Pandaboard cuenta con un puerto de expansión que también puede ser configurado para tener una entrada análoga de señal, siendo necesario en conversor A/D adicional. NI myRIO, tiene dos puertos de conectores MXP con

entradas análogas para señales de 0 a 5V, que realiza la conversión A/D con una resolución de 16 bits.

Memoria de almacenamiento.

Según la jerarquía de memoria, los datos que se usan con menos frecuencia y algunos datos del programa en ejecución se almacenan en una memoria de acceso rápido, que son externas al hardware del sistema embebido. El almacenamiento de Raspberry Pi se maneja mediante una memoria SD que se puede introducir en el hardware. NI myRIO tiene puertos USB 2.0 que se puede usar para almacenamiento de datos. Y pandaboard tiene puerto Micro-SD: admite todo tipo de tarjetas, hasta 64 GB (es donde se almacenará el sistema operativo y las aplicaciones).

Entorno de desarrollo.

Para trabajar con Raspberry Pi es necesario instalar en ella un sistema operativo ya sea en Linux o Windows y puede ser controlada mediante un teclado conectado a ella por HDMI o RCA y un ratón conectado por USB o mediante conexión SSH entre el PC y la Raspberry, lo cual permite controlarla desde el computador. Pandaboard, al igual que Raspberry Pi, también está diseñada para trabajar con sistemas operativos como Linux, en un entorno de desarrollo GNU C++, Android, entre otros. Puede ser controlada mediante sus puertos USB y HDMI (para conectar teclado, ratón y pantalla). A su vez, NI myRIO puede funcionar con Linux como sistema operativo, pero cuenta con la inclusión de la versión completa de LabVIEW, lo que permite generar algoritmos programados en LabVIEW que se compilan al sistema embebido al convertir cada proyecto en LabVIEW en una aplicación “Stand Alone”, sin necesidad de instalar un sistema operativo como Linux o similar. Esto contribuye a la optimización de recursos, pues sin la necesidad de instalar un sistema operativo, se genera un ahorro de la memoria del embebido.

LabVIEW es un lenguaje de programación paralela, mientras que C++ y otros lenguajes como MATLAB son de tipo serial o secuencial. En los lenguajes de programación secuenciales, el compilador no detecta automáticamente cada pieza de código que depende de otra para ejecutarse apropiadamente, en contraste con esto, al crear un ejecutable en LabVIEW, el compilador identifica las secciones de código que pueden trabajar de forma paralela, de acuerdo con su diagrama de bloques, permitiendo aprovechar los procesadores. Este aspecto es relevante pues es posible adquirir las señales de audio y procesarlas al mismo tiempo y esta característica es necesaria para realizar el proyecto.

Sistema embebido elegido.

Teniendo en cuenta los aspectos discutidos anteriormente se ha optado por seleccionar el embebido de National Instruments NI myRIO como el sistema para la implementación de los diferentes algoritmos en el proyecto, ya que sus características cumplen con las necesidades del mismo. NI myRIO cuenta con la memoria DDR3, que como se ha mencionado antes, es más rápida y consume menos energía respecto a las memorias DDR2 y LPDDR; cuenta con dos puertos de conectores MXP con entradas análogas para señales de 0 a 5V y una entrada Audio in de 3.5mm así como una salida Audio out de 3.5mm; es posible almacenar datos mediante una memoria externa USB y permite trabajar con LabVIEW que puede trabajar de forma paralela, optimizando recursos computacionales.

Adquisición y procesamiento de señales en el sistema NI myRIO

Para llevar a cabo la adquisición de la señal de voz con NI myRIO, primero es necesario un transductor y un acondicionamiento previo de la señal. El micrófono es el primer eslabón en la cadena de adquisición de la señal, para este caso, el transductor de entrada es un Sennheiser e835, dinámico y cardioide diseñado para la captura de la voz. No era necesario que el micrófono

tuviera alguna característica en especial para este proyecto, sin embargo se realiza un proceso de selección entre dos micrófonos dinámicos que son usados comúnmente para la captura de la voz. En el Anexo A encontrará sus respectivas respuestas en frecuencia y patrones polares.

Tabla 2. Especificaciones técnicas Sennheiser e835 y Shure SM 58.

Especificaciones técnicas	Sennheiser e835	Shure SM58
Respuesta en frecuencia	40-16000 Hz	50-15000 Hz
Sensibilidad (1kHz)	2,8 mV/Pa	1,85 mV/Pa
Impedancia nominal	350 Ω	300 Ω
Patrón polar	Cardioide	Cardioide

Aunque Shure SM58 parece ser más directivo en su patrón polar, según las especificaciones técnicas anteriores el micrófono Sennheiser e835 tiene mayor sensibilidad que el micrófono Shure SM58 y cuenta con una respuesta en frecuencia mucho más plana. Por eso se ha elegido como el transductor de entrada para este proyecto.

NI myRIO cuenta con varias entradas análogas. Una entrada de audio y 8 entradas análogas MXP. Para la señal de voz se usa la entrada Audio In que maneja un rango de -2.5V a 2.5V full scale y es desbalanceada, pero la salida del micrófono es balanceada y por lo tanto es necesario generar un acople entre ambos. Para ello fue elegido un circuito integrado de instrumentación que amplifica y desbalancea la señal de salida del micrófono. Para la señal de instrumento se usa la entrada MXP AIO que tiene un rango de 0V a 5V, es decir, esta entrada solo recibe voltajes positivos, por ende es necesario añadir un voltaje offset a la señal de instrumento, pues esta tiene partes positivas y negativas como toda señal de audio. Esto se hace mediante un circuito divisor de voltaje.

Para la elección del circuito integrado de instrumentación que se ubica entre el micrófono y la entrada Audio In del NI myRIO, se tiene en cuenta que el amplificador debe tener bajo ruido de entrada y de salida y un alto índice de rechazo en modo común (o CMRR) pues un alto índice

de rechazo en modo común elimina el ruido que aparece en la señal positiva o negativa del micrófono y la eliminación de ruido a la entrada y a la salida del integrado mejora la calidad de la señal. También se debe considerar la ecuación de ganancia de cada amplificador y su valor de R_g . Además de estas características se debe considerar el costo de cada integrado y su disponibilidad en el mercado de componentes.

Tabla 3. Especificaciones técnicas de 3 circuitos integrados.

	AD620	INA217	INA128
Ecuación de ganancia	$G = 1 + \frac{49.4 \text{ k}\Omega}{R_g}$	$G = 1 + \frac{10 \text{ k}\Omega}{R_g}$	$G = 1 + \frac{50 \text{ k}\Omega}{R_g}$
Input offset voltaje	15 a 30 μV	50+ (2000/Gain) μV	10+(100/Gain) μV
CMRR			
Ganancia de 1	90 dB	70 a 80 dB	86 dB
Ganancia de 10	110 dB		106 dB
Ganancia de 100	130 dB	100 a 116 dB	125 dB
Ganancia de 1000	130 dB		130 dB
Input voltage noise	9 a 13 $\frac{\text{nV}}{\sqrt{\text{Hz}}}$	1.3 $\frac{\text{nV}}{\sqrt{\text{Hz}}}$ para 1000 Hz	8 $\frac{\text{nV}}{\sqrt{\text{Hz}}}$ para 1000 Hz
Output voltage noise	72 a 100 $\frac{\text{nV}}{\sqrt{\text{Hz}}}$	90 $\frac{\text{nV}}{\sqrt{\text{Hz}}}$	

El integrado AD620 tiene un voltaje offset de entrada que va desde 15 hasta 30 μV , ofrece una relación de rechazo en modo común que va desde los 90 dB para una ganancia de 1, hasta 130 dB para ganancias mayores o iguales a 100. El INA217 tiene un voltaje offset de entrada que depende de la ganancia, ofrece una relación de rechazo en modo común de dB 70 a 80 dB a una ganancia de 1 y de 100 dB a 116 dB para una ganancia de 100. El ruido de entrada es dependiente de la frecuencia. Por último, el INA128, al igual que INA127 tiene un voltaje offset de entrada que depende de la ganancia, tiene una relación de rechazo en modo común de 86 dB para una ganancia de 1 y 130 dB para ganancias iguales o mayores a 1000. El ruido de entrada depende de la frecuencia.

Como la señal de salida del micrófono es balanceada, interesa que el rechazo en modo común sea elevado para eliminar cualquier ruido presente en la señal positiva o negativa del micrófono; como se observa en la tabla 3, el integrado AD620 tiene mayor rechazo en modo común que las otras dos opciones, únicamente en la ganancia de 1000 el AD620 y el INA128 tienen el mismo CMRR. Aunque el ruido de entrada de AD620 no es el más bajo entre las 3 opciones, su ruido de salida sí y finalmente interesa tener una señal de salida limpia. Teniendo en cuenta la discusión anterior, así como su costo y disponibilidad, se ha elegido el integrado AD620 como elemento a usar en la etapa de amplificación de la señal proveniente del micrófono. Su composición interna se puede ver en el Anexo B de este documento.

La ganancia de este integrado está dada por la resistencia externa R_g que se conecta entre los terminales 1 y 8 del integrado. La ecuación de ganancia es la siguiente:

$$G = \frac{49.4k\Omega}{R_g} + 1 \quad \text{Ecuación 9}$$

Para 1K Hz a la entrada del micrófono Sennheiser e835 este tiene una sensibilidad de 2.8m V/Pa, como se puede ver en el Anexo L. Sin embargo, este valor es medido a 1m de distancia entre el micrófono y la fuente y los micrófonos pueden ser usados normalmente a una distancia de 0cm a 10cm produciendo un incremento en bajas frecuencias por el denominado “efecto de proximidad”. Pensando en evitar recortes de la señal a la salida del integrado se define una ganancia de 100 V/V y así obtener un voltaje de salida de 0.3V, suponiendo que su sensibilidad sea 2.8mV/Pa, valor que aumenta en la práctica. Según la Ecuación 9, para un R_g de 1k Ω la ganancia será de 50.4 V/V, y para un R_g de 500 Ω la ganancia será de 99.8 V/V. Estos

valores corresponden con la imagen dada por la hoja de datos del integrado, en donde se puede ver el valor de la ganancia dependiendo de la resistencia R_g (Revisar Anexo B).

Una vez que la señal de la voz y la señal del instrumento han sido adquiridas, son procesadas en el software de programación LabVIEW. La Figura 8 es un diagrama de conexiones a NI myRIO. En ella se pueden detallar entradas y salidas del sistema y la forma como se conectan micrófono y sintetizador (Revisar Anexo C para ver el circuito de entradas análogas del NI myRIO-1900).

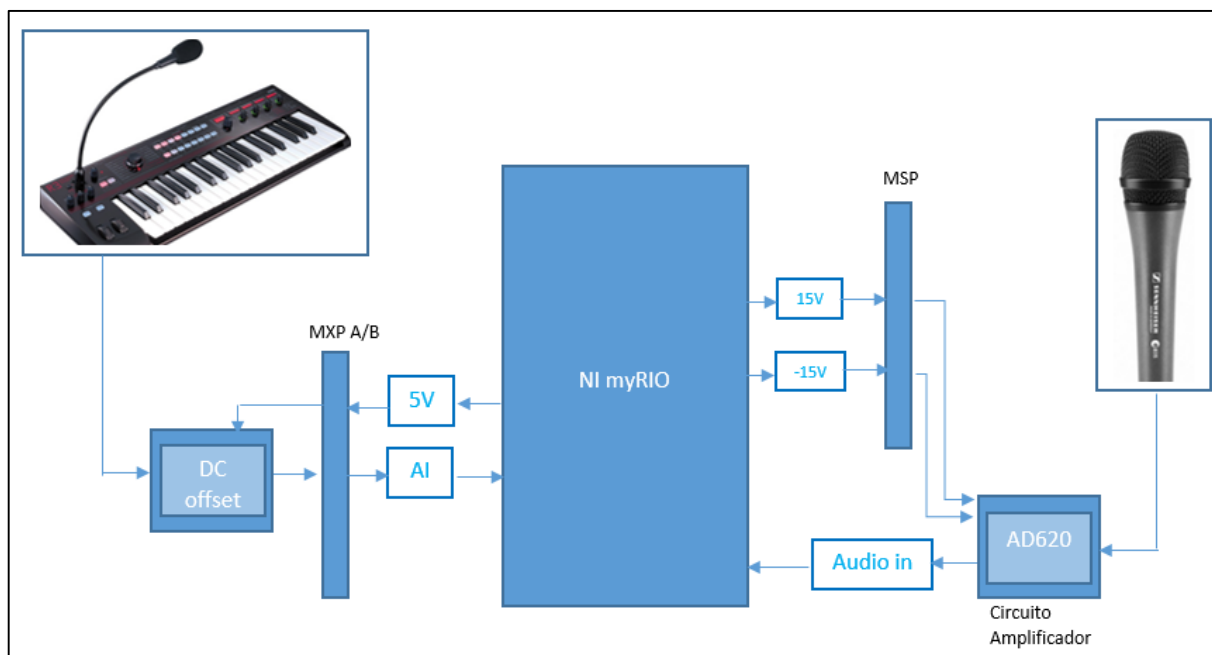


Figura 8. Diagrama de conexiones a NI myRIO.

Validación de la etapa de amplificación con el integrado AD620.

Con un R_g de 500Ω que obtiene 100 veces el voltaje de la entrada a la salida del integrado, se llevaron a cabo la medición de diferentes parámetros del integrado, entre ellos, función de transferencia, factor cresta, ganancia, THD y nivel de ruido, para el integrado en modo común y en modo diferencial, con una frecuencia de 1kHz para todos los casos. Se realizaron mediciones para dos generadores de señales para determinar si había alguna diferencia

en los resultados; los dos generadores fueron: Rigol DG1022 y Peak Tech 4025. Los resultados presentados a continuación muestran los valores obtenidos antes de que la señal de salida entrara en saturación. Estas mediciones se llevaron a cabo con la finalidad de determinar si el integrado AD620 cumplía con lo establecido por su fabricante. Los resultados obtenidos con el generador Rigol DG1022 no fueron satisfactorios y esto se atribuye a una probable imposibilidad que tiene el generador de generar señales de poca amplitud (amplitudes de 0.05 V a 0.35 V aproximadamente). Los resultados de las mediciones para este generador se muestran en el Anexo Q.

Medición en modo diferencial con el generador PeakTech 4025. Para hallar la función de transferencia, factor cresta y ganancia mediante las siguientes ecuaciones:

$$\text{Factor cresta} = \frac{V_{max}}{V_{rms}} \quad \text{Ecuación 10}$$

$$\text{Ganancia} = \frac{V_{pp} (\text{salida})}{V_{pp} (\text{entrada})} \quad \text{Ecuación 11}$$

Para mirar los resultados obtenidos de factor cresta y función de transferencia, dirigirse al Anexo Q. Los resultados de ganancia en modo diferencial medida con el generador de señales PeakTech 4025 se muestra en la siguiente imagen.

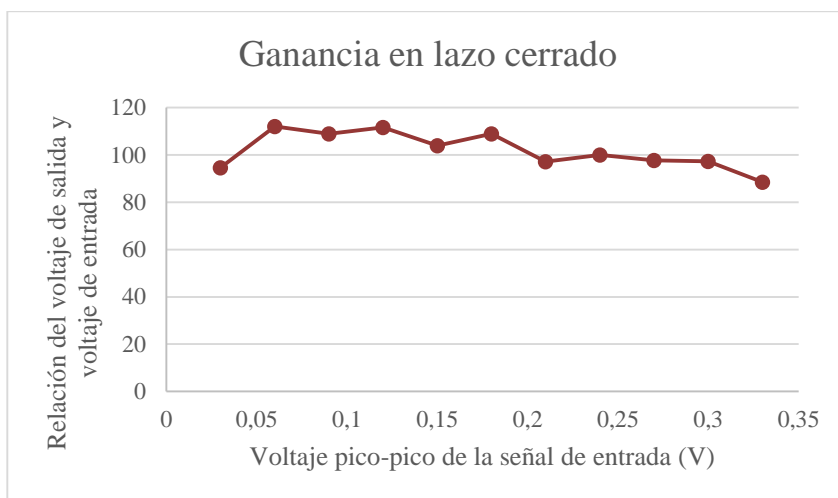


Figura 9. Ganancia resultado de medición en modo diferencial, generador PeakTeck 4025.

Para esta medición se esperaba tener una ganancia de 100 V/V, y este valor se vio reflejado en las mediciones. Por lo tanto el Integrado AD620 cumple con lo establecido por su fabricante y por ende es confiable que la ganancia que se va a obtener para una resistencia de 500Ω es la dada por la ecuación 9 y de ahí se puede partir para establecer la resistencia que se debe utilizar en el circuito de amplificación de la señal de la voz.

Un amplificador ideal es capaz de amplificar una señal senoidal pura para ofrecer una versión amplificada, en donde la forma de onda resultante es una señal senoidal pura de una sola frecuencia. Cuando ocurre la distorsión, la salida no es una copia exacta de la señal de entrada. Se considera que una señal tiene distorsión armónica cuando se presentan componentes frecuenciales adicionales a la frecuencia fundamental de la señal (Boylestad, S.F). Para obtener la distorsión armónica del integrado AD620, se utilizó el analizador de espectro del osciloscopio, para medir la señal de salida cuando esta empieza a recortarse. Esto ocurre cuando el voltaje de la señal de entrada es 0.15 Vpp, para los siguientes niveles de entrada se mantienen niveles de salida muy similares. Estos resultados se observan en la gráfica de la función de transferencia.

Los niveles de voltaje de los armónicos que aparecen en este punto se muestran en la siguiente tabla.

La ecuación para determinar la distorsión armónica es la siguiente:

$$THD = \frac{\sqrt{V1^2 + V2^2 + V3^2 \dots + Vn^2}}{V0} * 100 \quad \text{Ecuación 12}$$

Tabla 4. Voltaje de los armónicos de la señal. Generador PeakTeck 4025.

Armónico	Voltaje rms (Vrms)
A0	10
A1	0,2
A2	0,1

El resultado de la medición arroja una distorsión armónica de 0,021 o 2,1%.

Medición en modo común con el generador PeakTeck 4025. La medición de función de transferencia y factor cresta para el modo común también se puede ver en el Anexo Q. La siguiente imagen muestra los resultados de la medición de Ganancia del Integrado AD620 en modo común.

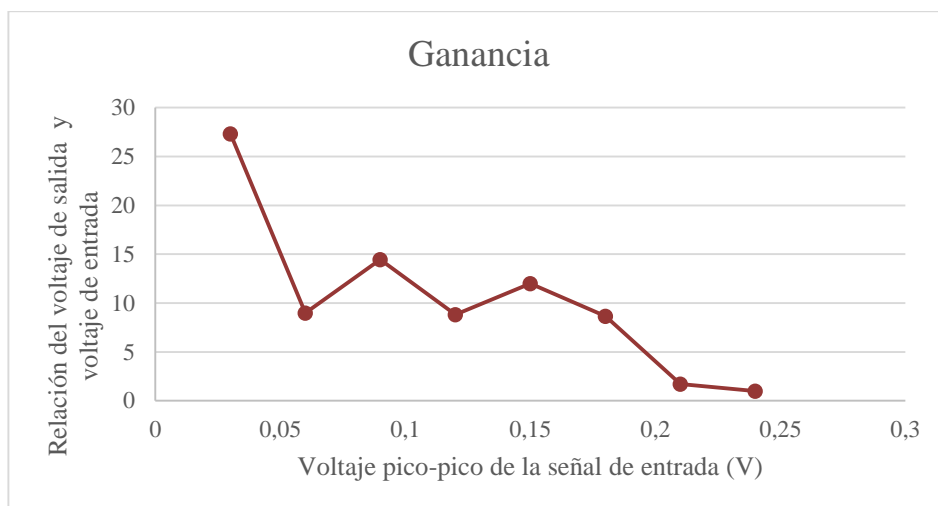


Figura 10. Ganancia resultado de medición en modo común, generador PeakTeck 4025.

Idealmente el amplificador de instrumentación debe amplificar las señales en modo diferencial e ignorar las señales en modo común. En la gráfica anterior se ve que la ganancia en modo común está alrededor de 10 V/V y aunque la teoría dice que para este caso la ganancia debe ser de 1V/V, en la práctica, las señales en modo común no son completamente ignoradas y por eso se obtienen niveles de ganancia superiores a 1.

Implementación y medición de los algoritmos (Banco de filtros, Spectrum y LPC)

Esta parte del documento inicia con una breve descripción de cómo fueron realizadas las mediciones de latencia, consumo de memoria RAM y análisis espectral de cada uno de los algoritmos. Los algoritmos cuya implementación y medición se describen en esta sección del documento, han sido implementados anteriormente por otros autores, sin embargo, en este proyecto han sido realizados a partir de la teoría de cada método y no son una réplica de otros algoritmos. La descripción de cada método se puede encontrar en el marco teórico del capítulo 3 de este documento.

Procedimiento de medición de latencia de los algoritmos

La latencia está directamente relacionada con la cantidad de tiempo que necesita un algoritmo para ejecutarse, dependiendo del tipo y la cantidad de datos que deben ser procesados.

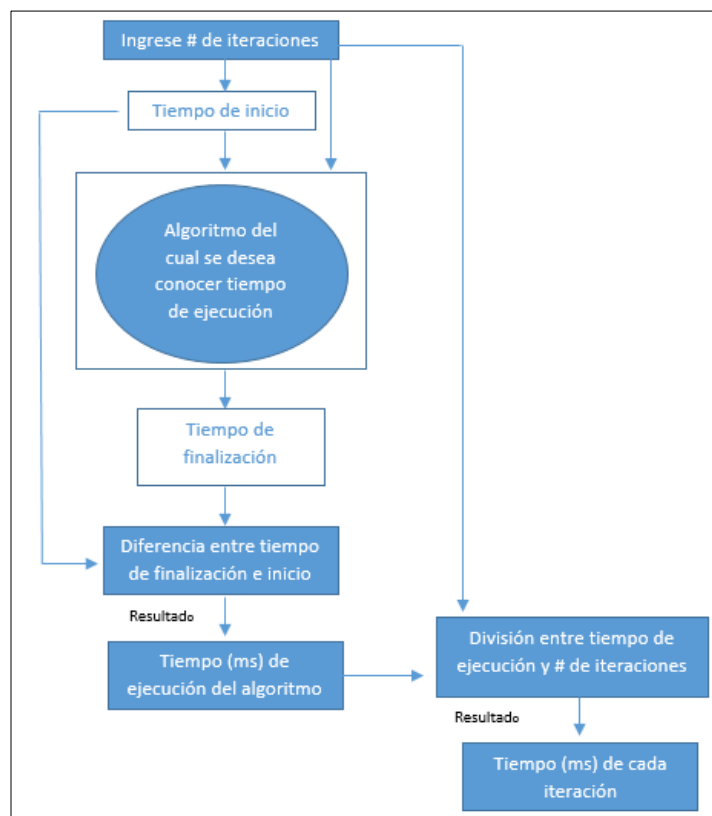


Figura 11. Diagrama de bloques del algoritmo usado para medir latencia y tiempo de ejecución de un algoritmo.

El algoritmo que corresponde al diagrama de la Figura 11 (Ir al Anexo M para ver su implementación en LabVIEW) utiliza un bloque de LabVIEW denominado “Flat Sequence Structure” el cual se asegura de que el algoritmo se ejecute secuencialmente; cada cuadro del Flat Sequence Structure se considera un subdiagrama y el código dentro del segundo cuadro no puede ejecutarse sin que el primero se haya ejecutado completamente. El algoritmo que se quiere medir, se inserta en el subdiagrama de la mitad dentro del ciclo For que se repetirá la cantidad de veces que sean determinadas por el usuario en “# of iterations”. El bloque con forma de reloj que aparece en el primer subdiagrama recibe el nombre de “Tick Count Function” y es un

temporizador. El primero contabiliza el tiempo de ejecución del primer subdiagrama y el segundo contabiliza el tiempo de ejecución del primer y segundo subdiagrama, por lo tanto, la diferencia entre los dos resultados es el tiempo de ejecución del segundo subdiagrama únicamente, que es donde se está ejecutando el algoritmo. Los resultados se dan en mili segundos y se muestran más adelante en este documento en la sección de cada algoritmo y en el análisis de resultados.

Procedimiento de medición de consumo de memoria RAM de los algoritmos

En LabVIEW cuando se programa, no se asignan variables sino que se crea un diagrama de bloques con conexiones, estas conexiones son la representación del flujo de datos de los algoritmos y estos bloques que en algunas ocasiones son funciones que son generadoras de datos, son las que se encargan de asignar, automáticamente, el almacenamiento de dichos datos, esto significa que la memoria usada por cada algoritmo realizado en LabVIEW es asignada automáticamente por los VI (Instrumentos virtuales). Para medir el consumo de memoria RAM de cada algoritmo, el cual depende de sus bloques, LabVIEW cuenta con una herramienta que recibe el nombre de Memory Usage que se puede encontrar en la barra de herramientas File >> VI Properties >> Memory Usage. Esta herramienta permite conocer el consumo de memoria del VI y el espacio que ocupa en el disco.

Procedimiento para la comparación espectral entre la señal de entrada y salida a cada algoritmo

La comparación espectral se realiza mediante la transformada rápida de Fourier que permite conocer la señal de la voz en el dominio de la frecuencia; La idea es realizar la FFT a la señal de la voz original, es decir, la señal de entrada a cada uno de los algoritmos y a su vez realizar la FFT a la señal de salida de cada uno de los algoritmos ya que con este procedimiento se conoce la amplitud de los armónicos de la voz en un instante de tiempo para ambas señales y

es posible comparar sus diferencias y similitudes. El algoritmo que obtenga la señal de salida con mayor similitud a la señal de entrada, será el que obtenga mejores resultados; los resultados se mostrarán con amplitud en escala lineal. El análisis espectral realizado a cada uno de los algoritmos se lleva a cabo mediante la misma señal de voz.

Caracterización de la señal de voz usada para las mediciones. Para cumplir con una de las partes del primer objetivo específico que consiste en hacer un análisis espectral para determinar la similitud entre las envolventes obtenidas con cada uno de los métodos (LPC, Spectrum y Banco de Filtros) y la envolvente original, se establecerán los requisitos de la señal de prueba y las características de la misma. Esta señal de prueba es utilizada en la comparación espectral de cada uno de los algoritmos.

La señal de prueba utilizada debe ser una señal de voz, sin embargo, existen muchas variaciones en la fonación que dependen del hablante y por ende ninguna persona produce los mismos sonidos; aunque mantengan ciertas relaciones formánticas (Báez, 2006). Por esta razón se realiza la caracterización de la señal de voz de una sola persona y su respectiva envolvente, mediante la FFT. La señal de prueba consiste en el sonido de la vocal A y tiene una duración de 10 segundos. Este análisis se realiza sobre la señal de entrada, antes de que esta sea procesada por cualquiera de los algoritmos.

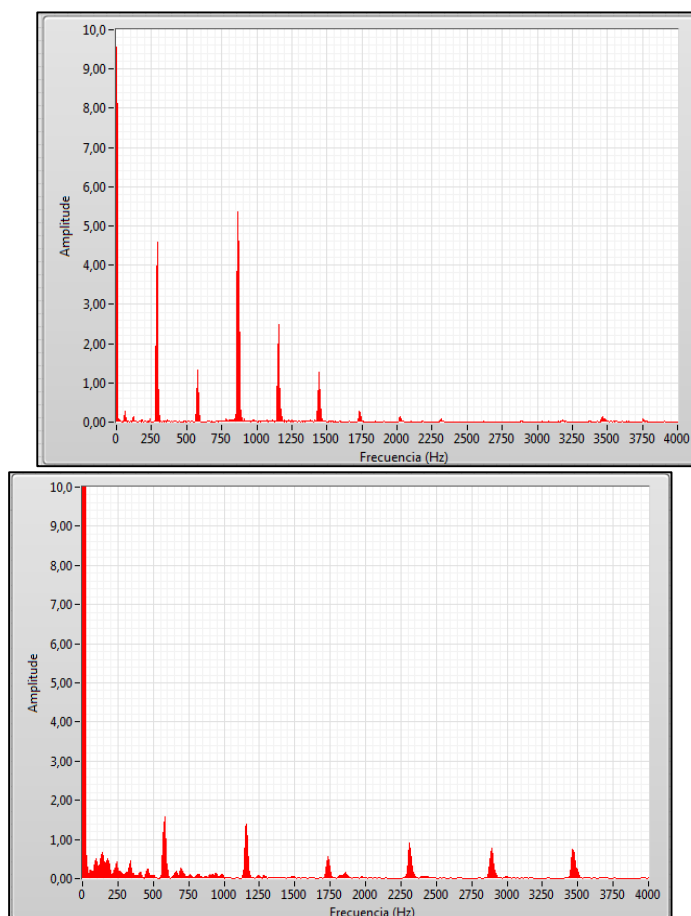


Figura 12. La figura de arriba es la FFT de la señal de la voz. La figura de abajo corresponde a la FFT de la envolvente de la señal de la voz.

El espectro de la figura anterior muestra los resultados en términos de RMS, es decir que su magnitud es medida en valores RMS y la amplitud se muestra en escala lineal. Hay importantes diferencias entre la señal de la voz y la envolvente de la misma. En las gráficas se puede ver que ambas señales comparten el armónico que aparece alrededor de los 600 Hz, el que aparece alrededor de los 1250 Hz y el que aparece en 1750 Hz, sin embargo, se puede ver que los armónicos de gran amplitud de la señal de voz, no aparecen en la FFT de la envolvente y esto se debe al método de detección de envolvente que implica pérdida de información frecuencial de la señal pues este método solo rescata las componentes en amplitud. La señal de la voz tiene un ancho de banda hasta 1750 Hz y en la envolvente aparecen armónicos hasta 3500 Hz. En las

gráficas se ve que las señales tienen un ruido de baja frecuencia, el cual es inherente al proceso de adquisición de la señal, sin embargo, en la parte de procesamiento de los algoritmos, se pone un filtro pasa altos con frecuencia de corte en 40 Hz que elimina este ruido, esto se puede apreciar más adelante en las mediciones espectrales para cada algoritmo.

En resumen, el análisis espectral de cada algoritmo se lleva a cabo con la señal de la figura 12. Las únicas características importantes de esta señal, es que debe ser un sonido de la voz, emitido por la persona con la que se realizaran el resto de pruebas y que debe ser el mismo sonido durante todo el intervalo de tiempo; como se mencionó antes, este sonido corresponde a la articulación de la vocal A; esto con la finalidad de que haya la menor cantidad de cambios en frecuencia y amplitud durante el intervalo de 10 segundos.

Implementación y medición de Banco de filtros

Implementación del banco de filtros. Como se menciona en el marco teórico, un banco de filtros es una colección de filtros digitales con una entrada común o una salida común. En este trabajo se ha implementado una colección de filtros paralelos que cubren un ancho de banda de 100Hz a 4kHz, considerando el ancho de banda de la voz humana. El diagrama de flujo del Banco de filtros implementado en este proyecto se muestra en la siguiente imagen:

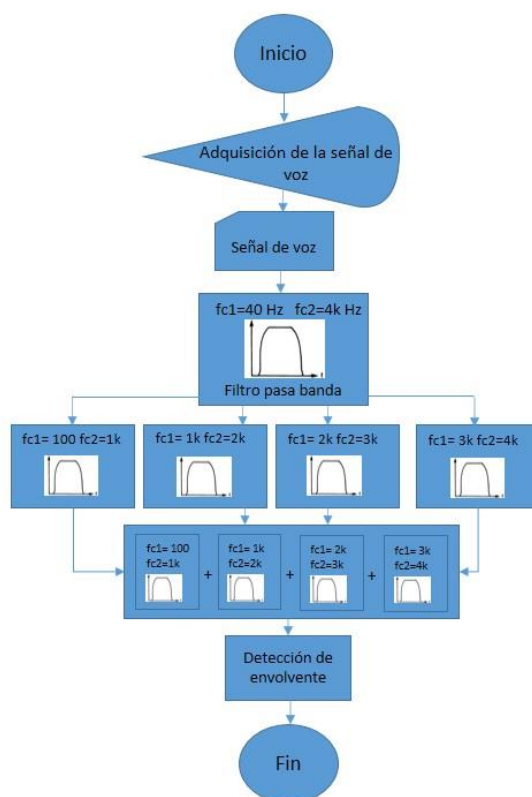


Figura 13. Diagrama de flujo del Banco de Filtros.

La teoría de Banco de Filtros no menciona el tipo de filtros digitales que se deben utilizar para implementarlo, por eso se realiza un proceso de selección entre los filtros digitales que pueden ser del tipo FIR cuyas siglas significan Finite Impulse Response o Respuesta Finita al Impulso, o del tipo IIR cuyas siglas significan Infinite Impulse Response o Respuesta Infinita al Impulso. En la Tabla 5 se comparan en términos generales, diferentes aspectos de estos dos tipos de filtro para elegir el que se ajuste mejor al proyecto.

Tabla 5. Comparación de los filtros FIR e IIR.

	Filtros FIR	Filtros IIR
Ventajas	Fase de la respuesta en frecuencia lineal. Son estables si son realizados como sistemas no recursivos.	Necesita menos coeficientes que el filtro FIR para obtener los mismos resultados, por lo tanto

Desventajas	Necesitan mayor cantidad de memoria que el filtro IIR para almacenar la muestra actual y muestras anteriores de la señal de entrada.	Necesitan menos registros de almacenamiento. Respuesta en fase no lineal.
--------------------	--	--

Los filtros FIR son estables¹ si son realizados como sistemas no recursivos, es decir, que no dependen de valores pasados de la salida. Los filtros FIR se emplean si se requiere que los filtros sean de fase lineal dentro de su banda de paso. Sin embargo, un filtro IIR tiene lóbulos laterales de menor amplitud en la banda de rechazo que un filtro FIR con la misma cantidad de coeficientes. Aunque los filtros FIR son capaces de tener una respuesta lineal en fase, requieren mucha memoria y también requieren un mayor número de coeficientes respecto al filtro IIR. En general, la implementación de filtros IIR involucra menos coeficientes, requiere menos memoria y tiene menor complejidad computacional (Proakis, 1998).

Para realizar una comparación empírica de los dos tipos de filtros digitales (FIR e IIR) para determinar cuál es el apropiado para el banco de filtros, ambos fueron implementados en LabVIEW. El bloque del filtro FIR en LabVIEW, recibe el nombre de *FIR windowed filter VI* y todos sus parámetros se muestran en el Anexo R. El parámetro “taps” del bloque determina el número de coeficientes del filtro y dependiendo de la cantidad de coeficientes es la selectividad del mismo. El valor por defecto para este parámetro es 25. Todos los módulos de filtros tipo FIR en LabVIEW utilizan la técnica “windowing”. Uno de los efectos de utilizar esta técnica es reducir el fenómeno de Gibbs que aparece como lóbulos cerca de las frecuencias de corte del

¹ En un sistema estable, toda secuencia de entrada $x(n)$ o $y(n)$ se traduce en la existencia de un par de números finitos M_x y M_y para todo n , entonces si para alguna entrada acotada $x(n)$ la salida no está acotada, o sea, es infinita, el sistema se clasifica como inestable (Proakis, John G. 1998).

filtro; estos lóbulos aparecen con más frecuencia a medida que los coeficientes aumentan. El parámetro “window” del bloque determina la ventana utilizada para suavizar este comportamiento, según John Proakis autor del libro Tratamiento Digital de Señales, se debe elegir una ventana que tenga una cima y decaiga gradualmente hacia cero, en lugar de abruptamente como ocurre en una ventana rectangular (La ventana rectangular viene por defecto en el bloque FIR windowed Filter VI). A continuación se muestra la respuesta en magnitud del filtro, usando una ventana Rectangular y una ventana Hamming respectivamente, para 20 coeficientes.

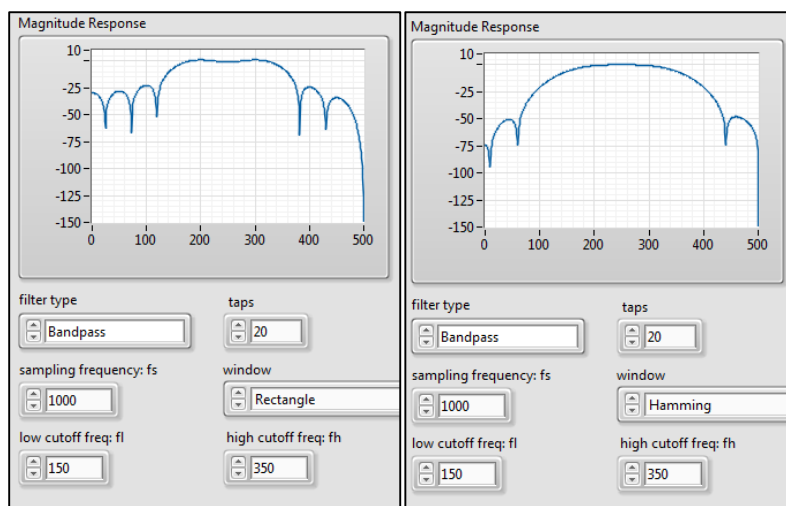


Figura 14. La figura de la izquierda es la respuesta en frecuencia de filtro FIR pasabanda con ventana rectangular y la figura de la derecha es la respuesta en frecuencia del filtro FIR pasabanda con ventana Hamming.

El decaimiento de la ventana Hamming es gradual a diferencia de la ventana rectangular y por esta razón tiene menor cantidad de lóbulos alrededor de la banda de paso del filtro. Aunque con la ventana Hamming se logra un mejor resultado, este método de diseño tiene varias desventajas, como bandas de transición más anchas y dificultad para definir frecuencias de corte con atenuación específica; en conclusión, aunque el diseño por ventanas no requiere gran cantidad de recursos computacionales, no es la mejor técnica para el diseño de filtros.

El bloque del filtro IIR de LabVIEW, recibe el nombre de *Butterworth Filter VI*, cuyos parámetros se muestran en el Anexo R. Para este caso, es el parámetro “order” el que especifica los coeficientes del filtro. Su valor por defecto es 2, porque como fue mencionado anteriormente, los filtros IIR necesitan menos coeficientes para obtener los mismos resultados que los filtros FIR.

Los filtros IIR en LabVIEW, se pueden diseñar con las técnicas habituales y se ha elegido hacer un diseño Butterworth ya que no tiene presencia de rizado en la banda de paso o en la banda atenuada y tiene una respuesta suave en todas las frecuencias ayudando a evitar distorsiones. El funcionamiento interno de este VI (Instrumento Virtual), consiste en hallar los coeficientes Butterworth según los parámetros dados por el usuario y luego llama a un filtro IIR en cascada; este diseño puede resultar ventajoso si se requiere un orden de filtro elevado.

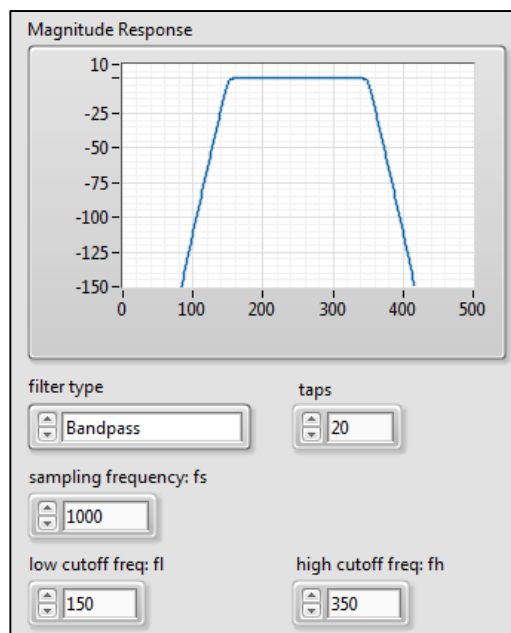


Figura 15. Respuesta en frecuencia de filtro IIR para 20 coeficientes.

En la Figura 15 se aprecia que el filtro IIR es más selectivo y no tiene lobulaciones y este resultado se consigue con igual número de coeficientes que fueron usados con los filtros FIR (20

coeficientes). Esto demuestra que los filtros IIR necesitan una menor cantidad de coeficientes que los filtros FIR y por ello requieren menos memoria del sistema embebido para ser procesados y además su respuesta en frecuencia no tiene lobulaciones y es plana.

A partir de los anteriores resultados se determinó que el banco de filtros se debe conformar de bloques Butterworth Filter VI, por sus ventajas respecto a los filtros FIR.

Para cada banda de frecuencia del banco de filtros se utilizaron ciertas frecuencias de corte que se muestran en la Tabla 6. La primera frecuencia del banco de filtros es 100 Hz porque como se puede ver en la caracterización de la voz realizada anteriormente, no existen componentes frecuenciales por debajo de 250 Hz en la señal de la voz, así que la frecuencia de corte se elige en 100 Hz previendo la aparición de un armónico entre 100 y 250 Hz, y las señales que puedan aparecer por debajo de 100 Hz se consideran ruido de baja frecuencia o armónicos con muy poca amplitud y no son relevantes para el análisis. La implementación del algoritmo Banco de Filtros en LabVIEW se muestra en el Anexo D.

Tabla 6. Frecuencias de corte para el banco de filtros.

Filtro	Frecuencia de corte inferior	Frecuencia de corte superior
Filtro 1	100 Hz	1100 Hz
Filtro 2	900 Hz	2100 Hz
Filtro 3	1900 Hz	3100 Hz
Filtro 4	2900 Hz	3900 Hz

La respuesta en frecuencia obtenida con las frecuencias de corte de la tabla 6, se muestra en la Figura 16. Se puede ver que la respuesta no es plana y presenta rizados alrededor de las frecuencias de corte de los filtros.

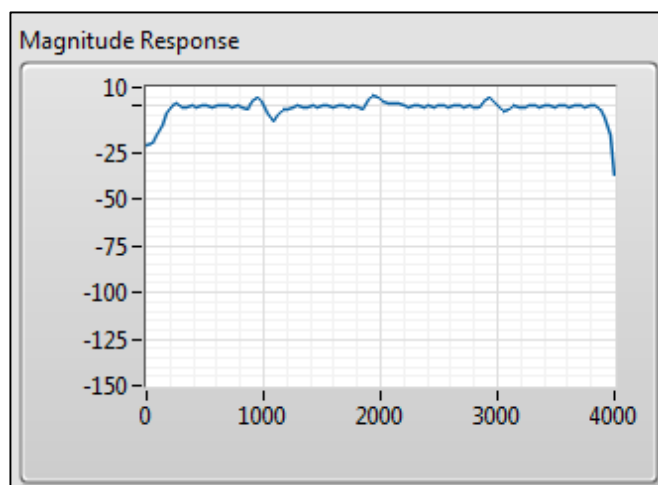


Figura 16. Respuesta en frecuencia de banco de filtros 1.

Buscando una respuesta en frecuencia plana, se establecen las siguientes frecuencias de corte. En el siguiente diseño del banco de filtros, la frecuencia de corte superior del primer filtro es la frecuencia inferior del siguiente filtro.

Tabla 7. Frecuencias de corte para el banco de filtros con respuesta plana.

Filtro	Frecuencia de corte inferior	Frecuencia de corte superior
Filtro 1	100 Hz	1000 Hz
Filtro 2	1000 Hz	2000 Hz
Filtro 3	2000 Hz	3000 Hz
Filtro 4	3000 Hz	4000 Hz

Con estas frecuencias de corte se obtiene una respuesta en frecuencia plana con algunas lobulaciones de poca amplitud. Este es el resultado esperado y se definen las frecuencias de la Tabla 7 como las frecuencias de corte del banco de filtros.

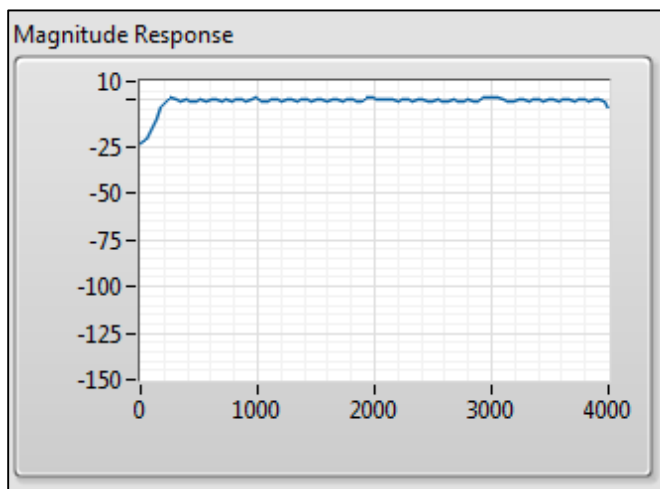


Figura 17. Respuesta en frecuencia de banco de filtros 2.

Para finalizar con la descripción de la implementación, se muestra un diagrama de bloques del banco de filtros realizado.

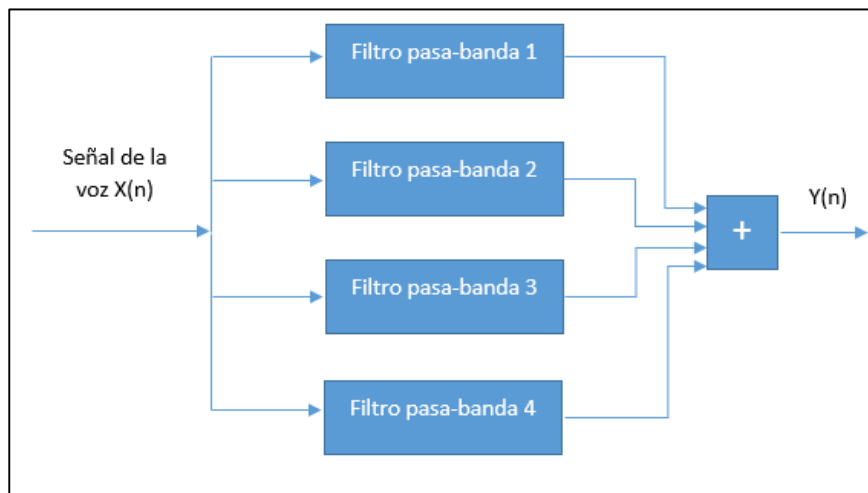


Figura 18. Diagrama de bloques del algoritmo Banco de Filtros.

Medición del banco de filtros. En esta sección del documento se presentan las mediciones de latencia, consumo de memoria RAM y análisis espectral del Banco de filtros. El número de bloques de cada algoritmo es una medida que también fue propuesta en los objetivos específicos, así que los resultados de esto también se muestran en esta sección.

Tabla 8. Resultados de las mediciones para Banco de filtros

	Resultado de medición de latencia (Tiempo de 100 iteraciones (ms))	Kylo bytes de memoria usada	Número de bloques
Banco de filtros	20556	3384,2K	11

Como se observa en la tabla anterior, el tiempo de ejecución de 100 iteraciones del algoritmo de banco de filtros da un resultado de 20556 ms, es decir que el tiempo de una sola iteración es 205,56 ms; dados estos resultados, se puede considerar que el algoritmo no se ejecuta en tiempo real, pues para que así sea, se dice que el algoritmo debe tardar menos de 30 ms en ejecutarse. Por otro lado, el resultado de la medición de memoria RAM usada por el algoritmo es de 3384,2K bytes este resultado es satisfactorio pues el sistema embebido cuenta con una capacidad de 512 MB de memoria, es decir que el consumo del algoritmo está por debajo del límite del sistema. Por último, el número de bloques utilizados en la implementación del Banco de Filtros es 11. Esta cantidad no es relevante y no se puede sacar un análisis de este resultado si no se compara con el resultado de los otros dos algoritmos.

El análisis espectral de las señales de entrada y salida al banco de filtros arroja los siguientes resultados:

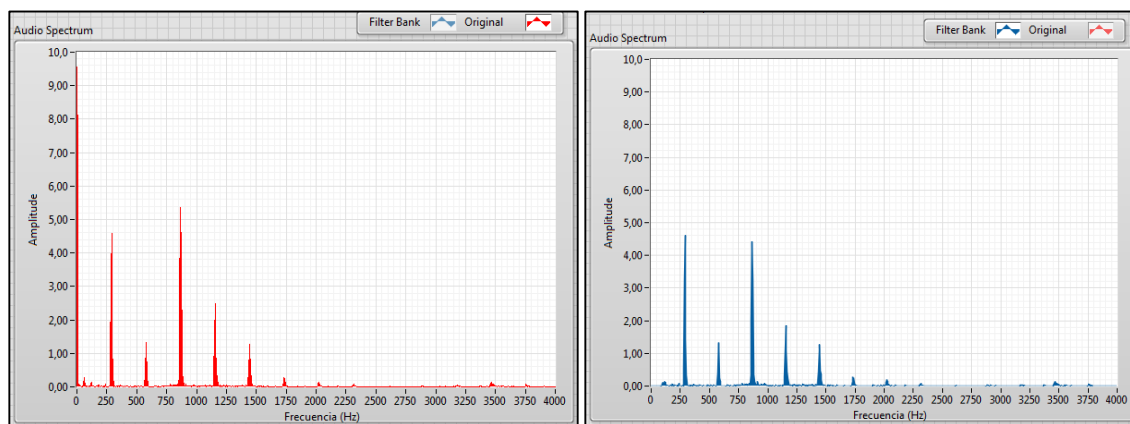


Figura 19. La gráfica de la izquierda muestra la FFT de la señal de voz original (señal de entrada) y la gráfica de la derecha muestra la FFT de la señal de salida del banco de filtros.

Las gráficas anteriores muestran la FFT de la señal de la voz antes y después de ser procesada por el Banco de filtros. Uno de los principales cambios es que la señal de salida del banco de filtros no tiene el ruido de baja frecuencia, que como se había dicho anteriormente, es inherente a la adquisición de la señal por el NI myRIO y esto es debido a que la señal es filtrada por un filtro pasa bandas con frecuencias de corte $f_{c1} = 40$ Hz y $f_{c2} = 4000$ Hz, para asegurarse que la voz que vaya a ser procesada no se salga de este ancho de banda y a su vez eliminar el ruido de baja frecuencia que daña la calidad de la señal. Otro cambio es que los armónicos de la señal de salida del banco de filtros han perdido amplitud respecto a la señal de entrada y puede que aunque el bloque del LabVIEW *Butterworth Filter VI* con el que se implementaron los filtros no presente una ganancia del todo unitaria en la práctica. Se debe rescatar que la amplitud de estas graficas es lineal.

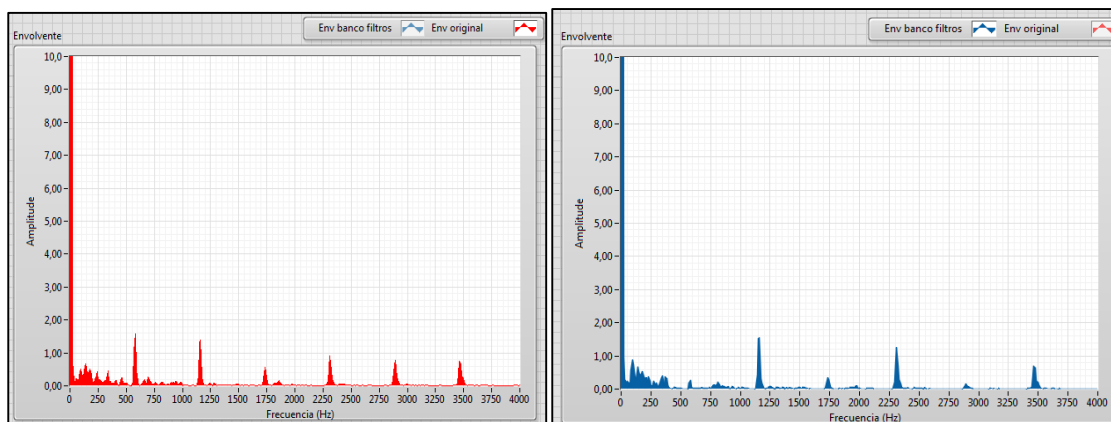


Figura 20. La gráfica de la izquierda muestra la FFT de la envolvente de voz original (señal de entrada) y la gráfica de la derecha muestra la FFT de la envolvente de la señal de salida del banco de filtros.

En cuanto al análisis de la envolvente de la voz, se observa que hay cambios considerables entre la envolvente de entrada y la de salida. Se recalca nuevamente, que la señal de la envolvente pierde armónicos respecto a la señal de la voz y esto es debido a proceso de detección de envolvente que se realiza pues al hallar la envolvente que busca rescatar únicamente las componentes en amplitud de la señal, algo que se puede apreciar de mejor manera en el

dominio del tiempo. Además de esto, al procesar la señal con el banco de filtros, nuevamente se pierde amplitud en los armónicos, pero en este caso es más considerable pues algunos desaparecen casi por completo.

Tabla 9. Error cuadrático medio de las señales de voz y de las envolventes.

Error cuadrático medio entre las señales de voz de entrada y salida	Error cuadrático medio entre las envolventes de entrada y salida
0,009	0,26
0,03	0,26
0,001	0,16
0,04	0,17
0,01	0,26
0,009	0,27
0,2	0,2

El error cuadrático medio se halla con la intención de determinar de forma numérica la similitud de las señales que están siendo comparadas. La señal de voz analizada tiene una duración de 10 segundos, por lo tanto, se halla el error cuadrático medio en diferentes instantes de tiempo. Idealmente, el error cuadrático medio debe ser igual a cero si las señales comparadas son completamente iguales. En resultado promedio del error cuadrático para las señales de voz es de 0.042 y para las envolventes es de 0.22. El valor obtenido para las señales de voz es más cercano a cero que el obtenido para las envolventes y esto significa que las señales de voz son similares entre si mientras que las envolventes presentan diferencias.

Implementación y medición de Spectrum

Implementación de Spectrum. A continuación se muestra el diagrama de flujo del algoritmo Spectrum, para facilitar el entendimiento de la implementación:

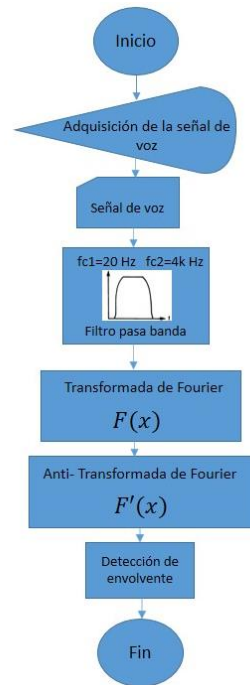


Figura 21. Diagrama de flujo de Spectrum

Este método reconstruye la señal a partir de la transformada rápida de Fourier y la anti-transformada de Fourier. Como se ha mencionado antes, esta transformada se puede interpretar como un camino para pasar del dominio del tiempo al dominio de la frecuencia y dicha señal puede reconstruirse a partir de sus coeficientes espectrales mediante la anti-transformada de Fourier, que básicamente es el camino de regreso al dominio temporal. La FFT impone la restricción de que el número de muestras temporales que toma de la señal de entrada sea exactamente una potencia de 2, y la función obtenida tendrá una cantidad de muestras igual a la mitad de las que tenía la función temporal (Martí, 1987). Para este caso el tamaño de la FFT es 12000 y por lo tanto, la función en el dominio de la frecuencia tiene 6000 muestras. Para realizar la Transformada rápida de Fourier o FFT en LabVIEW, se utiliza el bloque denominado FFT VI que convierte la señal de voz en el dominio del tiempo al dominio de la frecuencia, mediante DFT o Transformada Discreta de Fourier que funciona usando el algoritmo de la Transformada

Rápida de Fourier. La Transformada de Fourier se puede desarrollar transformando una señal periódica en una señal no periódica mediante el alargamiento sin límite del periodo (Ambardar, 2002), pero la DFT requiere que la función de entrada tenga duración finita, es decir, requiere solo un periodo de la señal periódica que se extiende de forma infinita. La entrada para la DFT puede estar compuesta de números reales y complejos y el siguiente paso, es aplicar la anti-transformada de Fourier, mediante el bloque denominado Inverse FFT VI, para regresar la señal al dominio del tiempo. Ir a Anexo F para ver su implementación en LabVIEW.

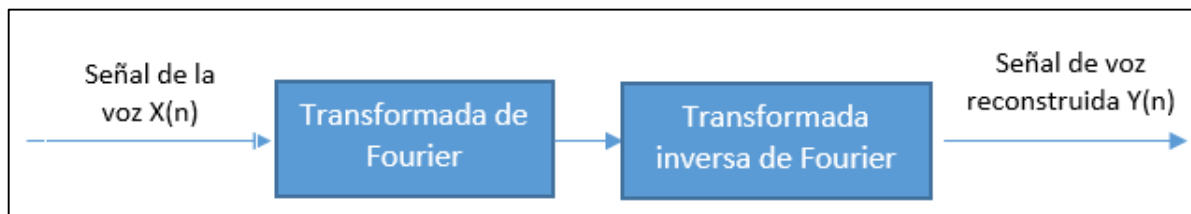


Figura 22. Diagrama de bloques del algoritmo Spectrum.

Medición de Spectrum. En esta sección se muestran los resultados de la medición de tiempo de latencia, consumo de memoria RAM y comparación y análisis espectral de las señales de entrada y salida del método. Además de esto se muestra el número de bloques necesarios para implementar el algoritmo.

Tabla 10. Resultados de las mediciones para Spectrum

	Resultado de medición de latencia (Tiempo de 100 iteraciones (ms))	Kylo bytes de memoria usada	Número de bloques
Spectrum	20795	3437,4K	3

En este caso el tiempo de latencia de una sola iteración es de 207,95 ms y evidentemente este algoritmo tampoco se realiza en tiempo real, además de esto, es mayor al tiempo de ejecución del Banco de Filtros y por lo tanto, en la práctica la latencia se percibe con facilidad en

este método. También usa mayor cantidad de memoria respecto al banco de filtros, pero a pesar de esto, su resultado es aceptable puesto que como se dijo anteriormente, el NI myRIO tiene 512 MB de memoria RAM. El número de bloque utilizado en la implementación de este algoritmo es 3 y comparando con el número de bloques del Banco de Filtros que era 11, se puede concluir que el consumo de memoria RAM y el tiempo de latencia son independientes del número de bloques, ya que el banco de filtros con mayor número de bloques obtuvo menos tiempo de latencia y menor consumo de memoria RAM, es decir que estos tiempos están relacionados sobre todo con las funciones que realizan cada bloque y no con su cantidad.

El análisis espectral de las señales de entrada y salida al método Spectrum, se muestran en las siguientes imágenes:

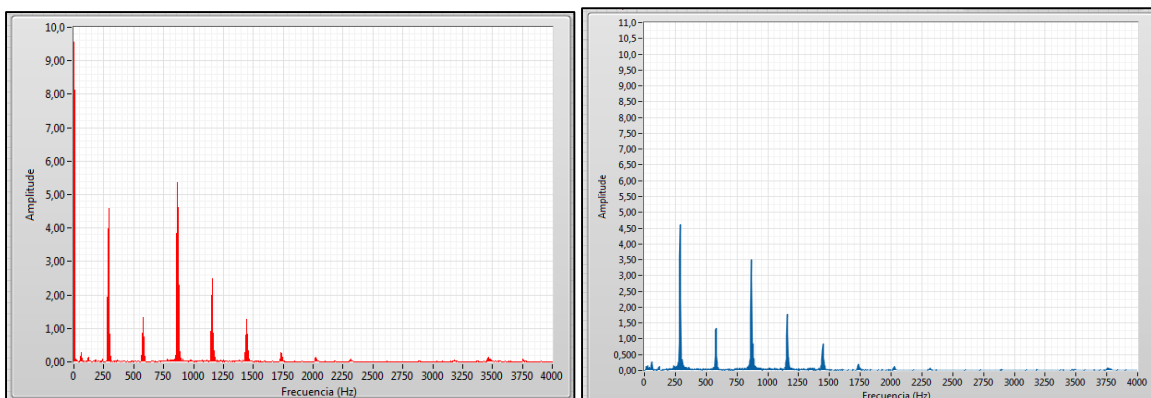


Figura 23. La gráfica de la izquierda muestra la FFT de la señal de voz original (señal de entrada) y la gráfica de la derecha muestra la FFT de la señal de salida del método Spectrum.

La grafica de la izquierda corresponde a la FFT de la señal de prueba, nuevamente se puede ver que hay ruido de baja frecuencia, y que es eliminado al procesar la señal con el algoritmo, para este caso, se aplicó un filtro pasa banda con frecuencias de corte: 20 Hz y 4000 Hz. Se filtró desde 20 Hz, pues a la salida de método se seguía observando ruido en bajas frecuencias con 40 Hz con frecuencia de corte. Para este caso se observan pérdidas de ganancia

en los armónicos en la señal de salida del método, es decir que hay pérdida de información al realizar la transformada de Fourier y luego realizar la anti-transformada para devolver la señal al dominio del tiempo: esto en teoría es una reconstrucción de la señal que pasa del dominio del tiempo la dominio de la frecuencia y luego regresa al dominio del tiempo y es posible que durante este proceso se presente una pérdida de información que cause las diferencias de amplitud entre los armónicos de las dos señales.

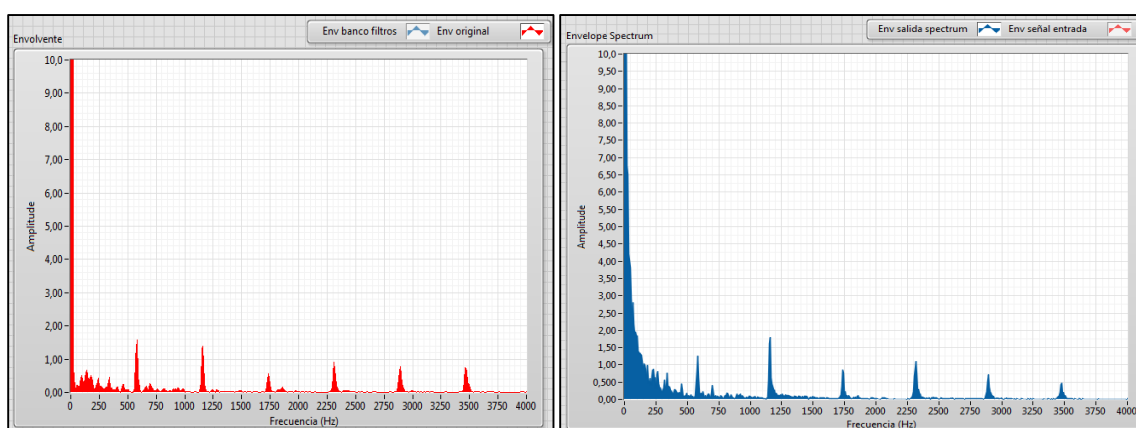


Figura 24. La gráfica de la izquierda muestra la FFT de la envolvente de voz original (señal de entrada) y la gráfica de la derecha muestra la FFT de la envolvente de la señal de salida del método Spectrum.

Por otro lado, las señales de las envolventes tienen mayor similitud para Spectrum que para Banco de filtros, en este caso se ve que todos los armónicos de la señal de entrada se mantienen a la salida del método, pero aparece ruido en baja frecuencia a pesar de que la señal de entrada ha sido filtrada, haciendo FFT de la señal de salida en diferentes partes, se obtiene que el ruido aparece luego de la Anti-transformada de Fourier.

Tabla 11. Error cuadrático medio de las señales de voz y de las envolventes.

Error cuadrático medio entre las señales de voz de entrada y salida	Error cuadrático medio entre las envolventes de entrada y salida
0,08	0,3
0,9	0,2
0,07	0,5
0,05	0,2
0,22	0,34

0,23	0,32
0,17	0,3

Nuevamente, se recuerda que la amplitud de las gráficas mostradas en esta sección están en escala lineal y por lo tanto las diferencias entre la señal de entrada y salida para el método Spectrum no son significativas. El promedio de los valores de error cuadrático medio entre las señales de voz de entrada y salida es 0.24, se recuerda que el error cuadrático medio debe aproximarse a cero si las señales son parecidas, sin embargo en este caso el resultado es de 0,24 y es mayor al obtenido con el Banco de filtros. Ahora, el promedio de los valores de error cuadrático entre las envolventes es de 0,3, que no se aleja del valor obtenido con el banco de filtros para este caso. Sin embargo, los dos valores de error cuadrático medio son mayores para este método que para el método banco de filtros, esto significa que entre estos dos métodos, el banco de filtros obtiene una señal de salida con mayor similitud a la original.

Implementación y medición de Linear Predictive Coding (LPC)

Implementación de Linear Predictive Coding. La finalidad de este método es predecir una muestra de voz mediante una combinación de muestras de voz pasadas. La predicción de la muestra de voz se realiza a través de un filtro FIR que utiliza unos coeficientes que son hallados mediante la resolución de una matriz de autocorrelación. La autocorrelación se realiza entre dos bloques de la señal multiplicados, normalmente, por una ventana Hamming y desfasados entre sí. Según lo anterior, para implementar el método, la señal de voz debe ser adquirida y luego darle peso a una sección de esta señal mediante una ventana Hamming y normalizarla dividiéndola por su máximo valor en amplitud, después se realiza una autocorrelación entre la señal inventanada y la señal original. El resultado se considera una matriz que es resuelta mediante el método de Levinson Durbin y ese resultado son los coeficientes LPC. Los coeficientes LPC modelan la

componente espectral del segmento usado como señal original. En el Anexo G se muestra el algoritmo implementado en MATLAB usando el método de autocorrelación para hallar los coeficientes.

Por otro lado, MATLAB cuenta con una función llamada LPC que se encarga de hallar los coeficientes LPC de la señal inventanada. Estos se convierten en los coeficientes de un filtro FIR cuya salida es la señal predicha o esperada. Cuando se tiene esta señal esperada se halla el error de predicción que es el resultado de la resta entre la señal original y la señal predicha el cual se encarga de determinar si los coeficientes hallados son los adecuados, dependiendo de lo alejado que esté su valor de cero. Con esta función, deja de ser necesario usar el método de autocorrelación para hallar los coeficientes del filtro. Para un mejor entendimiento se muestra a continuación el diagrama de flujo del método LPC.



Figura 25. Diagrama de flujo de LPC

Los resultados de la implementación de LPC en MATLAB con la función `lpc`, se muestra en las siguientes imágenes. El algoritmo con esta función se muestra en el anexo H.

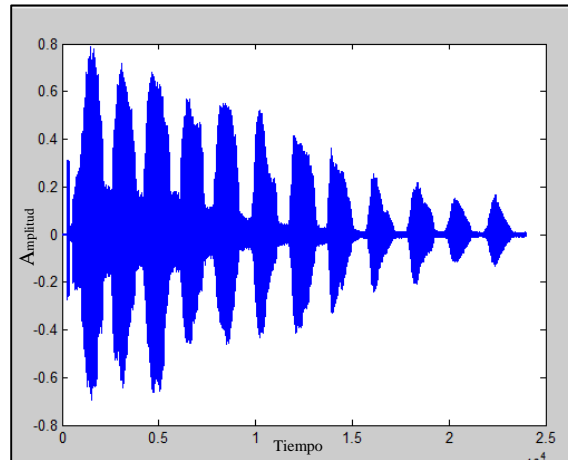


Figura 26. Señal original. Señal de entrada al algoritmo LPC.

La imagen anterior es la señal de entrada al modelo LPC. No tiene ninguna característica especial, ni era necesario que tuviera alguna, simplemente es una señal de voz grabada durante un intervalo de tiempo la cual será procesada por el algoritmo descrito en el diagrama de flujo anterior.

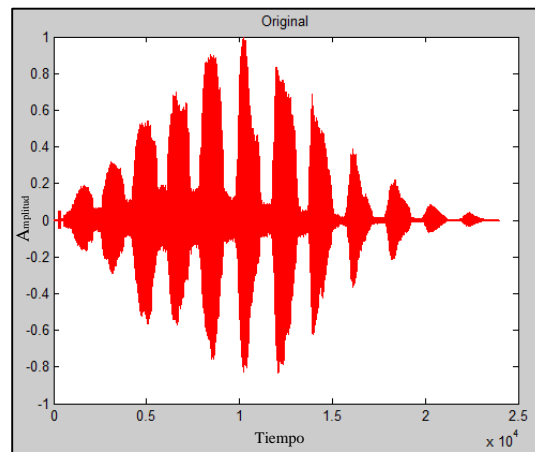


Figura 27. Señal original multiplicada por una ventana Hamming.

La imagen anterior muestra la señal original multiplicada por una ventana Hamming, por ello se ve la forma de la ventana en la señal original, esto con la finalidad de darle peso a algunas

partes de la señal. Si luego de realizar todo el procedimiento, el resultado obtenido no es el deseado se pueden variar los pesos de la señal para tratar de conseguir un mejor resultado.

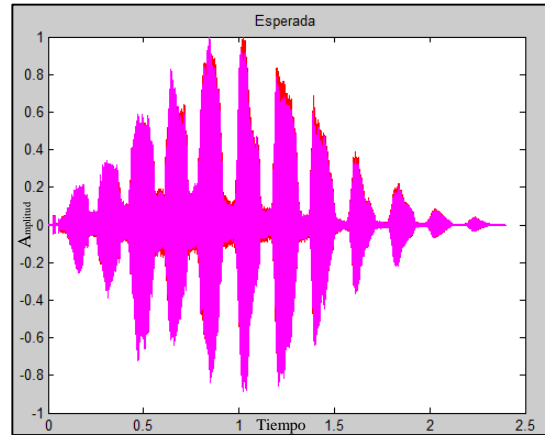


Figura 28. Salida del filtro FIR. Señal esperada.

La figura anterior muestra la señal de salida del filtro FIR que utiliza los coeficientes LPC hallados para hallar la señal esperada o predicha; esta señal esperada se dibuja sobre la gráfica de la señal original y es de color rosa. Para este caso fueron hallados 10 coeficientes, a medida que la cantidad de coeficientes hallados aumenta, aumenta el tiempo de procesamiento del algoritmo, pero el filtro se vuelve más selectivo. Se puede ver que la señal predicha es similar a la señal original, tienen la misma duración y las mismas amplitudes en casi todo el intervalo de tiempo.

Para determinar qué tan parecidas son las dos señales (original y predicha) se halla el error de predicción entre las dos. Se espera que los valores del error de predicción se aproximen a cero. El resultado se muestra en la siguiente gráfica:

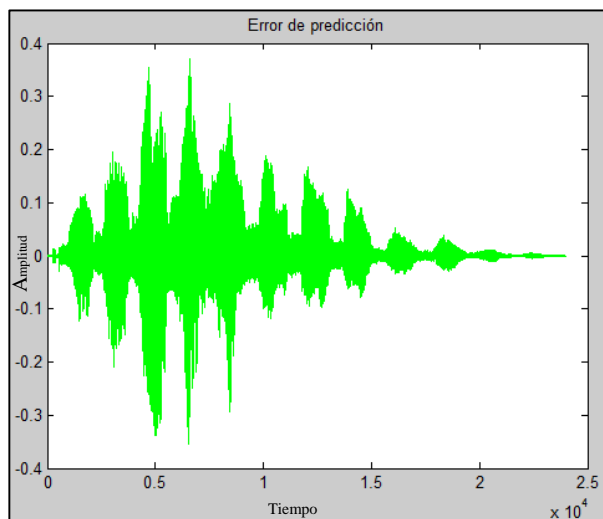


Figura 29. Error de predicción obtenido.

Sin un factor de aprendizaje, los resultados obtenidos del error de predicción son más cercanos a cero que el obtenido con algoritmo de auto-correlación descrito anteriormente (el error de predicción del método de auto-correlación se puede ver en el Anexo G). El error cuadrático medio hallado mediante los datos de la gráfica anterior, ayuda a entender que tan parecidas son las dos señales. El resultado es un valor muy cercano a cero, lo que significa que las señales son parecidas.

Tabla 12. Error cuadrático medio entre señal filtrada y señal original.

Error cuadrático medio	1.2602e-05
-------------------------------	------------

Para implementar el algoritmo LPC en LabVIEW con la función LPC anterior, se utiliza el módulo MathScript de LabVIEW que permite integrar código .m en el entorno de programación de LabVIEW, como se muestra en el Anexo I.

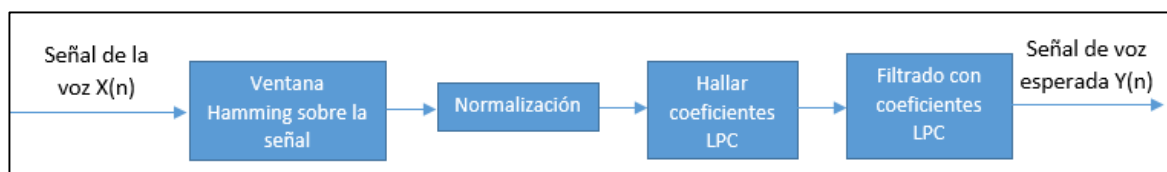


Figura 30. Diagrama de bloques del algoritmo LPC.

Medición de Linear predictive coding. En esta sección del documento se muestran las mediciones de Latencia, memoria RAM, análisis espectral y número de bloques necesarios para implementar el algoritmo.

Tabla 13. Resultado de las mediciones para LPC

	Resultado de medición de latencia (Tiempo de 100 iteraciones (ms))	Kylo bytes de memoria usada	Número de bloques
LPC	20562	3648,2K	10

No hay una diferencia significativa en los tiempos de latencia de banco de filtros y LPC, ya que el tiempo de latencia para el banco de filtros fue de 20556 ms, es decir, que el algoritmo que tiene mayor latencia es Spectrum y Banco de filtros y LPC tienen casi el mismo tiempo. Por otro lado, es el que una más Kylo bytes de memoria, así que para este caso se determina que la cantidad de memoria usada no está directamente relacionada con el tiempo de latencia, ya que aunque es el algoritmo con mayor cantidad de memoria usada, no es el que presenta mayor tiempo de latencia. Y nuevamente se comprueba que el número de bloques es una medida irrelevante que no puede ser relacionada con las otras medidas y no determina la eficiencia de los algoritmos.

El análisis espectral de las señales de entrada y salida al método LPC, se muestran en las siguientes imágenes:

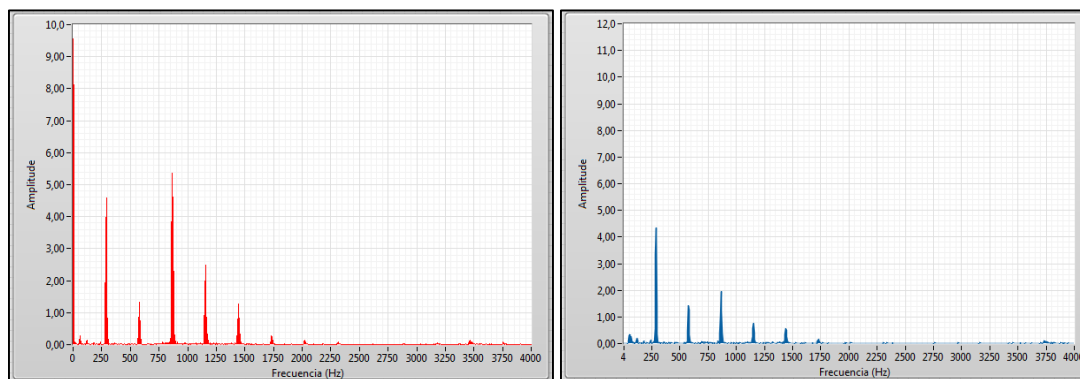


Figura 31. La gráfica de la izquierda muestra la FFT de la señal de voz original (señal de entrada) y la gráfica de la derecha muestra la FFT de la señal de salida del método LPC.

Como se muestra en las imágenes, las pérdidas de amplitud de los armónicos en la señal de salida son considerables, sobre todo en el armónico fundamental ubicado entre 750 Hz y 1000 Hz, esto puede significar que este método presenta una pérdida de información considerable, sobre todo al momento de inventanar la señal de voz en donde solo se le da peso a algunas secciones de la señal. Como se ve en la figura de la derecha, el ruido en baja frecuencia presente en la señal de entrada desaparece a la salida del método ya que al inventanar la señal con la ventana Hamming, la información en baja frecuencia se filtra al igual que la información en alta frecuencia.

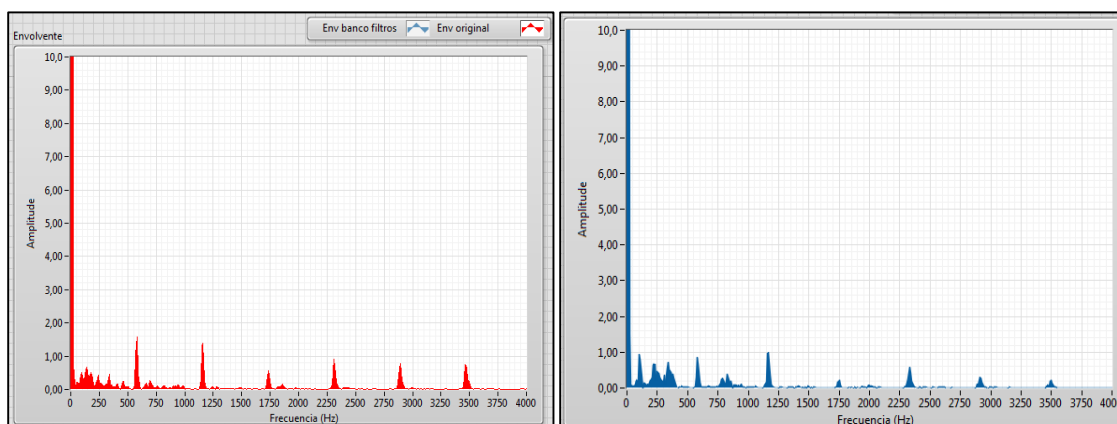


Figura 32. La gráfica de la izquierda muestra la FFT de la envolvente de voz original (señal de entrada) y la gráfica de la derecha muestra la FFT de la envolvente de la señal de salida del método LPC.

Por otro lado, en las FFT de las envolventes de la voz, aparecen los mismos armónicos en las dos gráficas, aunque nuevamente a la salida del método aparecen con menos amplitud. Para un mejor entendimiento de las diferencias que se presentan en las señales comparadas, se halla el error cuadrático en diferentes instantes de tiempo y se promedian los resultados para conocer qué tan alejado está el error cuadrático de cero. Se recuerda que cero es el valor que se obtiene en el error cuadrático medio cuando las señales son iguales.

Tabla 14. Error cuadrático medio de las señales de voz y de las envolventes.

Error cuadrático medio entre las señales de voz de entrada y salida	Error cuadrático medio entre las envolventes de entrada y salida
2,78	0,6
2,65	0,21
3,7	0,67
2,5	0,43
3,25	0,67
2,15	0,9
2,2	0,32

El promedio de los resultados de error cuadrático medio para la señal de la voz es de 2.74, siendo este el valor más alto entre los 2 algoritmos, por lo tanto, es el método de LPC el que presenta mayor pérdida de información durante el proceso de reconstrucción de voz y eso lo hace menos eficiente para este proyecto, en donde se requiere el método que obtenga la señal de salida con mayor similitud a la señal original. Por otro lado, el promedio de los valores obtenidos en el error cuadrático de las envolventes de la voz es de 0.54 y también es mayor que el obtenido con Spectrum y Banco de filtros. Es decir, que los valores de error cuadrático medio son mayores en LPC tanto para la voz, como para la envolvente.

Detección de envolvente

Para llevar a cabo el método de detección de envolvente descrito en el marco teórico, se utiliza un nodo MathScript en el diagrama de bloques de LabVIEW que permite realizar los

pasos descritos en Mathworks mediante programación basada en texto usando archivos .m. Este proceso de detección de envolvente es posterior a la reconstrucción de la señal de la voz como se muestra en el diagrama de flujo de todos los algoritmos, según se ve en los análisis espectrales anteriores, el proceso de detección de envolvente adhiere ruido de fondo a la señal de salida pues todas las señales de la envolvente presentan este ruido de baja frecuencia a pesar de que al inicio de Spectrum y Banco de filtros hay un filtro pasa banda que lo elimina y en LPC hay una ventana Hamming. Dirigirse al Anexo J para ver la programación en LabVIEW.

Proceso de modulación AM

Una vez se han realizado las mediciones y caracterización de todos los métodos para reconstrucción de la voz y se ha realizado la detección de la envolvente para cada uno se ha definido que el método con mejores resultados en las mediciones es el Banco de filtros (hecho que será discutido con mayor profundidad en el análisis de resultados) y es posible realizar la modulación AM entre la voz y el instrumento musical. Como se mencionó en la sección “Adquisición y procesamiento de señales en el sistema NI myRIO” de este capítulo, la señal del instrumento es adquirida por una entrada análoga MXP y debe cumplir con algunas características, las cuales son descritas a continuación:

Caracterización de la señal del sintetizador.

Inicialmente se propuso modular en amplitud la señal de una guitarra, mediante la envolvente de la voz. Sin embargo, la modulación AM debe satisfacer que la frecuencia de la señal portadora, que en este caso es la señal del instrumento, debe ser mayor a la componente frecuencial superior de la señal moduladora (Martín Fernandez, 2002). Ahora, en una guitarra eléctrica con una afinación estándar, las frecuencias no descienden por debajo de los 80 Hz ni superan los 2 kHz, aproximadamente. Por eso se ha realizado la modulación con un sintetizador

el cual puede generar señales de alta frecuencia. Para las pruebas se usó un sintetizador Korg de referencia X5D. A continuación se muestra el espectro en frecuencia de las notas (Do, Re, Sol, La) pertenecientes a la quinta y última octava del sintetizador Korg, las cuales se eligieron aleatoriamente entre todas las notas que componen la octava. Para hallar dicho espectro, la señal ha sido adquirida mediante una entrada análoga MXP del sistema embebido y luego se ha realizado la FFT sobre cada nota adquirida. Los resultados se muestran en las siguientes imágenes en donde el eje X es frecuencia en Hertz y el eje Y es amplitud en valores que representan el voltaje de entrada de la señal en escala lineal. Vale la pena resaltar que la escala escogida fue la quinta por ser la última octava del sintetizador y por lo tanto, es la que tiene los tonos más altos y se necesitaban componentes en alta frecuencia para la modulación, como ya fue mencionado.

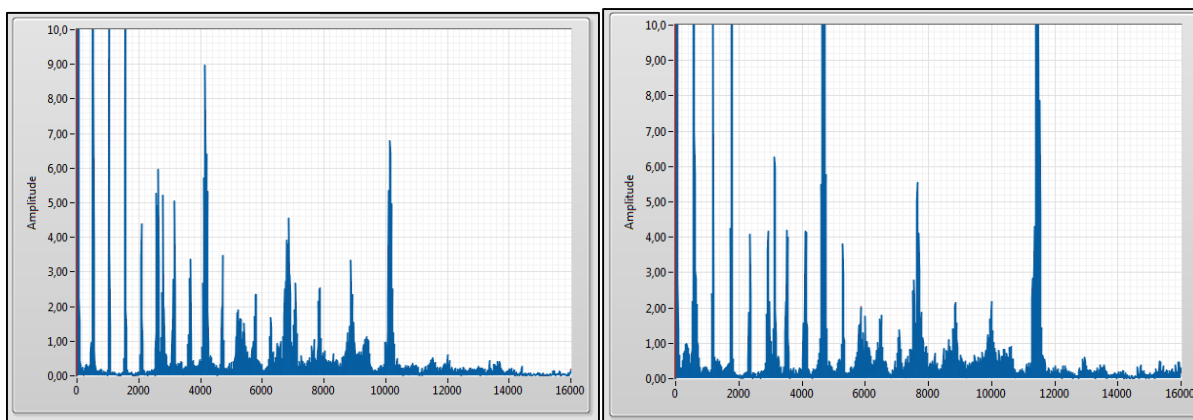


Figura 33. Espectro en frecuencia de la nota Do y la nota Re, respectivamente.

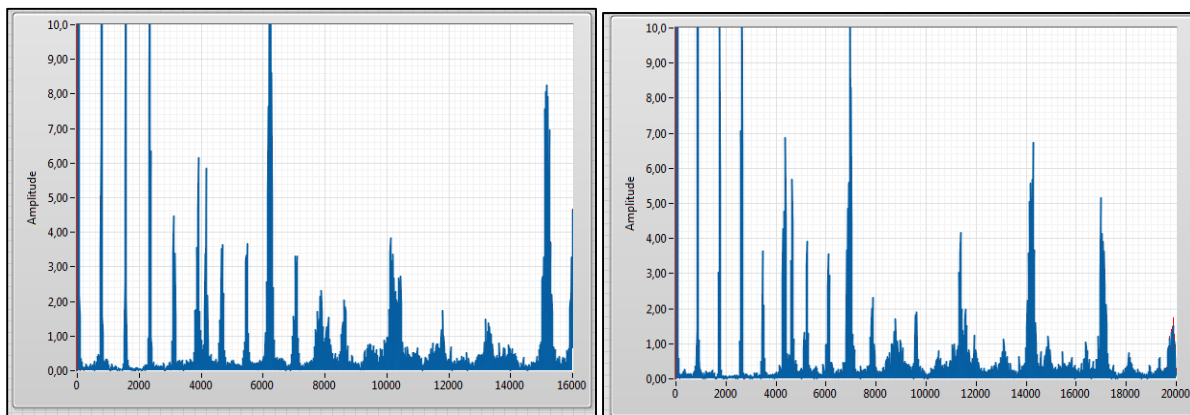


Figura 34. Espectro en frecuencia de la nota Sol y la nota La, respectivamente.

Según las pruebas anteriores, se determina que cada nota ejecutada tiene componentes de media y baja frecuencia, sin embargo a medida que sube la nota en la escala aparecen armónicos de alta frecuencia con gran amplitud casi llegando a 20000 Hz. La característica fundamental de la señal del instrumento es que debe tener componentes de alta frecuencia, al menos 10 veces superiores al componente frecuencia más alto presente en la señal de la voz. Para determinar si esto se cumple, se realizó el mismo análisis espectral a diferentes sonidos emitidos por una persona, la misma que fue utilizada para obtener la señal de prueba para las mediciones de los algoritmos. Los resultados de estas mediciones se muestran a continuación.

Caracterización de la señal de voz.

Para determinar si la componente frecuencial más alta presente en la señal de voz es 10 veces menor a la componente frecuencial más alta presente en el sintetizador se llevó a cabo la FFT de diferentes sonidos que emite la misma persona. A continuación se muestra la FFT sobre la señal de la voz y su envolvente para las vocales A y E (El análisis se realizó sobre todas las vocales, para ver las demás gráficas diríjase al Anexo T).

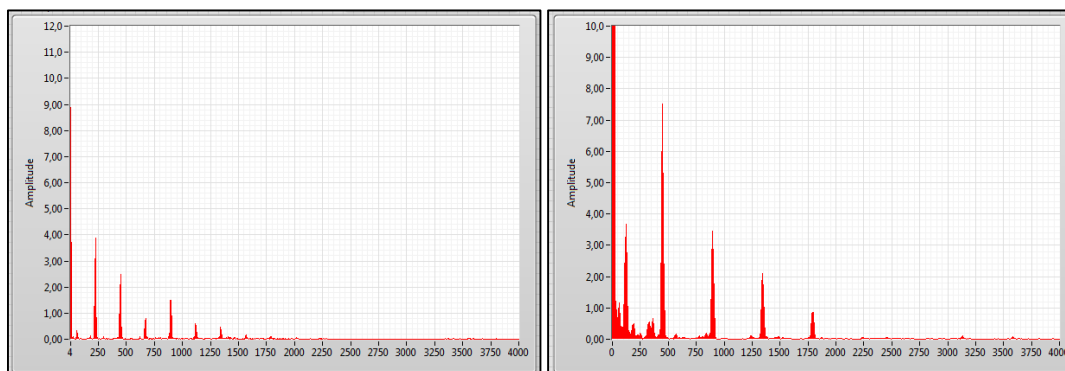


Figura 35. FFT de señal de entrada y su envolvente. La señal corresponde a la vocal "A".

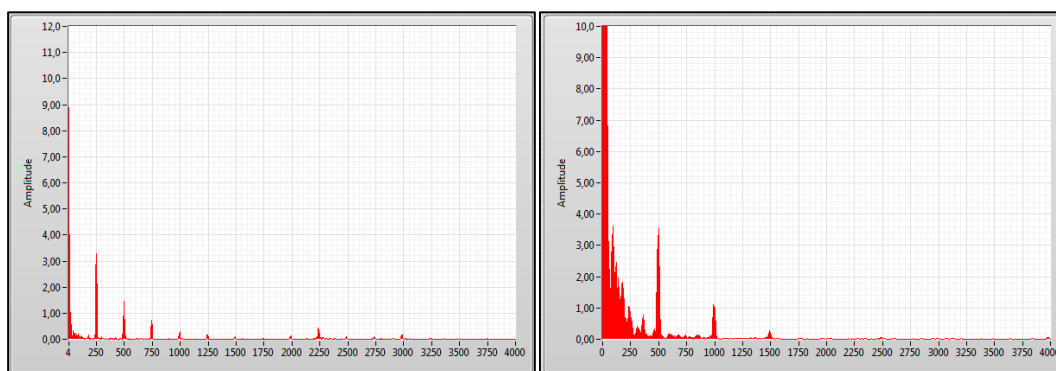


Figura 36. FFT de señal de entrada y su envolvente. La señal corresponde a la vocal "E".

Como se puede ver en las anteriores imágenes el ancho de banda de las vocales (excepto en la envolvente de la vocal A) llega hasta 1000 Hz y los armónicos de mayor amplitud se encuentran entre 250Hz y 500 Hz. La señal del sintetizador tiene componentes frecuenciales por encima de 10000 Hz y por ende el requerimiento para llevar a cabo la modulación en amplitud se cumple.

Como se ha mencionado antes, la envolvente de la señal tiene una pérdida de la información frecuencial pues en el método de detección de envolvente se adquieren las componentes en amplitud de la señal y por ello la envolvente ha perdido información. Por esta razón, la señal de la voz y su envolvente son completamente diferentes y el proceso de

modulación se puede llevar a cabo con cada una de ellas. Por lo tanto, en el proceso de modulación se han obtenido dos señales, una es resultado de la modulación entre la señal de voz y el instrumento y la otra es resultado de la modulación entre la envolvente de la voz y el instrumento.

Modulación AM en LabVIEW.

Una vez se han definido las señales que serán usadas en la modulación se procede a realizar el algoritmo para llevarla a cabo. El siguiente diagrama de bloques representa el algoritmo de modulación AM implementado en LabVIEW. Esta modulación AM, es una modulación normal en donde la señal modulada se transmite junto con la señal portadora (Ir a Anexo K para ver la implementación en LabVIEW).

Como ya fue mencionado, con el proceso de modulación AM se obtuvieron dos sonidos. El primero usaba la señal de la voz completa como señal moduladora y el segundo usaba su envolvente.

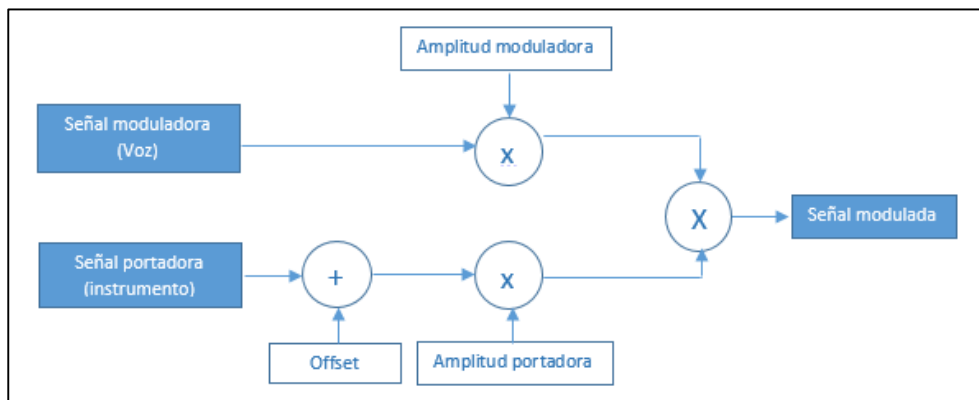


Figura 37. Diagrama de bloques de modulación AM realizada entre la voz e instrumento

Aplicación de encuesta para valoración subjetiva del sonido obtenido

Uno de los objetivos específicos de este proyecto establece que se debe realizar una encuesta a una muestra de músicos para determinar si el sonido les resulta agradable, además de

establecer una valoración subjetiva de la inteligibilidad de la palabra, calidez y brillo del sonido según la muestra seleccionada, con la finalidad de caracterizarlo.

Para cumplir con este objetivo, en primer lugar, se define el conjunto o muestra sobre la cual se realiza la encuesta y una delimitación geográfica. En un principio fue planteado aplicar la encuesta sobre una población de músicos de la ciudad de Bogotá, luego esta población se replantea pensando en que además de una población de músicos, sería útil una población de ingenieros de sonido que pueden hallar más aplicaciones posibles al sonido obtenido. Debe ser tenido en cuenta que los conceptos y conocimientos respecto a la producción musical o audiovisual cambian con el semestre en el que el estudiante se encuentra, por eso, se ha elegido una muestra de estudiantes de quinto semestre o inferiores y otra de estudiantes de sexto semestre o superiores y determinar si los resultados varían dependiendo del semestre que cursa el estudiante; también se busca determinar si sus resultados varían respecto a los de la población de músicos.

Es evidente que la muestra debe tener ciertas nociones y/o conocimiento de los conceptos relacionados con la producción musical y audiovisual, por ende, el tipo de muestra es no probabilística ya que depende de las características de la investigación; su elección no depende de una fórmula de probabilidad, sino de un proceso de toma de decisiones (Sampieri, N.F).

Según menciona Sampieri en su libro Metodología de la investigación, para las muestras no probabilísticas, el tamaño de la misma es irrelevante, ya que no interesa generalizar los resultados a toda la población, lo que se necesita en este caso es responder a las preguntas de la investigación. Para esta ocasión, las muestras están orientadas hacia la investigación cualitativa y se pueden clasificar como muestras diversas o de máxima variación, esto quiere decir, que con

ellas se busca obtener diferentes perspectivas e identificar diferencias o coincidencias para el caso estudiado (Sampieri, N. F). Según lo anterior se han definido las siguientes muestras:

- Estudiantes de ingeniería de sonido que se encuentra cursando sexto semestre o semestres superiores. Tamaño de la muestra: 20 individuos.
- Estudiantes de ingeniería de sonido que se encuentran cursando quinto semestre o semestres inferiores. Tamaño de la muestra: 20 individuos.
- Músicos de profesión. Tamaño de la muestra: 6 individuos.

Descripción de los audios utilizados en las encuestas.

Para realizar las encuestas, se graba un primer sonido obtenido con el dispositivo implementado, usando la voz como señal moduladora y un sintetizador como señal portadora; y se graba un segundo sonido usando la envolvente de la voz como señal moduladora y el mismo sintetizador como portadora. Los audios son grabados en el software Pro Tools como se muestra en el diagrama de flujo de la Figura 38.

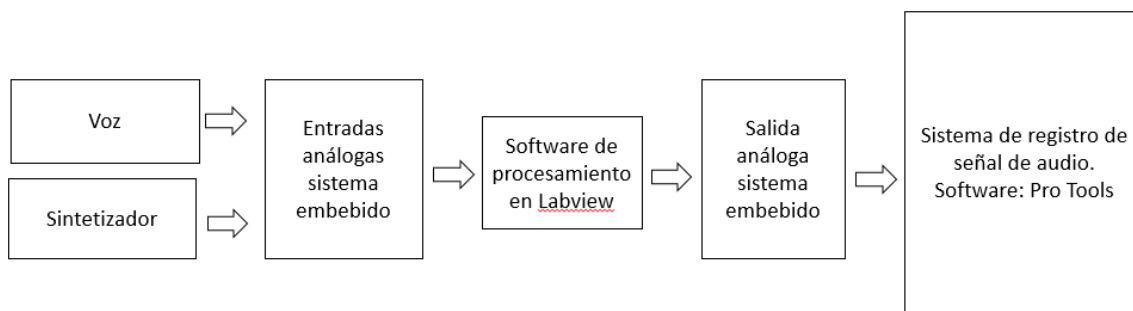


Figura 38. Diagrama de flujos de grabación del sonido para las encuestas.

Cada audio es utilizado en las encuestas y en los dos la frase mencionada es “Modulación AM”, de esa manera se puede comparar si alguno de los dos es inteligible o no.

Categorización de las variables de medición.

Como se menciona al principio de esta sección interesa evaluar la inteligibilidad, brillo y calidez del sonido obtenido, además de determinar si el sonido resulta agradable a los individuos encuestados con el objetivo de caracterizarlo. Y por último el agrado que genera el sonido sobre el individuo encuestado.

Las categorías establecidas son:

- **Agrado:** Para determinar si el sonido resulta agradable, se pregunta al individuo si lo usaría en la producción de una canción, en una producción audiovisual o en el diseño sonoro del videojuego. Si las respuestas son positivas se infiere que el sonido ha resultado agradable al individuo.
- **Inteligible:** Para determinar la inteligibilidad del sonido se pregunta al individuo que tan inteligible le parece la palabra mencionada en los audios y que lo represente en la escala de Likert.
- **Calidez:** Para determinar la calidez del sonido se pregunta al individuo que tan cálido percibe el audio y que lo represente en la escala de Likert.
- **Brillo:** Para determinar el brillo del sonido se pregunta al individuo que tan brillante percibe el audio y que lo represente en la escala de Likert.

Descripción del formulario generado para la encuesta.

Cada individuo encuestado resolvió la encuesta individualmente, usando los audífonos Shure 240A como transductor de salida para todos los casos. La persona encuestada debía escuchar cada audio y responder el formulario correspondiente. Dicho formulario se compone de dos partes, la primera parte para evaluar el sonido obtenido con la señal de la voz como moduladora y la segunda parte para evaluar el sonido obtenido con la envolvente de la voz como

moduladora, cada parte consta de 6 preguntas, las primeras 3 preguntas son cerradas y tienen dos opciones de respuesta: Si o No, estas tres primeras preguntas están enfocada a determinar si el sonido resulta o no agradable al individuo dependiendo de si lo usaría en algún proceso creativo dentro de la ingeniería de sonido (producción de una canción, desarrollo sonoro de un videojuego o diseño sonoro de un cortometraje, película y otra producción audiovisual). Las últimas 3 preguntas utilizan el escalamiento tipo Likert para evaluar los parámetros de brillo, calidez e inteligibilidad del sonido, la escala utilizada es la numeración de 1 a 5, siendo 1 lo menos inteligible, menos brillante o menos cálido y siendo 5 muy inteligible, muy brillante o muy cálido (Se puede consultar la encuesta aplicada en el Anexo N).

Capítulo 6

Presentación y análisis de resultados

Análisis comparativo de los algoritmos para determinar su eficiencia

Como se describió en el desarrollo ingenieril, se midieron ciertas características en los algoritmos con la finalidad de compararlos y elegir el que obtuviera mejores resultados. Las características medidas fueron: Latencia, número de bloques de código necesarios en la implementación, consumo de memoria RAM y similitud de las envolventes obtenidas con la envolvente original.

En la siguiente tabla se muestra el resultado de la medición de latencia para cada algoritmo, estableciendo que se realizaran 100 iteraciones.

Tabla 15. Resultados de medición de latencia de cada algoritmo.

Algoritmo	Tiempo 100 iteraciones (ms)	Tiempo por iteración (ms)
Banco de filtros	20556	205,56
LPC	20562	205,62
Spectrum	20795	207,95

De acuerdo a los resultados anteriores, el algoritmo que menos tiempo necesita para ejecutarse es el Banco de filtros con 205,56 ms por iteración, le sigue LPC con 205,62 y por último Spectrum con 207,95. Esta medida está relacionada directamente con el tiempo que tarda cada bloque que hace parte del algoritmo en realizar su función, hay algunos bloques específicos en cada algoritmo que son generadores de información, por ejemplo, en el método de Spectrum, estos bloques serían los que realizan la transformada de Fourier y la anti-transformada y en LPC sería el nodo de MathScript que contiene toda la información del algoritmo LPC y en el banco de

filtros serían los bloques de filtros FIR. Según los resultados anteriores, son estos últimos lo que realizan su procesamiento en menor tiempo.

Otra característica pensada para el análisis comparativo entre LPC, Spectrum y Banco de Filtros, fue el consumo de memoria RAM. Por ende, se ha definido la cantidad de memoria volátil usada por cada algoritmo; el análisis de esta variable es importante pues la rapidez de un algoritmo también es definida por el ajuste de este a la memoria disponible dado que en ella se almacena el algoritmo y se ejecutan los datos adquiridos. NI myRIO cuenta con una memoria DDR3 de 512 MB. En la tabla 16 se muestra el consumo de memoria producido por cada algoritmo.

Tabla 16. Resultados de medición de consumo de memoria RAM.

Algoritmo	Kilo bytes de memoria usada	Tamaño del VI (Instrumento Virtual) en el disco
Banco de filtros	3384,2K	486,3K
LPC	3648,2K	485,6K
Spectrum	3437,4K	483,5K

Los resultados muestran que el algoritmo con menor consumo de memoria es el banco de filtros con 3384,2 Kilo bytes de memoria usada. Por otro lado, el algoritmo con menos espacio ocupado en el disco es Spectrum. Aunque se puede observar que algoritmo necesita menos memoria, las diferencias entre cada algoritmo no son significativas; por ejemplo, el que menos memoria requiere es el Banco de filtros con 3384,2 Kbytes y el que más necesita es LPC con 3648,2 Kbytes y los dos resultados no están demasiado alejados entre sí. Además, se puede observar que el tiempo de ejecución es independiente de la cantidad de memoria RAM consumida por cada algoritmo. El consumo de memoria puede asociarse a la complejidad de los procesos que debe realizar cada algoritmo y la cantidad de datos que necesite para realizarlos.

También se analizó el número de bloques necesarios en la implementación de los algoritmos. Según la tabla 17 y los resultados anteriores, la cantidad de bloques usados no está relacionada con los bytes de memoria usados por el algoritmo, ni tampoco con su tiempo de ejecución. Sin embargo, el algoritmo con menor cantidad de bloques es Spectrum, y este también es el que ocupa menos espacio en el disco aunque esta es solo una especulación, por lo tanto el número de bloques se considera una medida irrelevante que no aporta a la caracterización de los algoritmos.

Tabla 17. Número de bloques necesarios en la implementación de cada algoritmo.

Algoritmo	Número de bloques
Banco de filtros	11
LPC	10
Spectrum	3

Por último se quiso determinar cuál de los métodos, obtiene la señal de salida con mayor similitud a la señal de entrada. En las siguientes imágenes se observa el análisis espectral realizado a la señal de voz usando la transformada rápida de Fourier para conocerla en el dominio de la frecuencia; en ellas se representa la amplitud de los armónicos de la voz en un instante de tiempo, antes y después de ser procesada por cada uno de los algoritmos. Para los tres casos, la gráfica azul es la representación de la señal procesada y la gráfica roja representa la señal de entrada que es la misma para cada algoritmo según se ha descrito en el Desarrollo Ingenieril. Para verificar si la diferencia entre las dos señales es o no apreciable, se calculó el error cuadrático medio en diferentes instantes de tiempo.

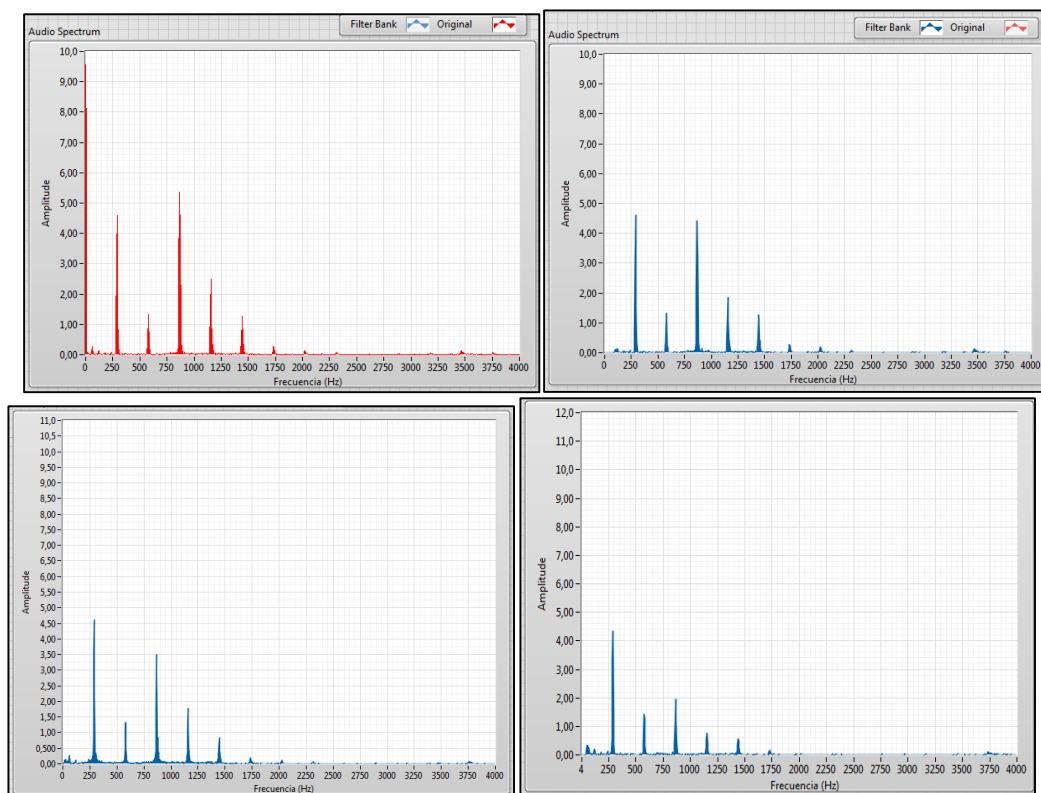


Figura 39. La gráfica superior izquierda muestra la FFT de la señal de voz original (señal de entrada), la gráfica superior derecha muestra la FFT de la señal de salida del banco de filtros, la gráfica inferior izquierda muestra la señal de salida del método Spectrum y la gráfica inferior derecha muestra la señal de salida del método LPC.

En la figura anterior se puede observar la FFT sobre las señales de voz a la salida de cada uno de los algoritmos, como fue mencionado en el desarrollo ingenieril, el banco de filtros, es el que obtiene una señal de salida con mayor similitud a la señal original, luego es el algoritmo que tiene menor pérdida de información, pues solo en los dos armónicos presentes entre 750 Hz y 1250 Hz existe un decaimiento de la amplitud que además no es significativo, considerando que la amplitud está en escala lineal. Ahora, respecto a la señal original, las señales de salida de Spectrum y LPC, presentan pérdidas de amplitud en todos los armónicos, sobre todo el LPC, y por lo tanto, se puede decir que es este algoritmo el que tiene mayor pérdida de información en el proceso de reconstrucción de la voz.

Tabla 18. Error cuadrático de las señales de entrada y salida para cada algoritmo.

Error cuadrático medio entre las señales de voz de entrada y salida para Banco de Filtros	Error cuadrático medio entre las señales de voz de entrada y salida para Spectrum	Error cuadrático medio entre las señales de voz de entrada y salida para LPC
0,009	0,08	2,78
0,03	0,9	2,65
0,001	0,07	3,7
0,04	0,05	2,5
0,01	0,22	3,25
0,009	0,23	2,15
0,2	0,17	2,2

El análisis realizado con las gráficas se reafirma al hallar el error cuadrático medio entre las señales. Se recuerda que el error cuadrático medio es la diferencia entre la señal de entrada señal original (que es la misma para los 3 algoritmos) y la señal de salida de cada algoritmo. Si las señales son completamente iguales, esta diferencia es igual a cero, así que entre más cerca estén de cero los resultados del error cuadrático medio, más similares son las señales analizadas. Según esto, Banco de filtros que obtiene los valores más cercanos a cero en diferentes instantes de tiempo según lo mostrado en la tabla anterior, es el método que obtiene una señal de salida con mayor similitud a la original. Y LPC, es el método que obtiene valores más alejados de cero en diferentes instantes de tiempo y por lo tanto es que obtiene una señal de salida con menos similitud a la original.

Para llevar a cabo el mismo análisis sobre la envolvente de la voz, se obtuvo la envolvente de la señal de entrada a cada algoritmo y la envolvente de la señal procesada. La transformada rápida de Fourier se realizó sobre las dos envolventes.

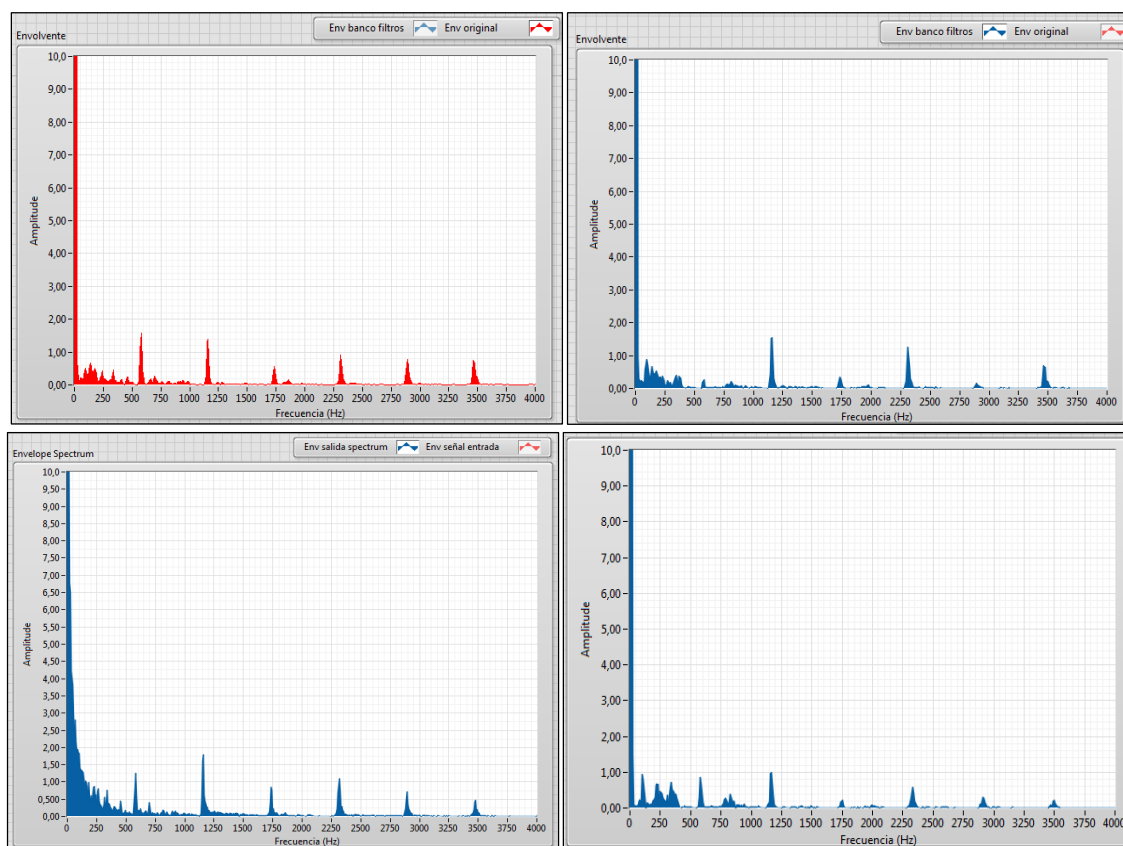


Figura 40. La gráfica superior izquierda muestra la FFT de la envolvente de voz original (señal de entrada), la gráfica superior derecha muestra la FFT de la envolvente de la señal de salida del banco de filtros, la gráfica inferior izquierda muestra la FFT de la envolvente de la señal de salida del método Spectrum y la gráfica inferior derecha muestra la FFT de la envolvente de la señal de salida del método LPC.

En las gráficas de la FFT de las envolventes de la voz, se observa principalmente que el ruido en baja frecuencia que aparece con la adquisición de la señal de la voz, permanece a pesar de que este ha sido filtrado al principio de cada algoritmo y por ende se destaca que el proceso de detección de envolvente hace aparecer este ruido de baja frecuencia.

A simple vista, parece que la envolvente con mayor similitud a la original es la obtenida con el método Spectrum, sin embargo las componentes en baja frecuencias aparecen con mayor amplitud y en un ancho de banda mayor en este método, así que es posible que en este caso no solo sea el proceso de detección de envolvente el que genere el ruido en baja frecuencia, sino también el proceso de reconstrucción de voz que es la transformada y anti-transformada de

Fourier. Por otro lado, las envolventes de LPC y Banco de Filtros tienen pérdidas significativas de amplitud en sus armónicos, es decir, que existe una pérdida de información generada durante el proceso.

Para finalizar este análisis se halla el error cuadrático entre la envolvente de entrada y las envolventes de salida justo como fue descrito anteriormente y se puede observar que los valores obtenidos para el Banco de Filtros y Spectrum son similares entre sí, sin embargo el menor valor sigue siendo el de Banco de filtros; Para LPC, aunque los valores no están tan alejados de cero como en la tabla 18, siguen siendo los valores mayores de los 3 algoritmos.

Tabla 19. Error cuadrático de las envolventes de las señales de voz de entrada y salida para cada algoritmo.

Error cuadrático medio entre las envolventes de la voz de entrada y salida para Banco de Filtros	Error cuadrático medio entre las envolventes de la voz de entrada y salida para Spectrum	Error cuadrático medio entre las envolventes de la voz de entrada y salida para LPC
0,26	0,3	0,6
0,26	0,2	0,21
0,16	0,5	0,67
0,17	0,2	0,43
0,26	0,34	0,67
0,27	0,32	0,9
0,2	0,3	0,32

Para finalizar, se concluye que el método de Banco de Filtros obtuvo menor tiempo de ejecución y menor consumo de memoria RAM. A su vez, Spectrum obtuvo menor consumo del espacio del disco y menor número de bloques de programación. Sin embargo, en el análisis espectral, la señal de salida con mayor similitud a la señal de entrada es la obtenida mediante el banco de filtros, tanto para el análisis de la señal de la voz completa como en el análisis de la envolvente. Se puede decir que Banco de Filtros y Spectrum tienen resultados muy parecidos, pero es Banco de Filtros el método con resultados más satisfactorios; por ende, el algoritmo elegido para ser implementado en el sistema es el Banco de filtros.

Presentación y análisis de resultados de encuesta realizada para valoración subjetiva del sonido obtenido

A partir de los resultados obtenidos mediante las encuestas se realiza el siguiente análisis que permite la caracterización del sonido obtenido con el dispositivo. Se recuerda que la encuesta se realizó para dos audios, el primero (Audio 1) usaba la señal de la voz como moduladora y el segundo (Audio 2) usaba la envolvente.

Resultado de encuesta realizada a la muestra de estudiantes de ingeniería de sonido que están cursando sexto semestre o superiores.

A continuación se presentan los resultados obtenidos con la realización de la encuesta a una muestra de estudiantes de ingeniería de sonido que se encuentran cursando sexto semestre o superiores. La tabulación de las encuestas se puede encontrar en el Anexo O.

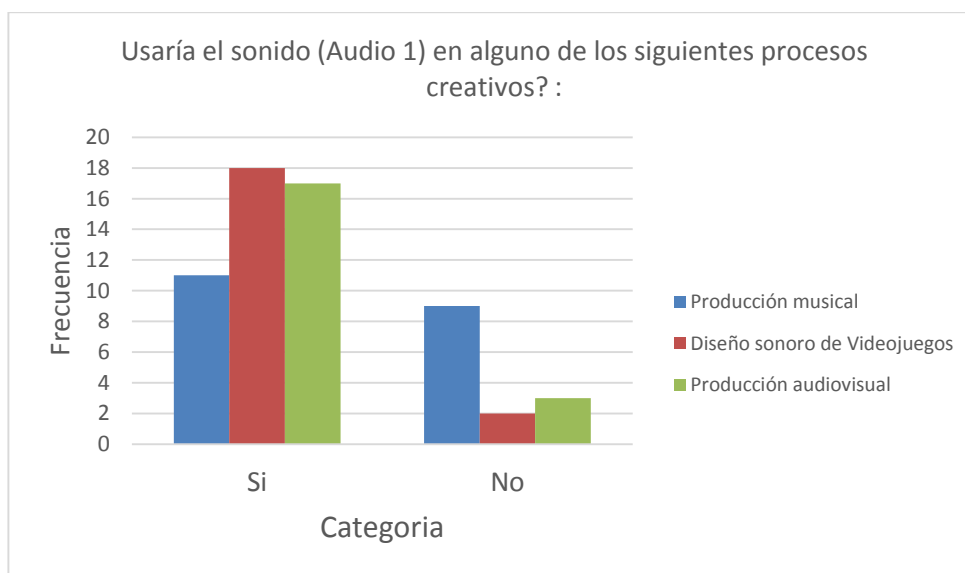


Figura 41. Gráfica de distribución de frecuencias.

Los resultados obtenidos en este caso fueron favorables para la categoría agradable. Se recuerda que determinar si el sonido resultaba agradable a la muestra encuestada era una de las variables propuestas en el tercer objetivo específico; con este fin se pregunta al individuo encuestado si usaría el sonido que está escuchando en la producción de una canción, en el diseño

sonoro de un videojuego o en el diseño sonoro de una producción audiovisual. Como se puede ver en la gráfica, los individuos de la muestra que Si usaría el Audio 1 en cada una de las producciones mencionadas, predomina sobre la cantidad de individuos que no lo usarían. Se concluye que la muestra de estudiantes de Ingeniería de Sonido de Sexto semestre en adelante, tuvo una reacción favorable hacia el primer audio y lo incluirían dentro de procesos creativos musicales y audiovisuales y que por lo tanto, lo consideran agradable.

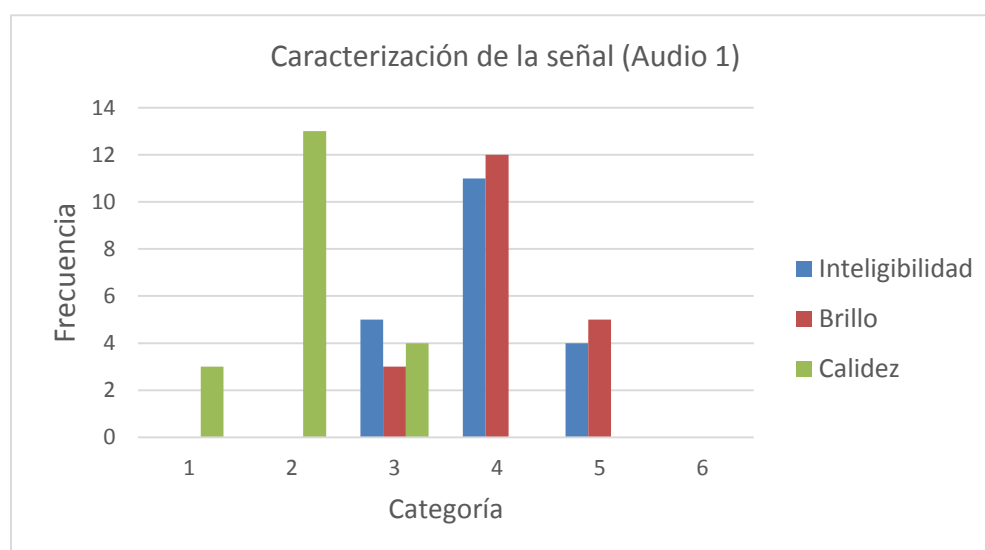


Figura 42. Caracterización del Audio 1

Según la Figura 42, la palabra mencionada en el Audio 1 resulta inteligible para la mayoría de la muestra. La categoría más repetida fue 4 en una escala de Likert de 1 a 5, donde 5 es Muy inteligible. El 50% de los participantes de la encuesta ha respondido 4 o por encima de 4 y nadie ha marcado 1 o 2 en la escala para calificar la inteligibilidad del Audio, por ende, los resultados se ubican en valores medios y altos y por lo tanto se concluye que la muestra ha calificado el sonido como inteligible.

El Audio 1 resulta brillante o rico en frecuencias altas para la muestra. La categoría más repetida en la escala fue 4, siendo 5 lo más brillante. El 50% de los participantes ha marcado 4 o por encima de 4 y los individuos se inclinan por un valor de 4.1 en la escala. Los valores marcados por toda la muestra están entre 3 y 5, por ende, nadie marcó los valores más bajos; por lo tanto, se concluye que la muestra ha calificado el sonido como brillante.

Por último, se determina que el Audio 1 no es cálido o rico en frecuencias bajas según la muestra. Para este caso, la categoría 2 fue la más seleccionada (13 individuos) y los individuos se inclinan por el valor de 2.05 en la escala. Según la Figura 42, nadie eligió categorías por encima de 3.

Las próximas figuras, muestran el resultado de las encuestas del Audio 2, es decir, el sonido obtenido de usar la envolvente como señal moduladora.

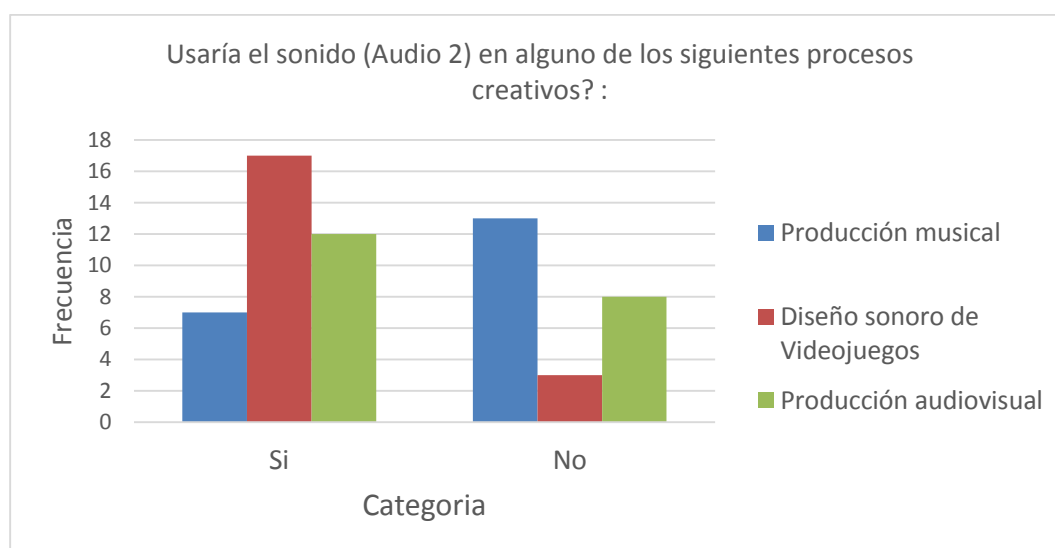


Figura 43. Gráfica de distribución de frecuencias.

La respuesta de los individuos encuestados fue desfavorable para la pregunta “¿Usaría el sonido en la producción de una canción?” donde el 65% de la muestra ha respondido que No. Sin embargo, el 85% de la muestra lo usaría en el diseño sonoro de videojuegos y el 60% lo usaría en

el diseño sonoro de una producción audiovisual. Así que aunque la muestra considera que no usaría el sonido obtenido con la envolvente de la voz en la producción de la canción, si lo usarían para diseño sonoro de videojuegos y audiovisuales. Puede que el sonido no haya resultado del todo agradable para la muestra, sobre todo para la rama de la producción musical, pero se rescata que la mayoría de la muestra lo usaría en la producción audiovisual o de videojuegos.

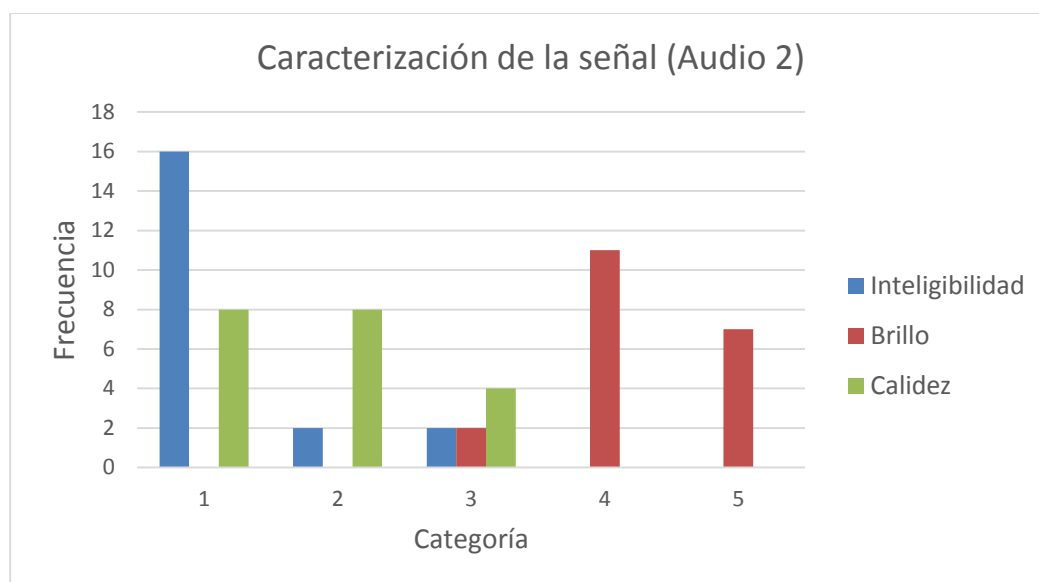


Figura 44. Caracterización del Audio 2.

Según la Figura 44, la muestra considera que el audio 2 no es inteligible. La categoría más repetida fue 1 que significa poco/nada inteligible y los resultados arrojan que los individuos se inclinan por el valor 1.3 dentro de la escala. El valor más alto para esta categoría fue 3 y el más bajo 1, pero nadie consideró valores altos de 4 o 5. De estos resultados se interpreta que el sonido no es inteligible para esta muestra.

Así mismo, la categoría seleccionada mayor cantidad de veces para calificar el brillo del sonido fue 4. 50% de los individuos de la muestra ha marcado 4 y la media indica que la

muestra se inclina por el valor de 4.25. Nadie calificó el sonido con valores de 1 o 2. Siendo 5 lo más brillante, el sonido se califica como brillante o rico en frecuencias altas.

Por último, de la Figura 44 se interpreta que el valor repetido mayor cantidad de veces es 1, el 50% de la muestra califica la calidez del sonido entre 2 y 3 y se inclina por un valor de 1.8 dentro de la escala. Es decir que la muestra percibe el Audio 2 como poco cálido o pobre en frecuencias bajas y este resultado coincide con la caracterización del brillo del Audio. Si es sonido es percibido como brillante, entonces es pobre en frecuencias bajas, esto significa que el brillo y la calidez son variables inversamente proporcionales.

Resultado de encuesta realizada a la muestra de estudiantes de ingeniería de sonido que están cursando quinto semestre e inferiores.

La encuesta se repitió para una muestra de estudiantes de ingeniería de sonido de primeros semestres y los resultados obtenidos se muestran a continuación.

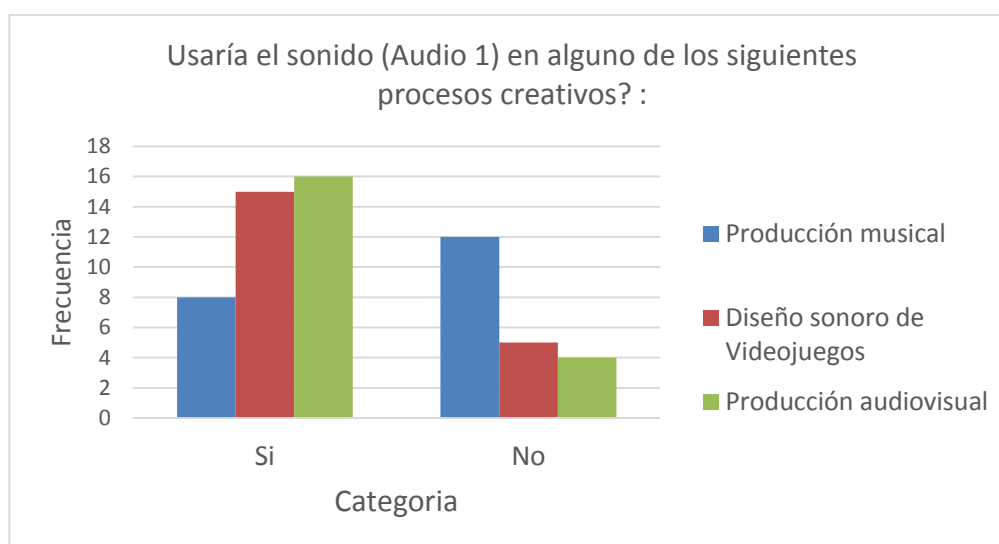


Figura 45. Gráfica de distribución de frecuencias

La muestra de estudiantes de Ingeniería de Sonido de quinto semestre e inferiores, ha respondido desfavorablemente a la pregunta “¿Usaría el sonido en la producción de una canción?”, pues el 60% (12 individuos) se inclinó por la categoría “No”. Sin embargo, 75% (15

individuos) de la muestra usaría el sonido en el diseño sonoro de videojuegos y 80% (16 individuos) usaría el sonido en el diseño sonoro de una producción audiovisual. Para este caso, el sonido ha sido percibido como agradable si se piensa en su aplicación para el sonido de un videojuego o para el sonido de una producción audiovisual, pero no tiene el mismo efecto si se piensa para aplicarlo en una producción musical.

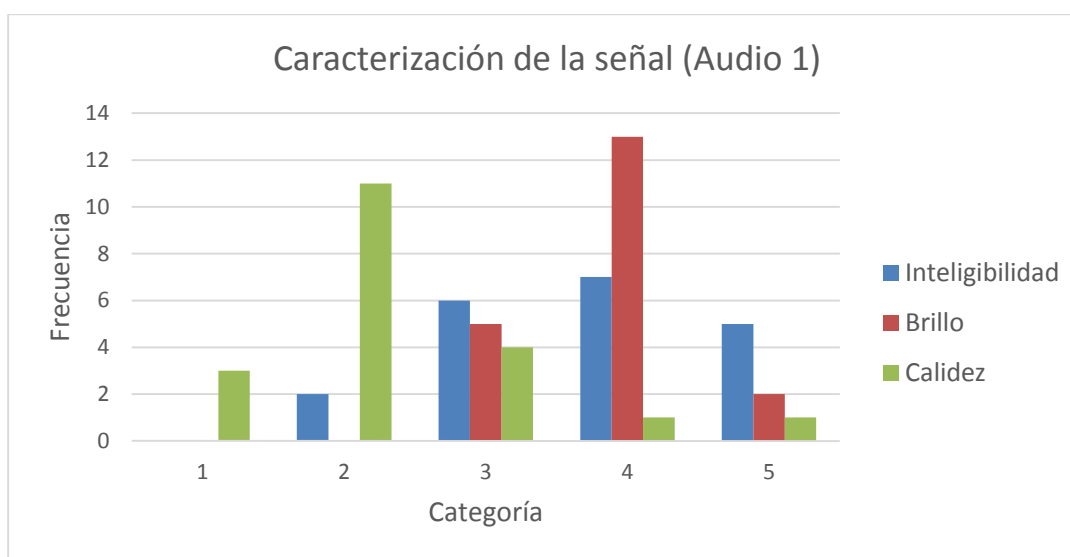


Figura 46. Caracterización de la señal (Audio 1)

Según la Figura 46, la muestra considera que el audio 1 es inteligible. La categoría más repetida para esta pregunta fue 4 y los resultados arrojan que los individuos se inclinan por el valor de 3.75 dentro de la escala de Likert.

Para la caracterización del sonido como brillante, el valor repetido mayor cantidad de veces para fue 4 y la media indica que la muestra se inclina por el valor de 3.85. Nadie calificó el sonido con valores bajos de 1 o 2. Siendo 5 lo más brillante, el sonido puede ser calificado como brillante o rico en frecuencias altas. Por último, se determina que el Audio 1 para la muestra de estudiantes de ingeniería de quinto semestre e inferiores, no es cálido o rico en frecuencias bajas.

Para este caso, el valor repetido mayor cantidad de veces fue 2 (11 individuos) y los individuos se inclinan por el valor de 2.3 en la escala.

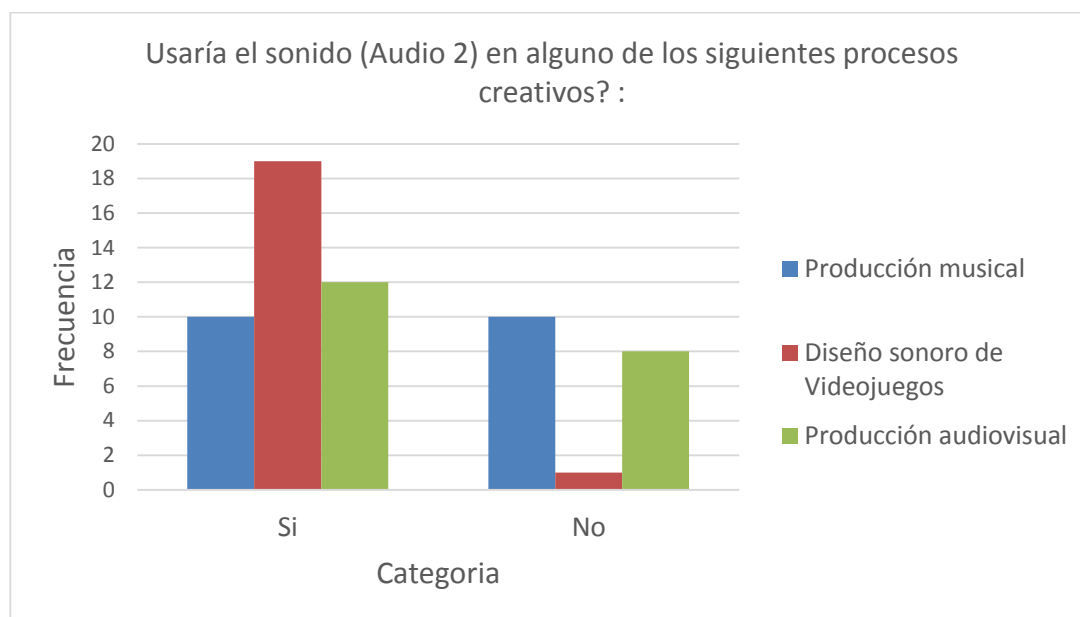


Figura 47. Gráfica de distribución de frecuencias

Para el segundo audio, los resultados de las encuestas han arrojado que el 50% de la muestra lo usaría en la producción de una canción, por ende, el otro 50% ha respondido que no lo usaría. De la misma manera que en la muestra anterior, los resultados han sido favorables en el uso del sonido para el diseño sonoro de un videojuego y el diseño sonoro de una producción audiovisual pues a estas preguntas, el 95% (18 individuos) y el 60% (12 individuos) respondió que si lo usarían, respectivamente. Es decir, que nuevamente el sonido es agradable si se piensa en aplicarlo en videojuegos o producciones audiovisuales, pero no se puede definir si lo es para usarlo en producciones musicales.

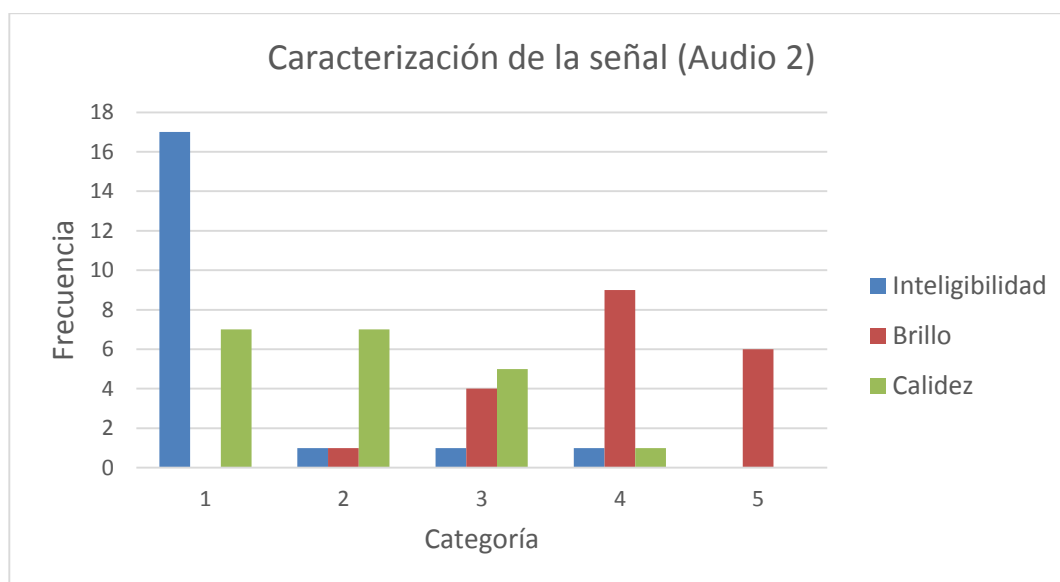


Figura 48. Caracterización de la señal (Audio 2)

Como se ve en la Figura 48, la muestra ha calificado el sonido como poco inteligible. La categoría repetida mayor cantidad de veces fue 1 y el valor por el que la mayoría de individuos dentro de la muestra se ha inclinado es 1.3. La diferencia para este caso es bastante notoria pues de 20 individuos, 17 eligieron 1 en la escala de Likert para calificar la inteligibilidad de la palabra. Los resultados para calificar el brillo del sonido, están alrededor de 4. Aunque el rango es de 3 y algunas personas se han inclinado por las categorías 2 y 3, la categoría más repetida fue 4, y se concluye que la muestra califica el sonido como brillante o rico en frecuencias altas. Finalmente, la muestra ha calificado el sonido como poco cálido. La categoría elegida mayor cantidad de veces fue 2 y la mediana indica que el 50% de la muestra se ha inclinado por las categorías 1 y 2 y según estos resultados se determina que los individuos dentro de la muestra se inclinan por la categoría 2, a pesar, que algunos hayan elegido 3 o 4 para calificar el sonido, y estos resultados, por ser tan pocos, son despreciables.

Resultado de encuesta realizada a muestra de músicos de profesión.

La última muestra evaluada estaba conformada por 6 músicos de profesión. A esta muestra se le preguntó únicamente si usaría el sonido en la producción de una canción, pues ninguno de los individuos participó antes en el diseño sonoro de videojuegos o audiovisuales.

El Audio 1 tuvo una reacción favorable, pues 5 individuos de 6 de la muestra se inclinaron por la categoría “Si”, lo cual determina que ellos usarían el sonido en la producción de una canción.

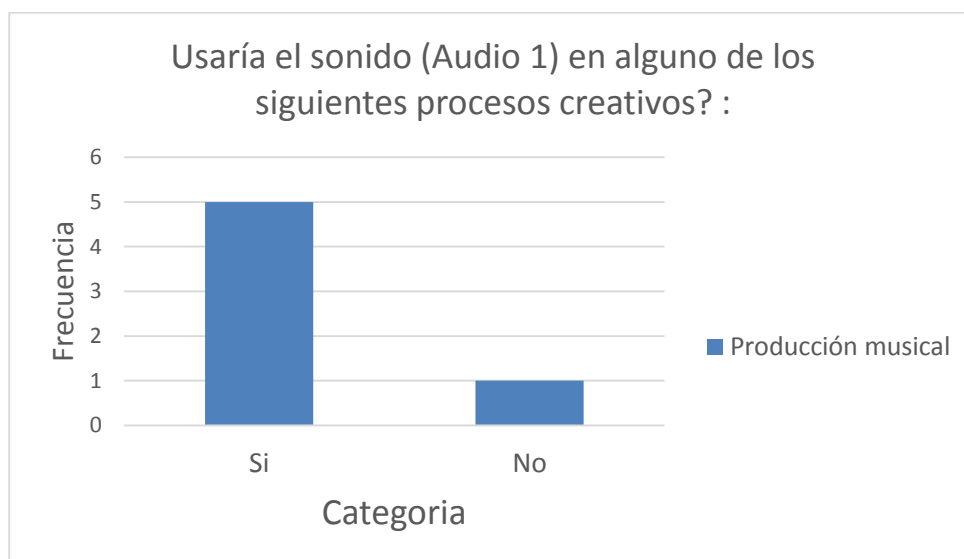


Figura 49. Gráfica de distribución de frecuencias

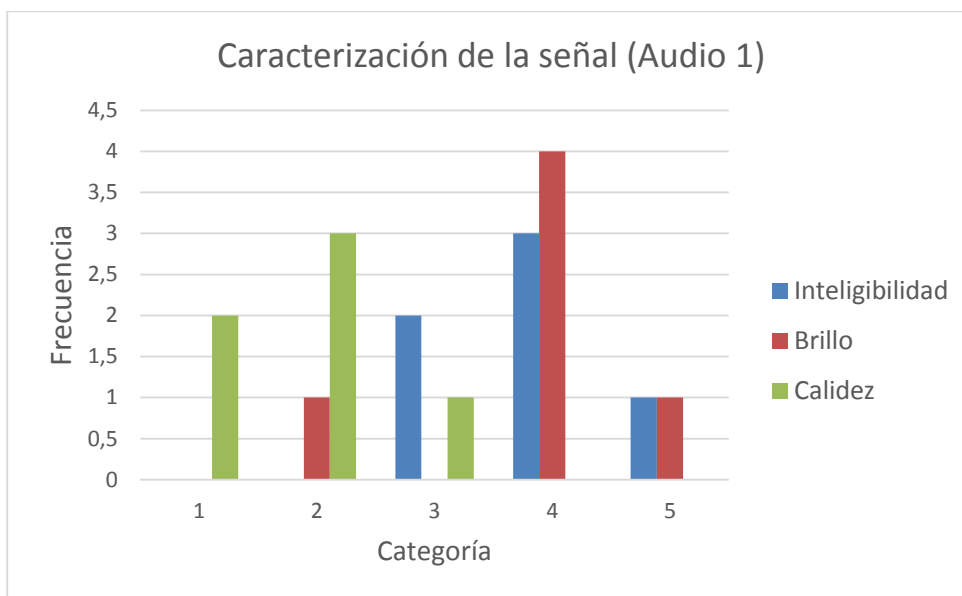


Figura 50. Caracterización de la señal (Audio 1)

La muestra de músicos ha elegido categorías entre 3 y 5 para calificar la inteligibilidad del sonido. Es decir que se inclinaron por valores medios y altos para calificarlo. La categoría respondida mayor cantidad de veces fue 4, y la media determina que la muestra se inclina por un valor de 3.83 para calificar la inteligibilidad del sonido. Esto significa que la muestra percibe el mensaje claramente. A su vez, la muestra ha calificado el Audio 1 como brillante, pues la mayoría de individuos se han inclinado por la categoría 4 y la muestra tiende a un valor de 3.83. Este resultado coincide con los obtenidos con las otras muestras, es decir que en definitiva el sonido es percibido como brillante. Dicho esto, la muestra de músicos, ha calificado el sonido como poco cálido, siendo 2 la categoría usada con mayor frecuencia, y siendo 1.83 el valor al que tiende la muestra.

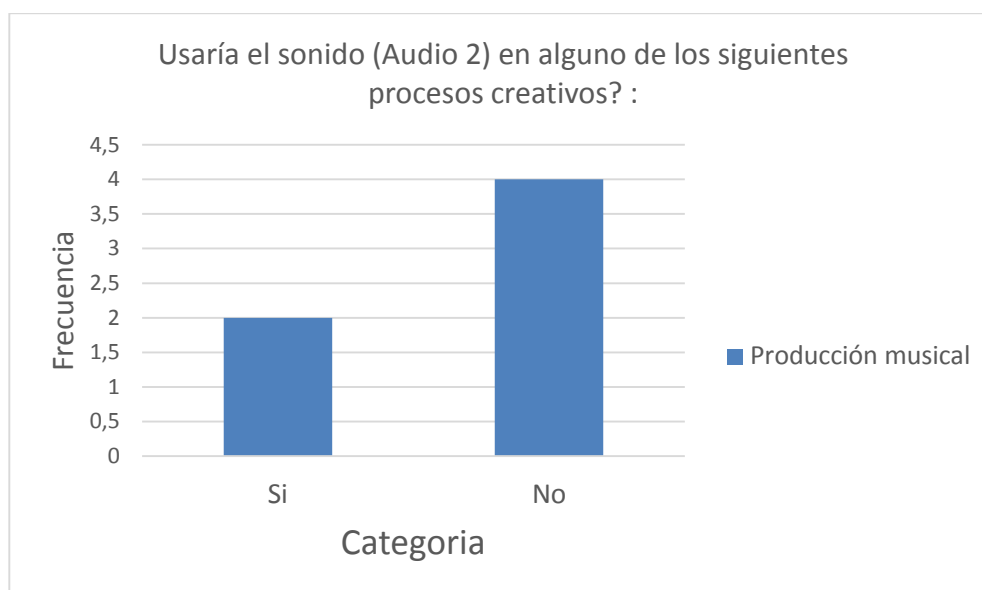


Figura 51. Gráfica de distribución de frecuencias

Para el Audio 2, los resultados han sido desfavorables pues el 67% (4 individuos) de los músicos han respondido que No usarían el sonido en la producción de una canción. Es decir que no lo perciben como agradable y no encuentran una aplicación de este dentro del proceso de producción musical.

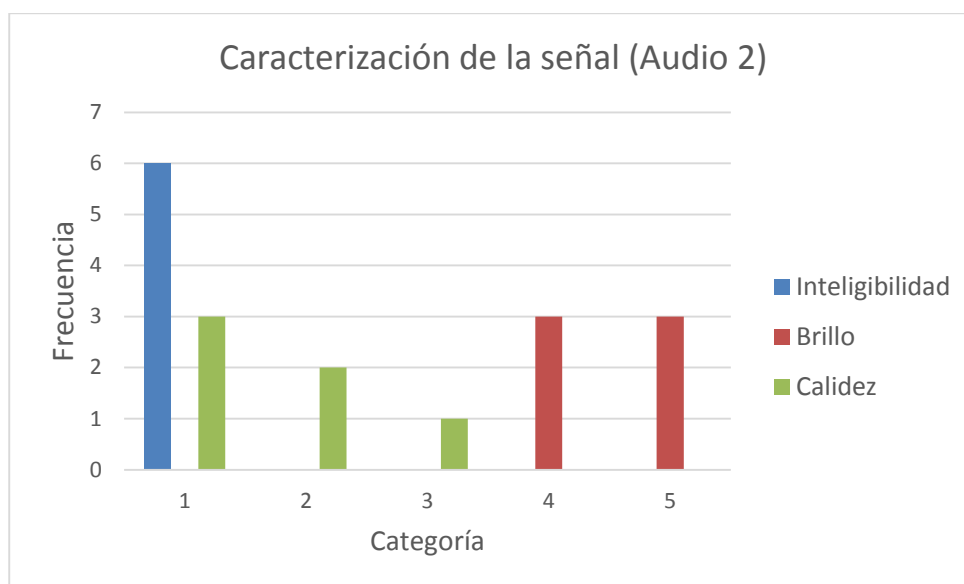


Figura 52. Caracterización de la señal (Audio 2)

En este caso, el resultado para calificar la inteligibilidad del sonido ha sido unánime, pues la totalidad de la muestra ha elegido la categoría 1, es decir, que la muestra califica el sonido con el menor puntaje para la inteligibilidad. Los resultados para calificar el Brillo del sonido han estado únicamente entre 4 y 5, siendo 4 la categoría repetida mayor cantidad de veces por los individuos. De esta manera, se determina que la muestra percibe el sonido como Brillante o rico en frecuencias altas. Los resultados para calificar la Calidez del sonido están entre las categorías 1 y 3, todos valores medios y bajos. La muestra se inclina por un valor de 1.6 para calificar la calidez y el valor repetido mayor cantidad de veces es 1. Por ende, se concluye que la muestra percibe el sonido como poco cálido o con ausencia en frecuencias bajas.

Conclusiones

El análisis de la medición de los algoritmos implementados, arroja las siguientes conclusiones:

- Según los resultados de las mediciones para cada algoritmo (LPC, Spectrum y Banco de Filtros) se observa que no hay diferencia significativa en los resultados de tiempo de latencia de cada uno. Los resultados de la medición del consumo de memoria RAM también son muy similares entre sí. Las mediciones coinciden en que el método con mejores resultados es el Banco de Filtros, pues es el que obtiene menos tiempo de latencia y menor cantidad de consumo de memoria RAM, sin embargo, LPC que tiene menos tiempo de latencia respecto a Spectrum no tiene menor consumo de memoria RAM que Spectrum; entonces, aunque la latencia está relacionada con el tiempo en que el procesador tarda en leer y escribir los datos a la memoria RAM y ejecutar las instrucciones que están almacenadas en ella, es posible que no haya una relación directamente proporcional entre la latencia y en el consumo de memoria de cada algoritmo, y que el consumo de memoria sea independiente del tiempo que tarda el procesador en adquirir los datos, escribirlos, leerlos y ejecutarlos.
- Teniendo en cuenta los resultados mencionados y que el número de bloques del banco de filtros es 11, el de LPC 10 y el de Spectrum es 3, se concluye que la cantidad de bloques que conforman el algoritmo no tienen relación con su consumo de memoria RAM o latencia; y por ende es una medida irrelevante que no aporta mayor información a la caracterización de los algoritmos. Sin embargo, vale la pena resaltar que el algoritmo que menos espacio ocupa en el disco es Spectrum con 483,5K y también es el que menos bloques de programación tiene y estos valores coinciden para los tres algoritmos, por lo

tanto, es posible que exista una relación entre el número de bloques y la memoria que el método ocupa en el disco.

- A partir de los resultados del análisis espectral realizado sobre la señal de salida y la señal de entrada a cada uno de los algoritmos, se concluye que el algoritmo que obtiene una señal de salida con mayor similitud a la señal de entrada es el Banco de Filtros, pues ha obtenido valores de error cuadrático medio más cercanos a cero en diferentes instantes de tiempo respecto a los otros algoritmos. Como se puede observar en el análisis de resultados, las diferencias entre las señales de salida de cada algoritmo y la señal de entrada, radica principalmente en las amplitudes de los armónicos, pues las señales de salida siguen teniendo los mismos armónicos que la señal de entrada para todos los casos, así que, de alguna manera, los métodos mantienen el contenido frecuencial de la señal de voz pero se presenta pérdida de información en las amplitudes.
- Las envolventes de salida de cada algoritmo, presentan diferencias más pronunciadas respecto a la señal de entrada, que en el caso de la señal de la voz completa. Para todos los algoritmos, la señal de salida presenta ruido de baja frecuencia que había sido eliminado al principio de cada algoritmo, es decir, que el método de detección de envolvente, adhiere ruido de baja frecuencia, lo cual es probable porque este método eleva las componentes de baja frecuencia para darle más peso sobre las componentes de alta frecuencia. A pesar de eso, se puede ver que para el método Spectrum, aparece un ruido de baja frecuencia con mayor ancho de banda que para LPC y Banco de filtros; por

ende, se puede concluir que la reconstrucción de la señal de voz mediante este método puede afectar la calidad de la señal.

El proceso de modulación AM entre la señal de voz y el instrumento arrojan las siguientes conclusiones

- La modulación AM se realizó con la señal de voz completa y con la envolvente como señales moduladoras. La diferencia entre los dos resultados, radica en que la señal de la voz contiene toda la información frecuencial, mientras que la envolvente solo tiene la información en amplitud de la señal. Por lo tanto, el sonido obtenido con la señal de la voz completa como moduladora es inteligible, mientras que el obtenido con la envolvente no. Sin embargo, la amplitud de la señal del instrumento varía de acuerdo a la amplitud de la envolvente que es lo que se quería lograr.

El análisis de las encuestas realizadas arroja las siguientes conclusiones:

- Los sonidos obtenidos mediante la modulación en amplitud (Audio 1 y Audio 2) fueron percibidos como brillantes para todas las muestras (Estudiantes de ingeniería de sonido cursando sexto semestre o superiores, estudiantes de ingeniería de sonido cursando quinto semestre o inferiores y músicos de profesión). Esto se relaciona con el hecho de que al realizar la modulación AM se transmiten tanto la señal portadora como la señal moduladora y teniendo en cuenta que los componentes en frecuencia de la señal portadora deben ser mucho mayores que los de la señal moduladora para que la envolvente de la señal resultante siga a la señal moduladora, el sonido obtenido es rico en frecuencias altas y medias.

- Para determinar el agrado generado por el sonido obtenido en la población encuestada, se ha preguntado a la población si usarían el sonido en algún proceso creativo musical o audiovisual. Los resultados arrojan que el sonido genera mayor aceptación si se piensa para ser usado en una producción audiovisual o un videojuego, sin embargo, la mayoría de la población piensa que no lo usarían en una producción musical. Esto puede significar que el sonido podría ser usado en géneros musicales muy específicos pero es difícil pensar en él como un procesamiento completamente aceptado para la música.
- El Audio 1 se clasifica como inteligible por todas las muestras. Por otro lado el Audio 2 se clasifica como poco inteligible; de manera que lo concluido anteriormente respecto al resultado de la modulación AM se confirma con las encuestas realizadas. La señal que se modula con la envolvente de la voz, pierde todo el contenido en frecuencia de la voz y solo se percibe una variación en amplitud del instrumento respecto al sonido que emita el usuario del dispositivo, pero no existe inteligibilidad del mensaje transmitido, por esto mismo es un sonido que puede ser utilizado con más facilidad en el diseño sonoro de un videojuego que en una producción musical en donde el entendimiento del mensaje transmitido tiene mayor importancia.

Recomendaciones

Para la caracterización subjetiva del sonido mediante encuestas, se recomienda realizar la captura de la señal de voz con diferentes tipos de micrófonos y a su vez, reproducir el sonido obtenido con diferentes transductores de salida para determinar si el resultado (Caracterización de brillo y calidez de la señal) se ve afectado por las características (Impedancia, respuesta en frecuencia, morfología, entre otros) de los transductores.

Para trabajos a futuro, se sugiere usar otro tipo de instrumentos musicales como señal portadora, siempre procurando que este genere frecuencias por encima de los 10 kHz, pues sería ideal que el dispositivo pudiera usarse con diferentes instrumentos. Así mismo, se recomienda utilizar otras señales de voz y realizar sobre ellas un análisis espectral para verificar su frecuencia fundamental y evaluar si es posible usar un instrumento que tenga mayor presencia de frecuencias medias y bajas para obtener un sonido que resulta más cálido en las pruebas subjetivas.

Se recomienda medir varios circuitos integrados durante el proceso de selección del circuito amplificador de la señal de voz, y comprobar las especificaciones dadas por el fabricante. Esto puede contribuir a una mejor elección del circuito integrado y por ende, a un mejor acondicionamiento de la señal de entrada.

Se sugiere integrar el sonido obtenido en una producción musical, audiovisual o en el diseño sonoro de un video juego, con la finalidad de comprobar empíricamente su facilidad de

integración a procesos creativos dentro de la ingeniería de sonido y la música.

Se recomienda realizar pruebas subjetivas a muestras de mayor tamaño con la finalidad de obtener más grados de libertad en la Prueba T y poder determinar con mayor precisión si los resultados obtenidos con cada muestra tienen diferencias significativas entre ellos o no.

Referencias

- Ambardar, Ashok (2002). Procesamiento de Señales Analógicas y Digitales. Editorial Thomson Learning. Segunda edición.
- Baez Dávila, Javier Alfredo (2006). Diseño de dispositivo para el reconocimiento de caracteres vocálicos, para ordenar comandos al televisor. Universidad de San Buenaventura. Bogotá, Colombia.
- Beltrán, Javier. Mejía, Marcela (2014). Desarrollo e implementación de un banco de pruebas de amplificadores de señal pequeña con válvulas al vacío y multitransformador. Universidad de San Buenaventura. Bogotá, Colombia.
- Bradbury, Jeremy (2000). Linear Predictive Coding. Obtenido de: http://my.fit.edu/~vkepuska/ece5525/lpc_paper.pdf
- Carballar, José A (1994). Software y Hardware de su PC. Editorial RA-MA, Madrid.
- Clot, Jaume & Hernando, Javier & Nadeu, Climent & Vallverdú, Francesc. Estudio comparativo y nuevas propuestas de técnicas de parametrización de la señal de voz para el reconocimiento del habla. Universidad politécnica de Cataluña. Barcelona, España.
- Colombini Gómez, Alejandro (2009). Hardware en la informática musical. Universidad politécnica de Valencia.
- Fernández, Marcos Martín (2002), Modulación en amplitud. Universidad de Valladolid.
- Hernández Sampieri, Roberto. Fernández Collado, Carlos (2010). Metodología de la investigación. Quinta edición. Mc Graw Hill.
- Lipiansky, Ed (2012). Ebedded systems hardware for software engineers. Mc Graw Hill.
- Martí, José (1987). Fast Fourier Transform como herramienta de análisis en fonética. Barcelona, España.
- Nashelsky, Boylestad Teoría de circuitos y dispositivos electrónicos. Editorial Pearson. Octava edición.
- Pallás Areny, Ramon (1993). Adquisición y distribución de señales. Editorial Marcombo. Barcelona, España.
- Proakis, Joshn G (1998). Tratamiento digital de señales. Tercera edición. Editorial Pretince Hall. UK.

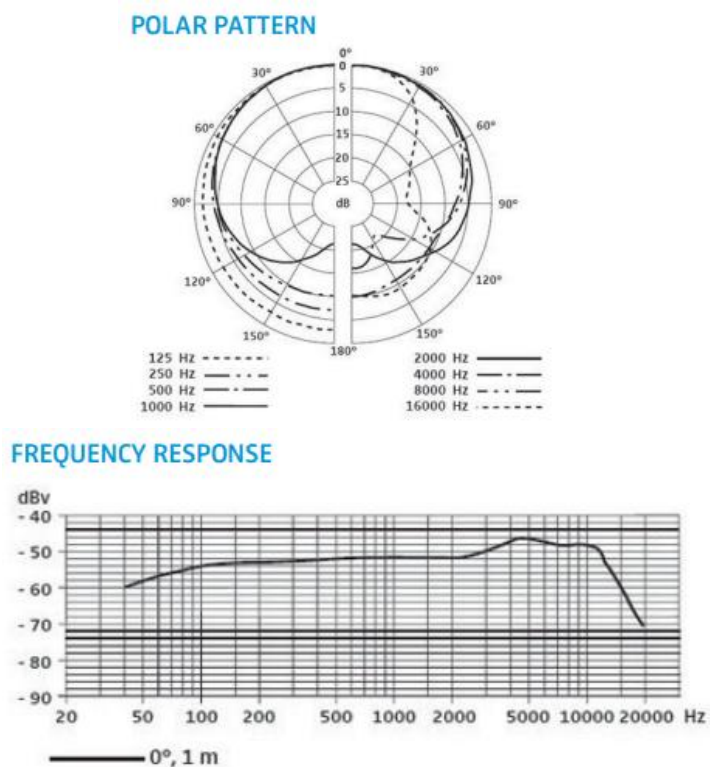
Ríos Ruiz, Wilson (2004). Aspectos legales del software libre o de código abierto. Obtenido de:
[http://www.acis.org.co/memorias/SalonInformatica/XXIVSalonInformatica/ASPECTOS
_LEGALES_DEL_SOFTWARE_LIBRE.doc](http://www.acis.org.co/memorias/SalonInformatica/XXIVSalonInformatica/ASPECTOS_LEGALES_DEL_SOFTWARE_LIBRE.doc).

Sinha, Priyabrata (2010). *Speech Processing in Embedded Systems*. Editorial Springer.

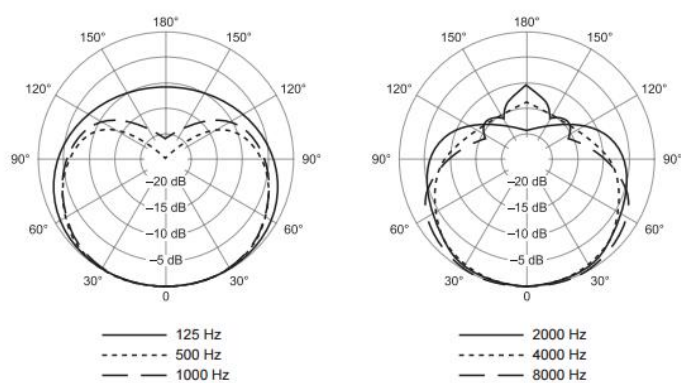
Zolzer, Udo (2002). *DAFX Digital Audio Effects*. Segunda edición. Editorial Wiley.

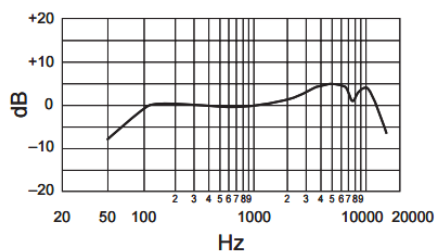
Anexos

Anexo A: Patrón polar y respuesta en frecuencia de Sennheiser e835 y Shure SM 58 respectivamente.



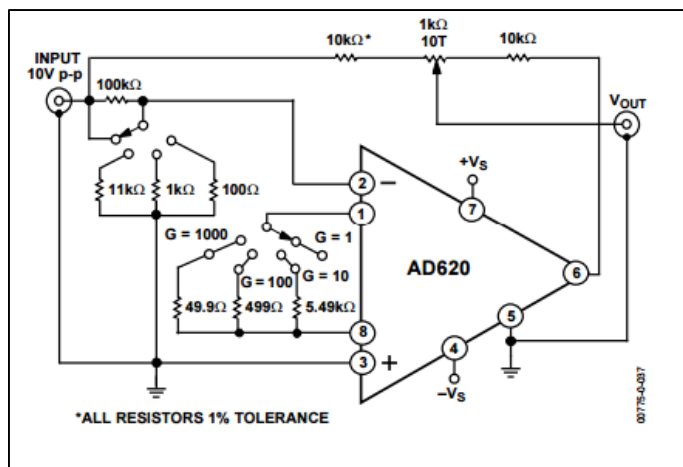
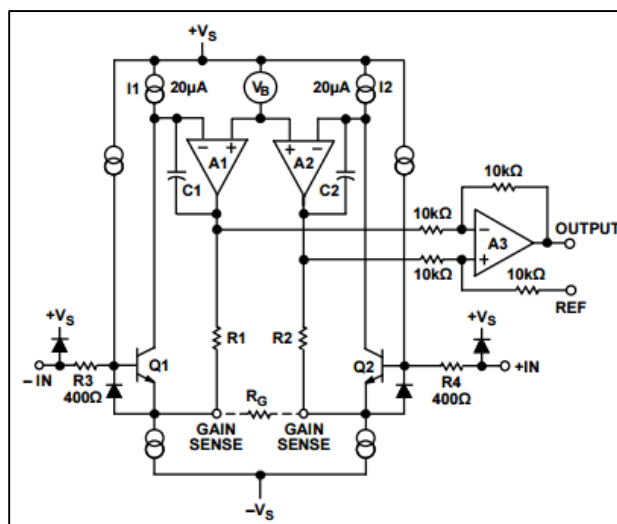
Recuperado de: http://en-de.sennheiser.com/downloads/download/file/3538/e_835_11_2012.pdf.

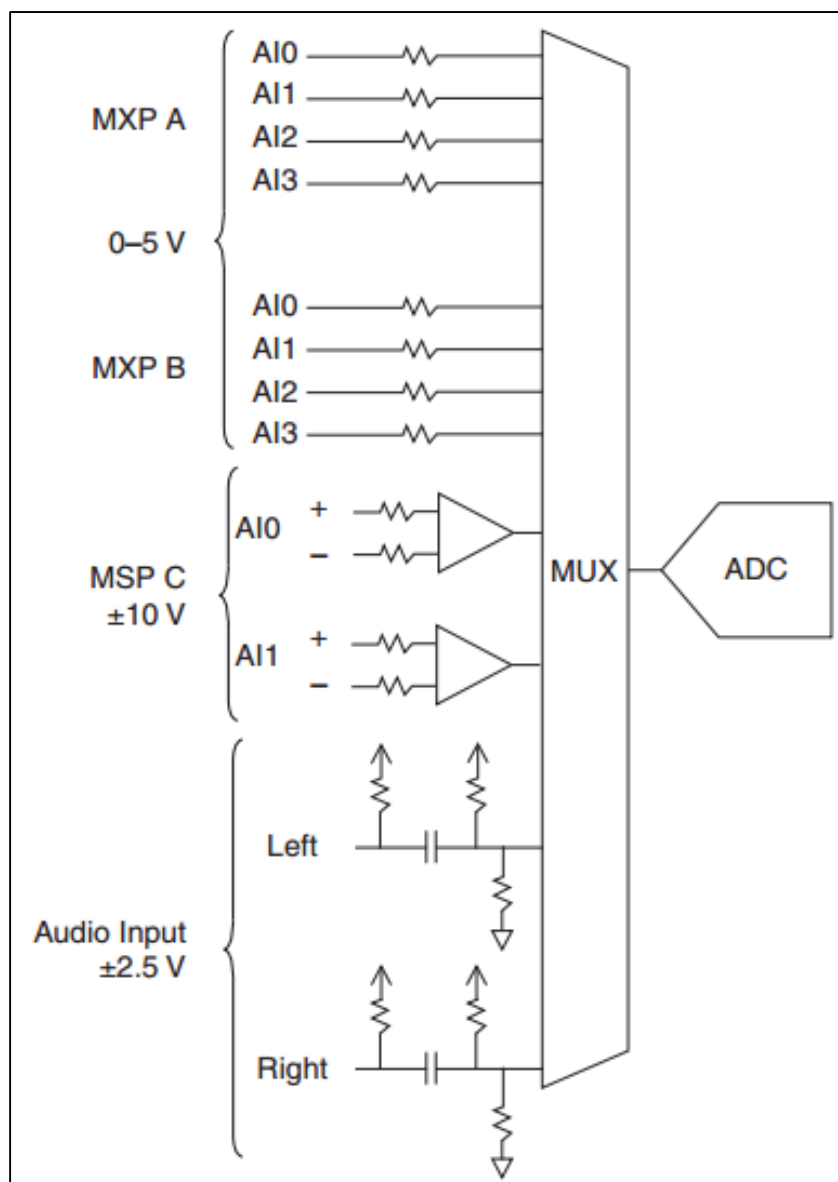




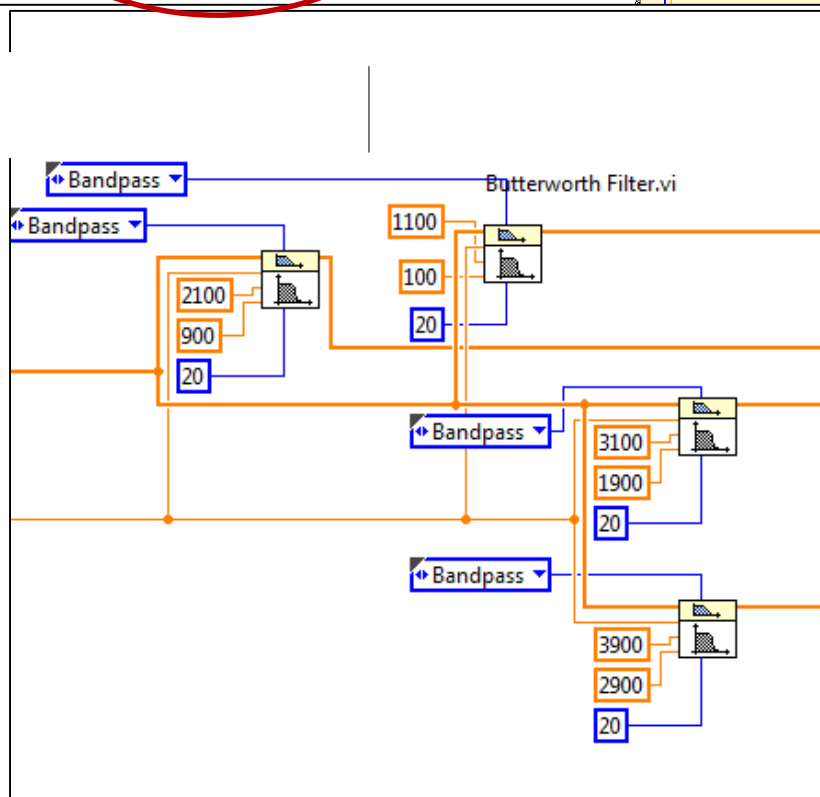
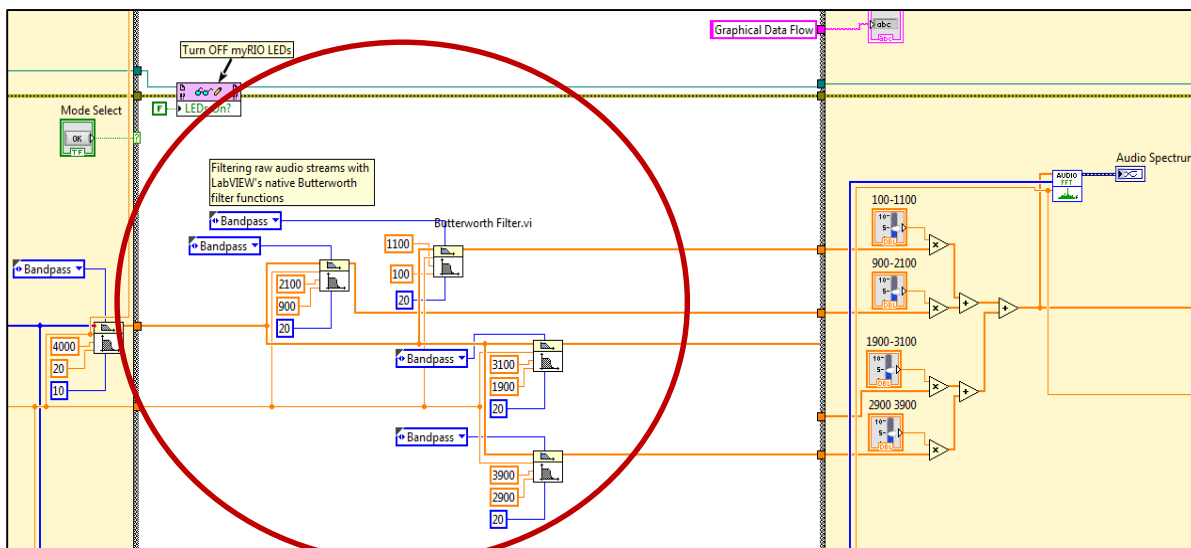
Recuperado de: http://cdn.shure.com/specification_sheet/upload/82/us_pro_sm58_specsheet.pdf.

Anexo B. Composición interna del integrado AD620.

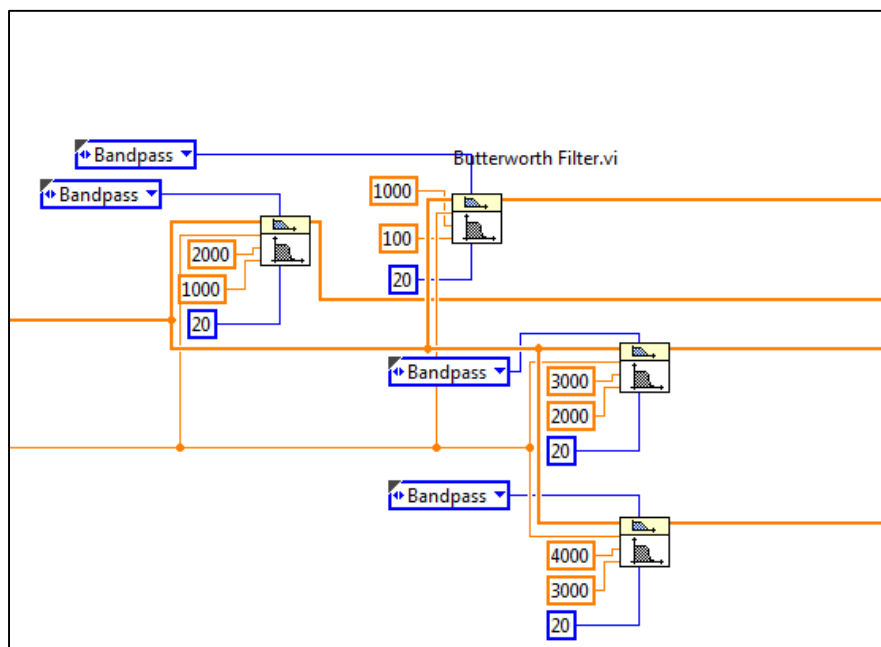


Anexo C. Circuito de entradas analógicas de NI myRIO-1900.

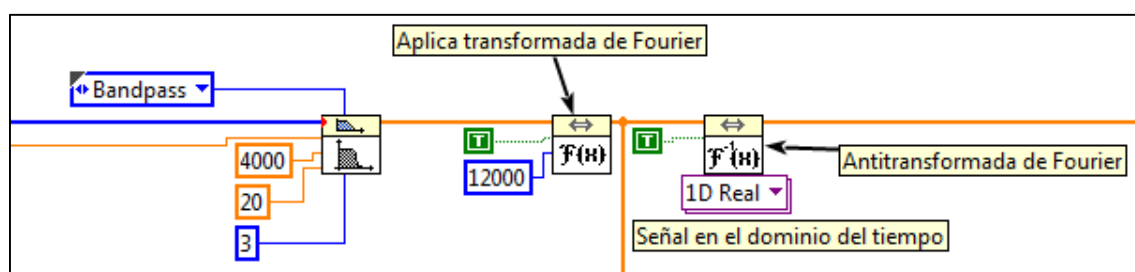
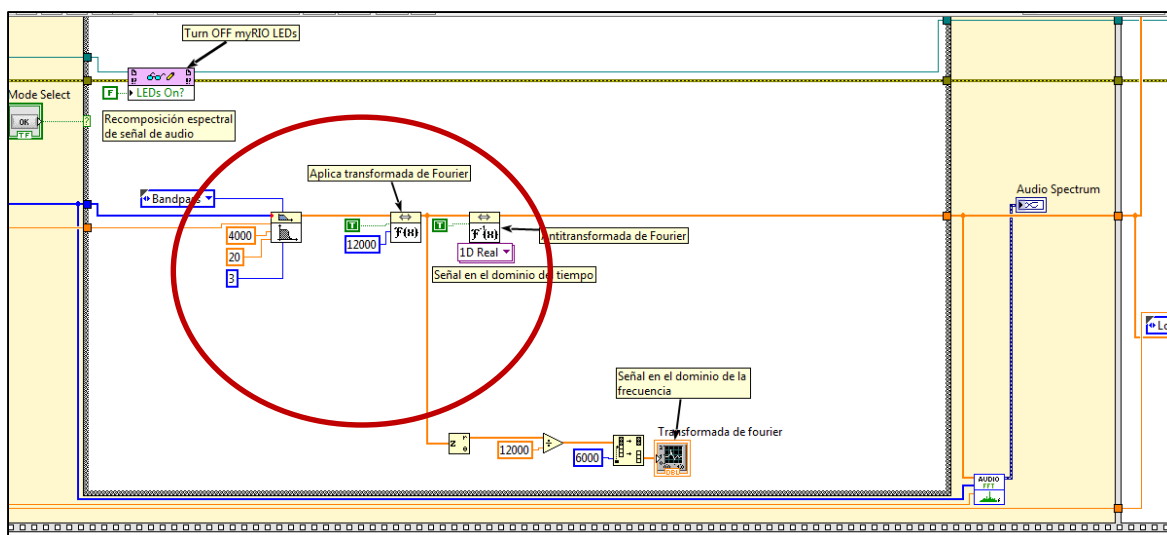
Anexo D. Implementación de banco de filtros en LabVIEW.



Anexo E. Implementación de banco de filtros en LabVIEW con respuesta en frecuencia plana.



Anexo F. Implementación de Spectrum en LabVIEW.



Anexo G. Implementación de LPC en MATLAB con método de Autocorrelación y Levinson Durbin.

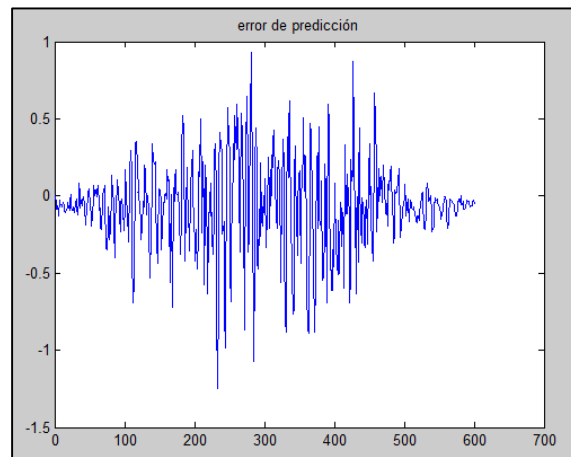
```

%% Adquirir señal
Fs=8000;
segsr=3;
recording=audiorecorder(Fs, 16, 1);
recordingblocking(recording,segsr);
A1=getaudiodata(recording);

%% Darle peso a una sección de la voz mediante una ventana
Anueva=A1(600:1200);
figure
plot(Anueva, 'r')
title('muestra de la señal original')
l=length(Anueva)
w=hamming(l);
x=w.*Anueva;
factor=0.01;
vmax=abs(max(x))
original= x/vmax;
figure
plot(original, 'g')
title('muestra de la señal original enventanada y normalizada')

%% Coeficientes LPC
R=xcorr(original,Anueva)
figure
plot(R, 'm')
title('correlacion del vector y el vector normalizado')
n=7;
[a,e] = levinson(R,n);
esperada = filter([0 -a(2:end)],1,original);
figure
plot(esperada, 'm')
title('salida esperada')
cant=length(esperada);
w=ones(cant,1); %Pesos
yesperada=esperada'*w;
e = (original - yesperada);
figure
plot(e)
error=mean(e)

```



Error de predicción entre la original y la señal esperada.

Anexo H. Implementación de LPC en MATLAB con función LPC.

```

Fs=8000;
segsr=3;
recording=audiorecorder(Fs, 16, 1);
recordblocking(recording, segsr);
A1=getaudiodata(recording);
sound(A1)
plot(A1)

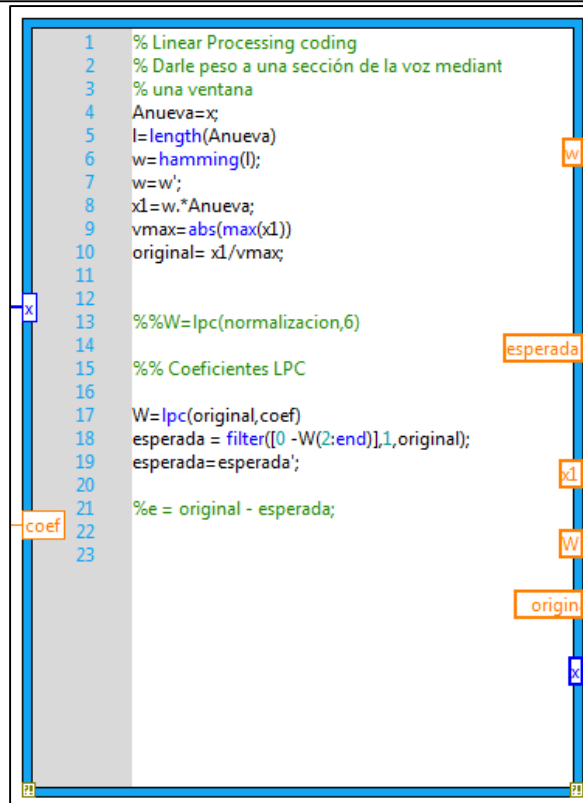
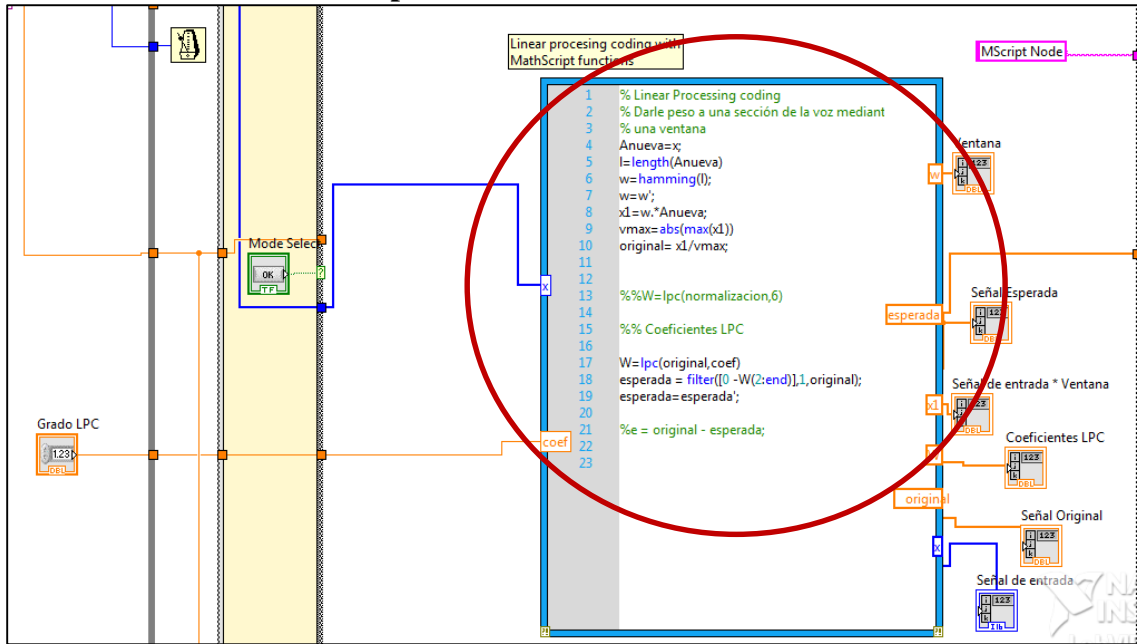
%% Darle peso a una sección de la voz mediante una ventana
Anueva=A1;
l=length(Anueva)
w=hamming(l);
x=w.*Anueva;
vmax=abs(max(x))
original= x/vmax;
plot(original, 'r')
hold on
title('Original')

%% Coeficientes LPC

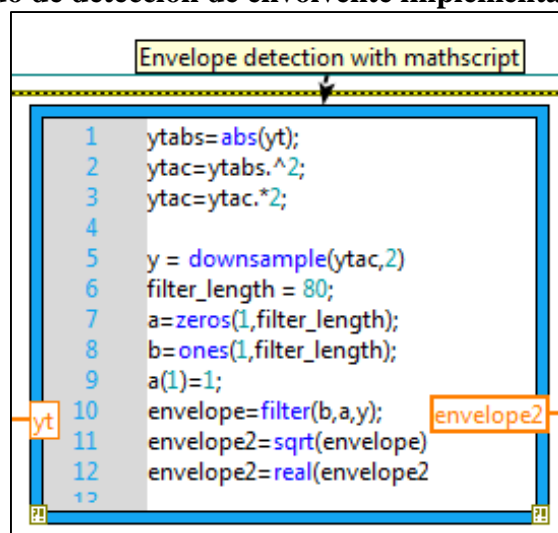
W=lpc(original,10)
esperada = filter([0 -W(2:end)],1,original);
plot(esperada, 'm')
hold on
title('Esperada')
e = original - esperada;
plot(e, 'g')
title('Error de predicción')

```

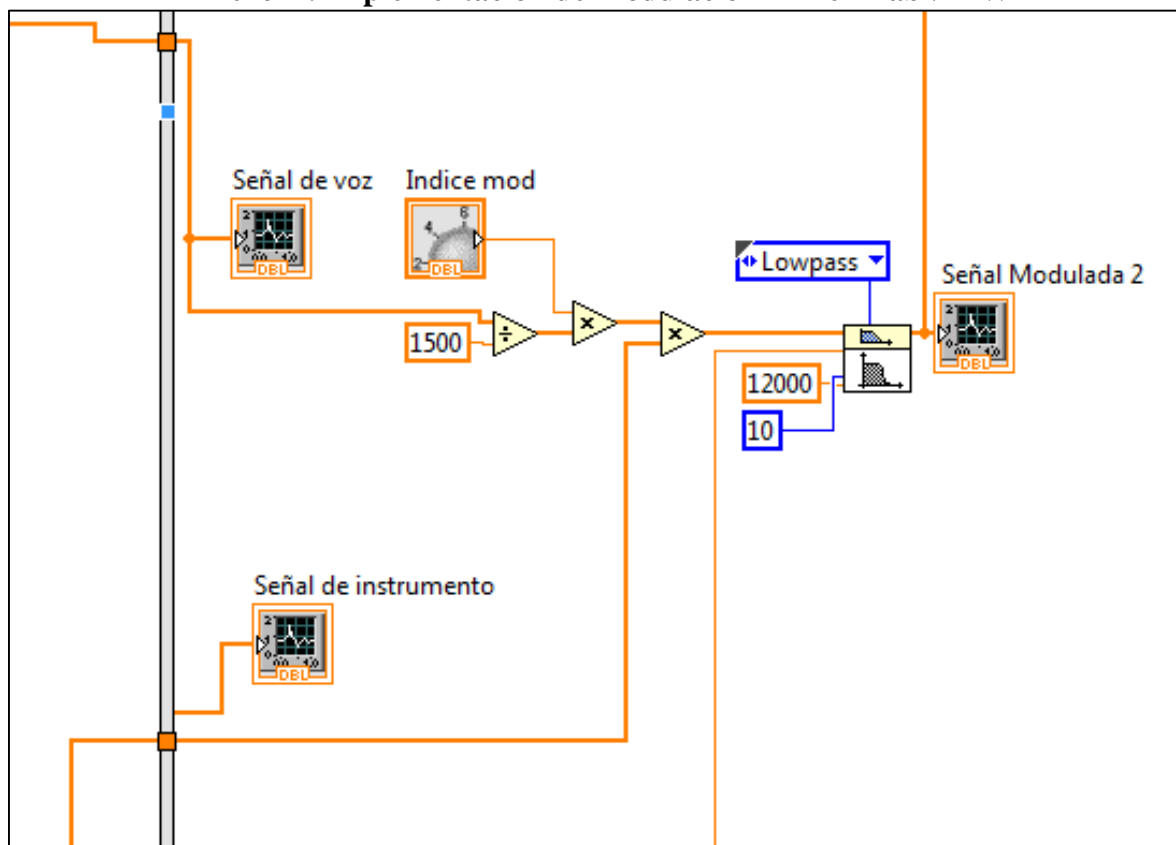
Anexo I. Implementación de método de LPC en LabVIEW

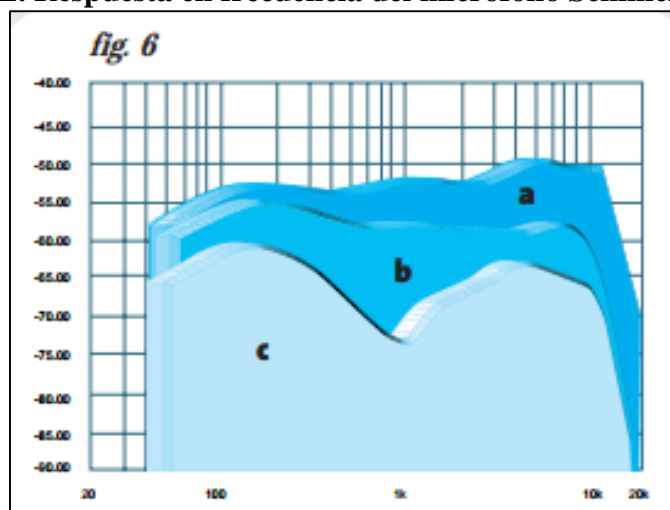


Anexo J. Método de detección de envolvente implementado en LabVIEW

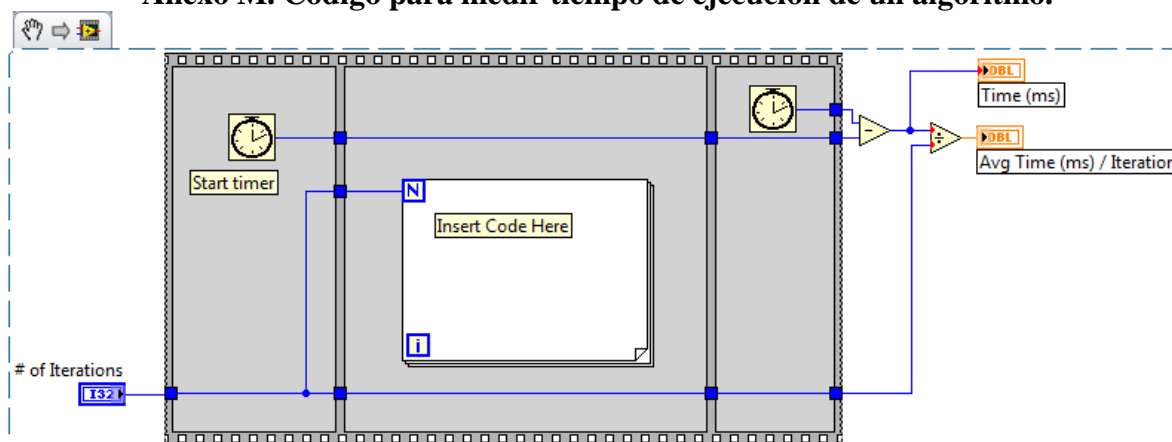


Anexo K. Implementación de Modulación AM en LabVIEW



Anexo L. Respuesta en frecuencia del micrófono Sennheiser e835

Recuperado de: <http://www.analfatecnicos.net/archivos/36.HablandoDeMicrofonos.pdf>.

Anexo M. Código para medir tiempo de ejecución de un algoritmo.

Obtenido de la página de National Instruments

Anexo N. Encuesta para determinar aceptación del sonido en una muestra de estudiantes de ingeniería de sonido y para determinar características subjetivas del mismo.

Encuesta para medir la aceptación del sonido obtenido mediante la modulación AM entre señal de la voz y un sintetizador.

Nombre: _____

Semestre: _____ Fecha: _____

La encuesta que usted está a punto de presentar es parte de los objetivos del proyecto de grado: *Desarrollo de un dispositivo por medio de un sistema embebido, que implemente el mejor método, obtenido del análisis comparativo entre LPC, Spectrum y Banco de filtros para modular la señal de un instrumento musical con la envolvente de la voz hallada*, del programa de Ingeniería de sonido de la Universidad de San Buenaventura. Los resultados serán utilizados con fines académicos únicamente.

En primera instancia usted escuchará el sonido obtenido mediante la modulación en amplitud de un sintetizador y la voz humana y por último, escuchara el sonido obtenido mediante la modulación en amplitud del mismo sintetizador y la envolvente de la voz. La modulación en amplitud o modulación AM consiste en la variación en amplitud de una señal portadora (sintetizador) con respecto a una señal moduladora (voz). Según su experiencia usted puede elegir o no responder las preguntas que se presentan a continuación.

Usted tiene libertad de decidir si quiere o no ser parte de esta investigación. Si está de acuerdo puede firmar y escribir su número de documento de identidad en la parte inferior de esta hoja. De antemano muchas gracias por su tiempo y colaboración.

Firma: _____

Documento: _____

Modulación entre sintetizador y voz

Para los siguientes enunciados, marque con una X su respuesta.

1. ¿Usaría el sonido escuchado en la producción de una canción?
 - a. Si
 - b. No

2. ¿Usaría el sonido escuchado en el diseño sonoro de un videojuego?
 - a. Si
 - b. No

3. ¿Usaría el sonido escuchado en el diseño sonoro de un cortometraje, película u otro tipo de producción audiovisual?
 - a. Si
 - b. No

Para los siguientes enunciados marque con una X el cuadro que corresponda a su experiencia, siendo 1 el resultado más bajo y 5 el resultado más alto.

4. ¿La palabra mencionada en el audio es inteligible? (La inteligibilidad se refiere al correcto entendimiento del mensaje que se está transmitiendo).

1	2	3	4	5

5. Qué tan brillante considera al sonido escuchado (Tenga en cuenta que el brillo se refiere a una presencia dominante de frecuencias altas).

1	2	3	4	5

6. Qué tan cálido considera al sonido escuchado (Tenga en cuenta que la calidez se refiere a una presencia dominante de bajas frecuencias).

1	2	3	4	5

Modulación entre sintetizador y envolvente de la voz

Para los siguientes enunciados, marque con una X su respuesta.

7. ¿Usaría el sonido escuchado en la producción de una canción?
- c. Si
d. No
8. ¿Usaría el sonido escuchado en el diseño sonoro de un videojuego?
- c. Si
d. No
9. ¿Usaría el sonido escuchado en el diseño sonoro de un cortometraje, película u otro tipo de producción audiovisual?
- c. Si
d. No

Para los siguientes enunciados marque con una X el cuadro que corresponda a su experiencia, siendo 1 el resultado más bajo y 5 el resultado más alto.

10. ¿La palabra mencionada en el audio es inteligible? (La inteligibilidad se refiere al correcto entendimiento del mensaje que se está transmitiendo).

1	2	3	4	5

11. Qué tan brillante considera al sonido escuchado (Tenga en cuenta que el brillo se refiere a una presencia dominante de frecuencias altas).

1	2	3	4	5

12. Qué tan cálido considera al sonido escuchado (Tenga en cuenta que la calidez se refiere a una presencia dominante de bajas frecuencias).

1	2	3	4	5

Anexo O. Tabulación de encuestas realizadas a una muestra de estudiantes de ingeniería de sonido de la Universidad de San Buenaventura.

Muestra de Estudiantes de ingeniería de sonido (sexto semestre y superiores)							
Audio 1							
	Semestre	Musical	Videojuego	Audiovisual	Inteligibilidad	Brillo	Calidez
1	8	No	No	No	3	4	3
2	8	No	Si	No	3	4	2
3	10	Si	Si	Si	5	4	2
4	9	Si	Si	Si	4	3	2
5	9	Si	Si	Si	4	5	2
6	9	Si	No	No	3	5	2
7	9	No	Si	Si	4	4	2
8	9	Si	Si	Si	4	4	2
9	6	No	Si	Si	5	3	2
10	6	Si	Si	Si	3	4	1
11	9	Si	Si	Si	4	3	1
12	8	Si	Si	Si	4	4	2
13	10	Si	Si	Si	5	4	1
14	10	No	Si	Si	3	4	3
15	9	Si	Si	Si	4	5	2
16	8	No	Si	Si	4	4	2
17	9	Si	Si	Si	4	5	2
18	9	No	Si	Si	4	5	3
19	9	No	Si	Si	4	4	3
20	8	No	Si	Si	5	4	2

Muestra de Estudiantes de ingeniería de sonido (sexto semestre y superiores)							
Audio 2							
	Semestre	Musical	Videojuego	Audiovisual	Inteligibilidad	Brillo	Calidez
1	8	No	No	No	1	4	1
2	8	Si	Si	Si	1	4	1
3	10	No	Si	No	1	5	3
4	9	No	Si	Si	1	4	1
5	9	No	Si	Si	1	4	2
6	9	No	Si	No	1	5	1
7	9	No	Si	Si	3	5	2
8	9	No	Si	Si	2	4	2
9	6	No	Si	Si	1	5	1
10	6	No	No	No	1	5	1
11	9	Si	Si	Si	1	4	1

12	8	Si	Si	Si	2	3	3
13	10	No	No	No	1	5	2
14	10	No	Si	Si	3	3	3
15	9	Si	Si	No	1	4	2
16	8	No	Si	No	1	4	2
17	9	No	Si	No	1	4	2
18	9	Si	Si	Si	1	5	2
19	9	Si	Si	Si	1	4	1
20	8	Si	Si	Si	1	4	3

Muestra de Estudiantes de ingeniería de sonido (Quinto semestre e inferiores)							
Audio 1							
	Semestre	Musical	Videojuego	Audiovisual	Inteligibilidad	Brillo	Calidez
1	5	No	No	No	2	4	1
2	5	Si	No	Si	5	4	3
3	5	Si	Si	Si	4	5	2
4	5	No	Si	Si	5	3	1
5	5	Si	No	Si	3	4	2
6	3	No	Si	Si	4	4	2
7	3	No	Si	SI	4	4	3
8	3	No	Si	Si	2	3	2
9	3	Si	Si	Si	4	3	3
10	4	No	Si	Si	4	3	5
11	5	Si	Si	Si	3	4	1
12	3	No	Si	Si	3	4	2
13	3	No	Si	Si	4	4	3
14	3	Si	Si	Si	3	4	2
15	3	No	Si	No	5	4	2
16	3	No	No	No	3	3	2
17	3	No	Si	Si	5	4	2
18	3	Si	Si	No	5	4	2
19	3	No	Si	Si	4	5	4
20	3	Si	No	Si	3	4	2

Muestra de Estudiantes de ingeniería de sonido (Quinto semestre e inferiores)							
Audio 2							
	Semestre	Musical	Videojuego	Audiovisual	Inteligibilidad	Brillo	Calidez
1	5	No	Si	No	1	4	2
2	5	Si	Si	Si	1	3	2

3	5	Si	Si	No	1	4	1
4	5	No	Si	Si	1	3	1
5	5	No	Si	Si	1	3	1
6	3	Si	Si	Si	1	5	1
7	3	No	Si	No	1	3	3
8	3	Si	Si	Si	1	4	3
9	3	No	Si	Si	2	4	3
10	4	Si	Si	No	3	4	2
11	5	No	Si	No	1	5	1
12	5	Si	Si	Si	1	4	2
13	3	Si	Si	Si	1	4	2
14	3	Si	Si	Si	4	5	3
15	3	No	Si	No	1	5	1
16	3	No	No	No	1	2	1
17	3	No	Si	Si	1	4	2
18	3	Si	Si	No	1	5	2
19	3	Si	Si	Si	1	5	4
20	3	No	Si	Si	1	4	3

Anexo P. Tabulación de encuestas realizadas a una muestra de músicos de profesión.

Muestra de músicos Audio 1				
	Musical	Inteligibilidad	Brillo	Calidez
1	No	3	4	2
2	Si	4	5	3
3	Si	4	4	2
4	Si	3	2	1
5	Si	4	4	1
6	Si	5	4	2

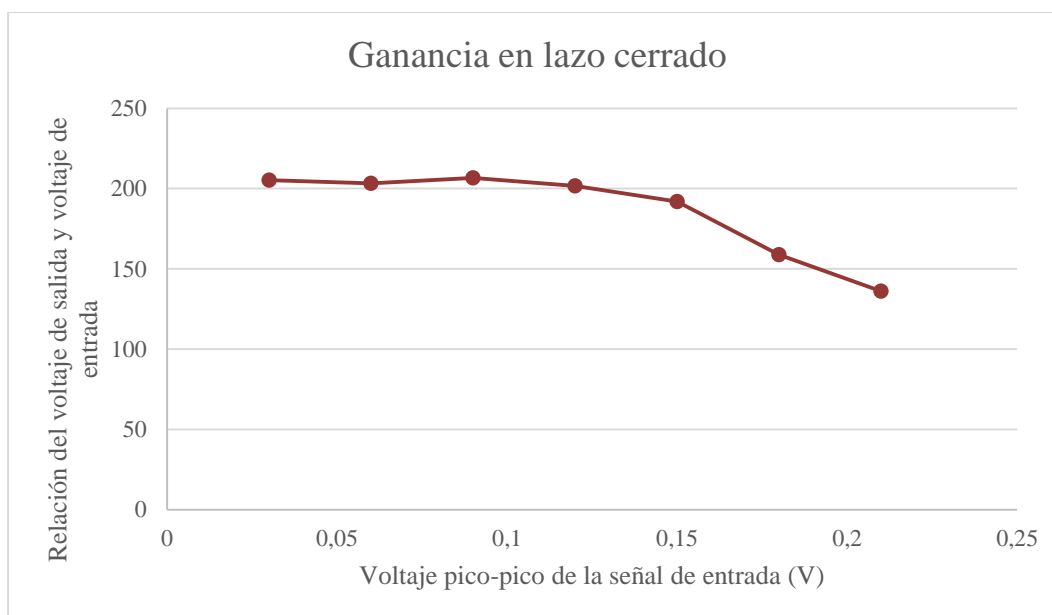
Muestra de músicos Audio 2				
	Musical	Inteligibilidad	Brillo	Calidez
1	No	1	4	2
2	Si	1	5	2
3	No	1	4	3
4	No	1	4	1
5	Si	1	5	1
6	No	1	5	1

Anexo Q. Mediciones del integrado AD620

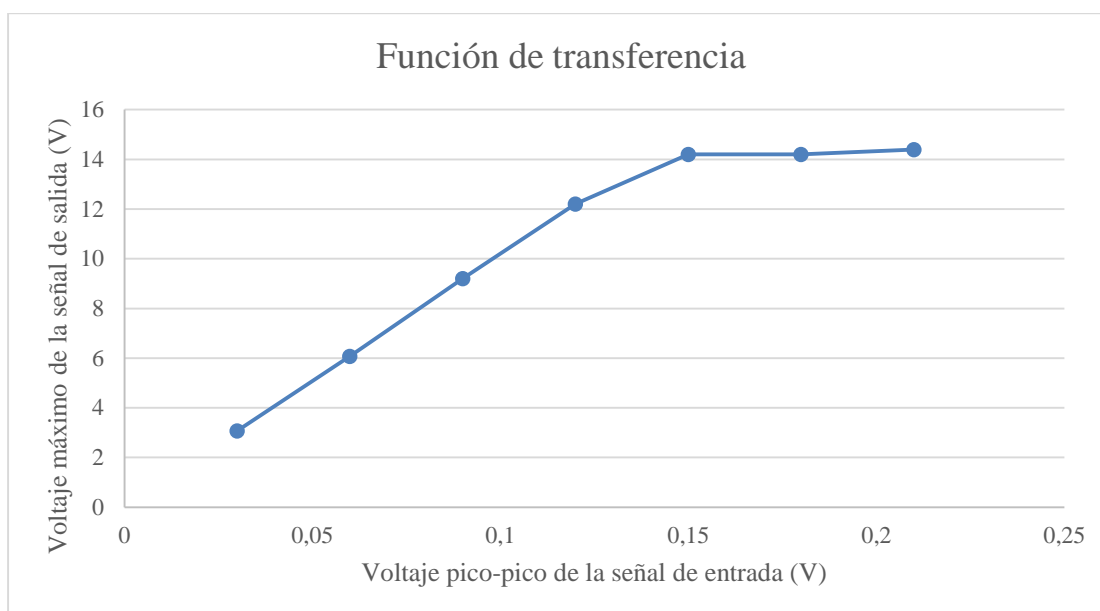
El factor cresta se define como el cociente entre el nivel del pico máximo y su valor eficaz y la ganancia utiliza el voltaje pico-pico de la salida del integrado y el voltaje pico-pico de la señal de entrada. Lo resultados obtenidos de esta medición, se muestran a continuación.

Resultados de medición del integrado en modo diferencial, Rigol DG1022.

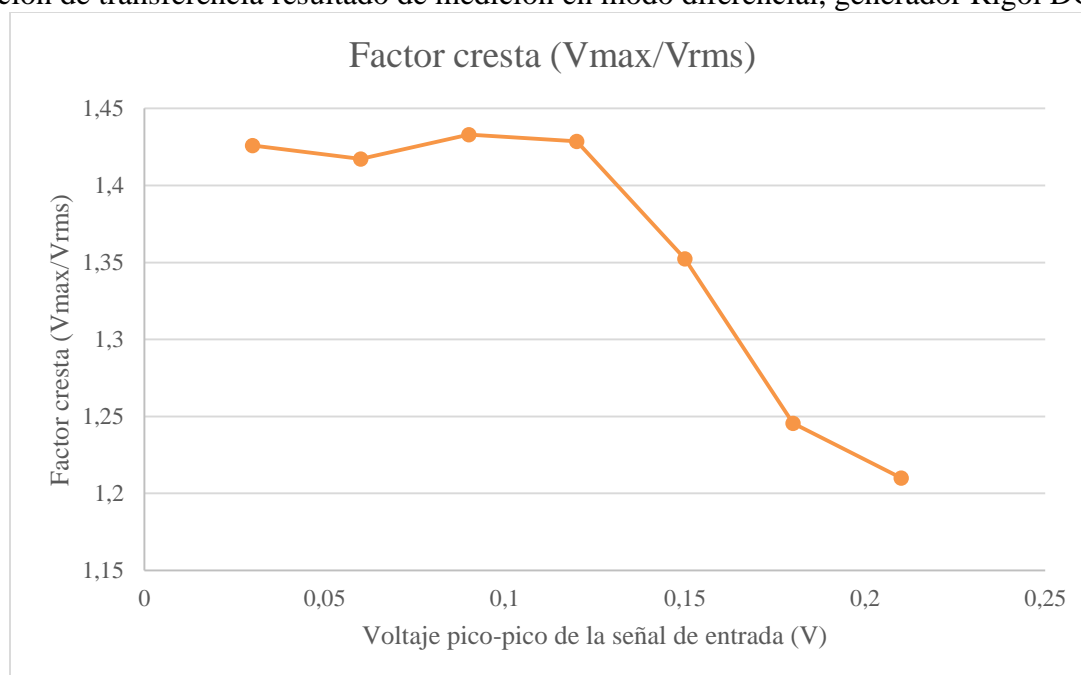
Voltaje del generador (Vp-p)	Vmax obtenido a la salida del integrado	Vmin obtenido a la salida del integrado	Vp-p obtenido a la salida del integrado	Vrms obtenido a la salida del integrado
0,03	3,08	-3,08	6,16	2,16
0,06	6,08	-6,16	12,2	4,29
0,09	9,2	-9,4	18,6	6,42
0,12	12,2	-12,2	24,2	8,54
0,15	14,2	-14,4	28,8	10,5



Ganancia resultado de medición en modo diferencial, generador Rigol DG1022.



Función de transferencia resultado de medición en modo diferencial, generador Rigol DG1022.



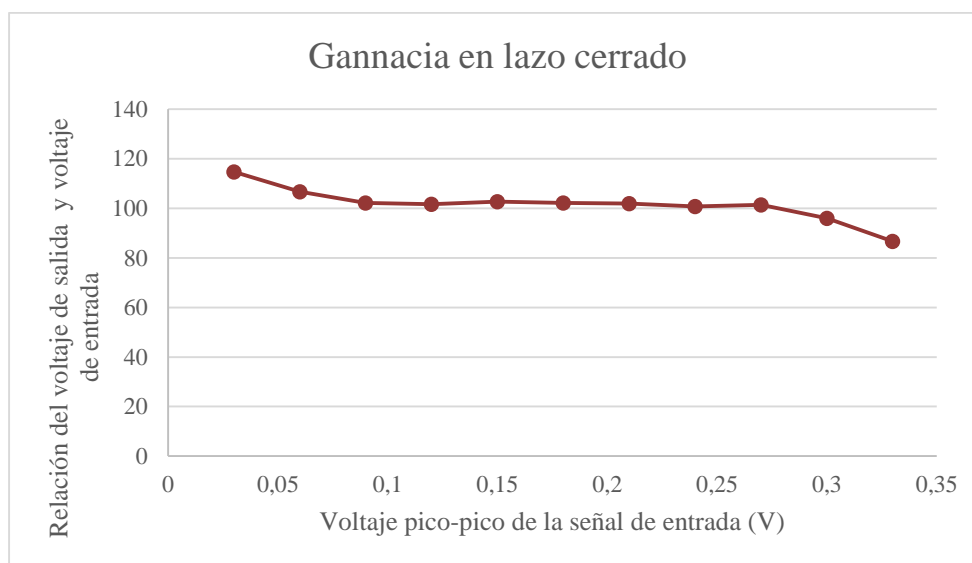
Factor cresta resultado de medición en modo diferencial, generador Rigol DG1022.

Tabla. Voltaje de los armónicos de la señal.

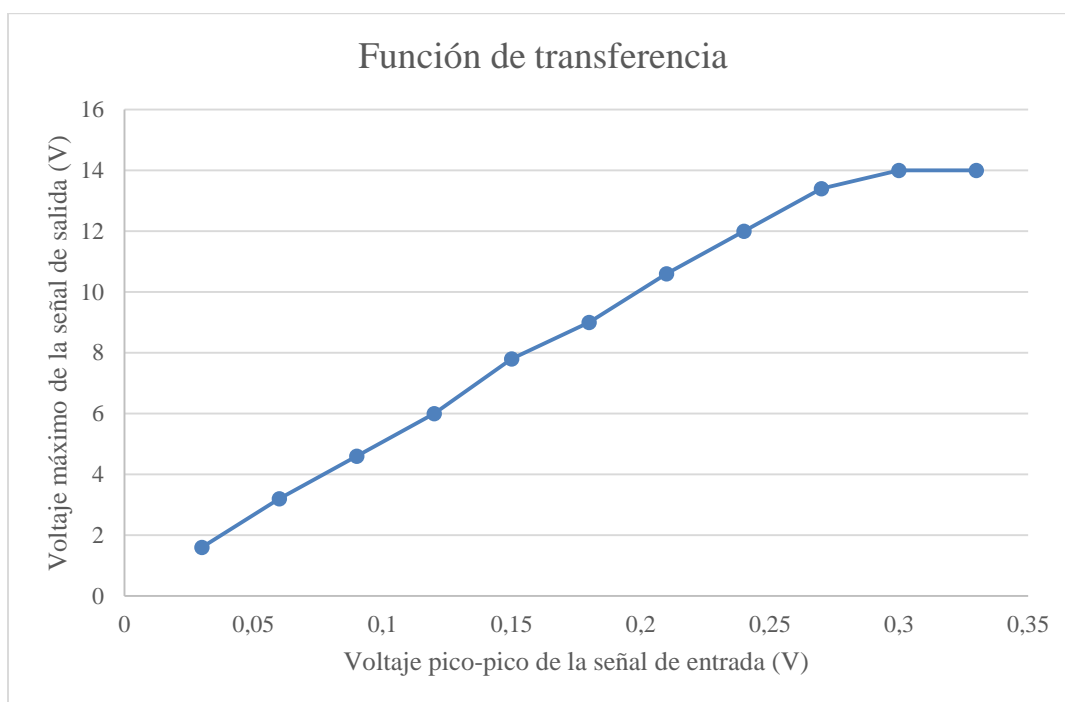
Armónico	Voltaje rms (V_{rms})
A0	8
A1	0,3
A2	0,1

Resultados de medición del integrado en modo común, Rigol DG1022.

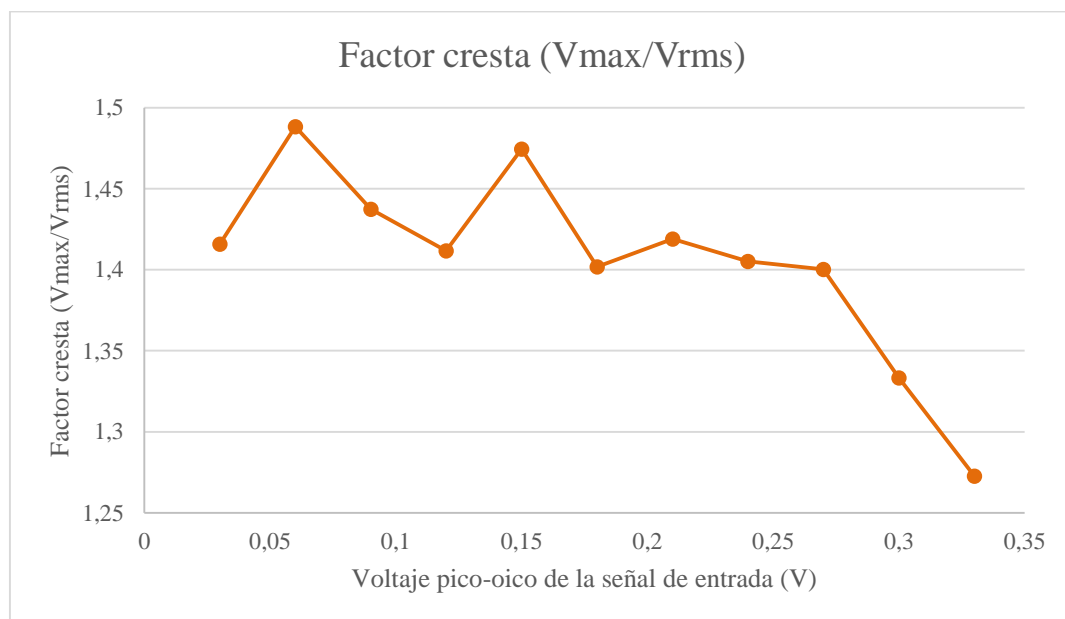
Voltaje del generador (Vp-p)	Vmax obtenido a la salida del integrado	Vmin obtenido a la salida del integrado	Vp-p obtenido a la salida del integrado	Vrms obtenido a la salida del integrado
0,03	1,6	-1,8	3,44	1,13
0,06	3,2	-3,2	6,4	2,15
0,09	4,6	-4,6	9,2	3,2
0,12	6	-6,2	12,2	4,25
0,15	7,8	-7,6	15,4	5,29
0,18	9	-9,4	18,4	6,42
0,21	10,6	-10,8	21,4	7,47
0,24	12	-12,2	24,2	8,54
0,27	13,4	-13,6	27,4	9,57
0,3	14	-14	28,8	10,5



Ganancia de transferencia resultado de medición en modo común, generador Rigol DG 1022.



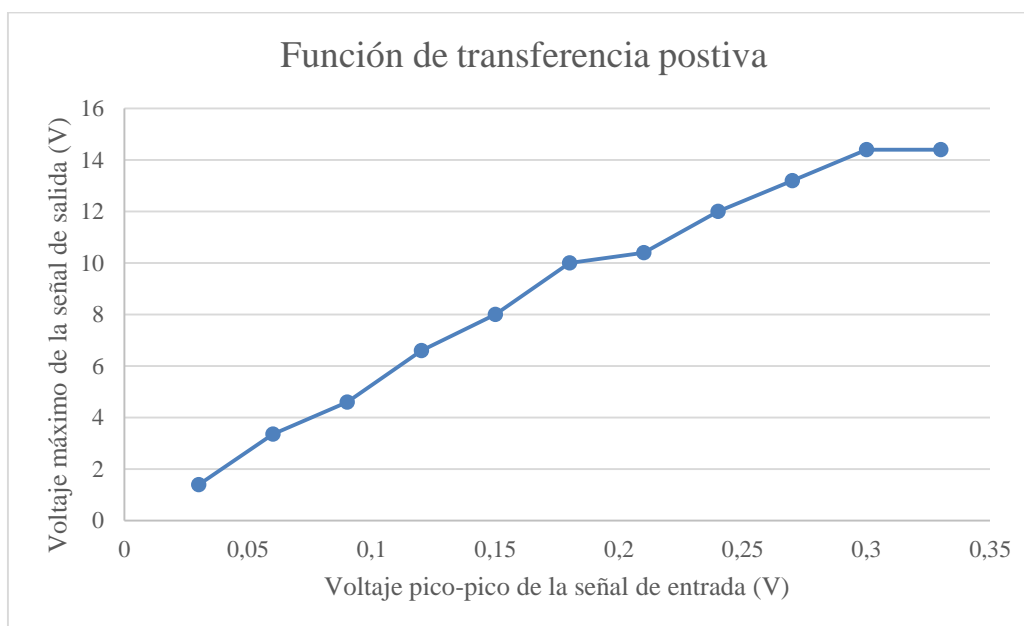
Función de transferencia resultado de medición en modo común, generador Rigol DG1022.



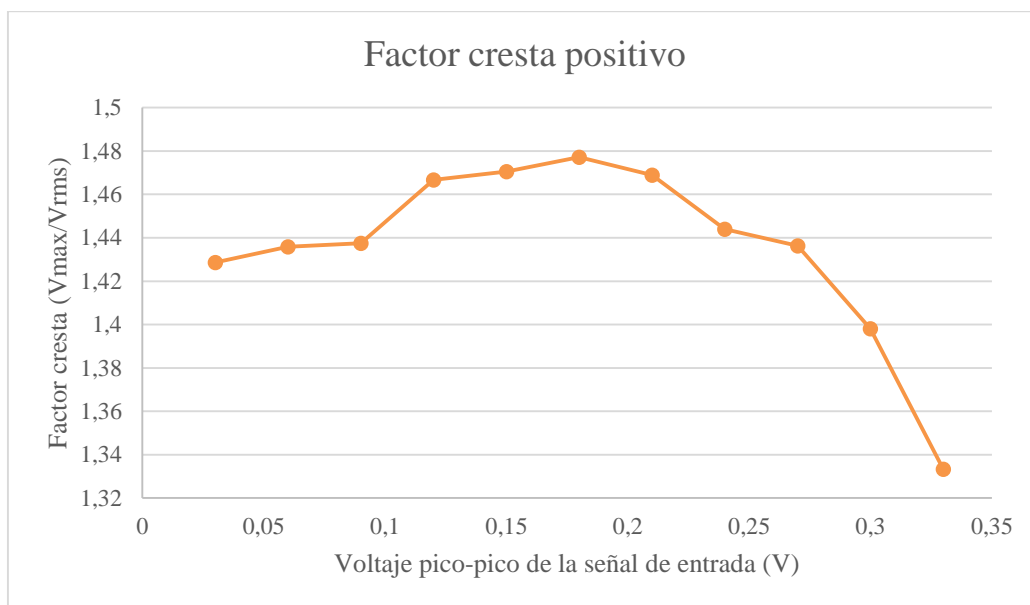
Factor cresta resultado de medición en modo común, generador Rigol DG 1022.

Resultados de medición del integrado en modo diferencial, generador PeakTeck 4025.

Voltaje del generador (Vp-p)	Vmax obtenido a la salida del integrado	Vmin obtenido a la salida del integrado	Vp-p obtenido a la salida del integrado	Vrms obtenido a la salida del integrado
0,03	1,4	-1,48	2,84	0,98
0,06	3,36	-3,36	6,72	2,34
0,09	4,6	-5	9,8	3,35
0,12	6,6	-6,8	13,4	4,7
0,15	8	-7,6	15,6	5,44
0,18	10	-9,6	19,6	6,77
0,21	10,4	-10	20,4	7,08
0,24	12	-12	24	8,31
0,27	13,2	-13,2	26,4	9,19
0,3	14,4	-14,8	29,2	10,3
0,33	14,4	-14,8	29,2	10,8



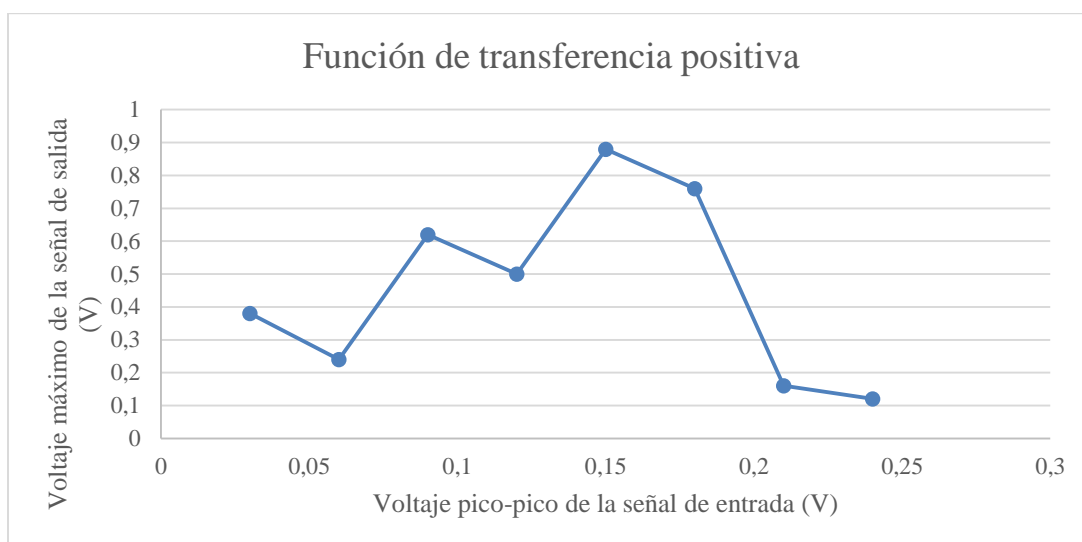
Función de transferencia resultado de medición en modo diferencial, generador PeakTeck 4025.



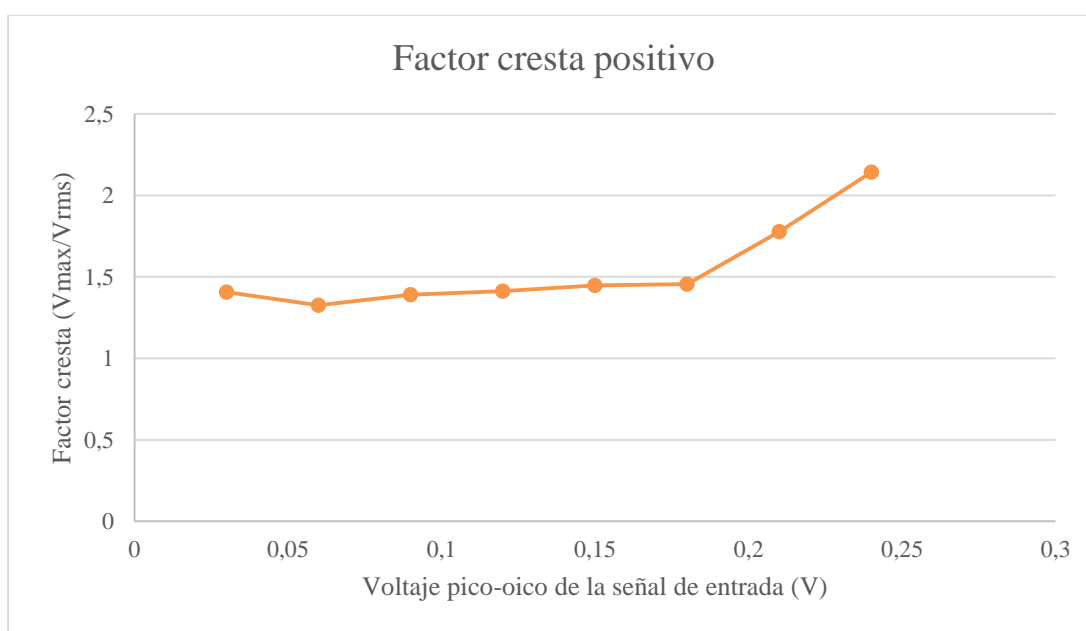
Factor cresta resultado de medición en modo diferencial, generador PeakTeck 4025.

Resultados de medición del integrado en modo común, generador PeakTeck 4025.

Voltaje del generador (Vp-p)	Vmax obtenido a la salida del integrado	Vmin obtenido a la salida del integrado	Vp-p obtenido a la salida del integrado	Vrms obtenido a la salida del integrado
0,03	0,38	-0,44	0,82	0,27
0,06	0,24	-0,3	0,54	0,181
0,09	0,62	-0,68	1,3	0,446
0,12	0,5	-0,56	1,06	0,354
0,15	0,88	-0,92	1,8	0,608
0,18	0,76	-0,8	1,56	0,522
0,21	0,16	-0,2	0,36	0,09
0,24	0,12	-0,12	0,24	0,056

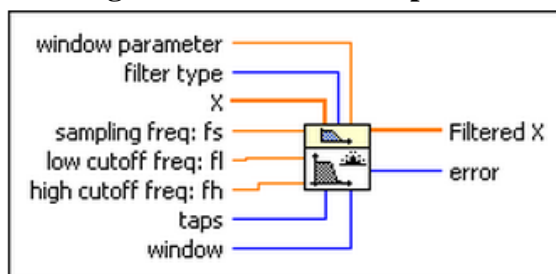


Función de transferencia resultado de medición en modo común, generador PeakTeck 4025



Factor cresta resultado de medición en modo común, generador PeakTeck 4025.

Anexo R. Bloques de filtros digitales de LabView implementados en el banco de filtros



FIR Windowed Filter VI.

Fuente: zone.ni.com/reference/en-XX/help/fir_windowed_filter

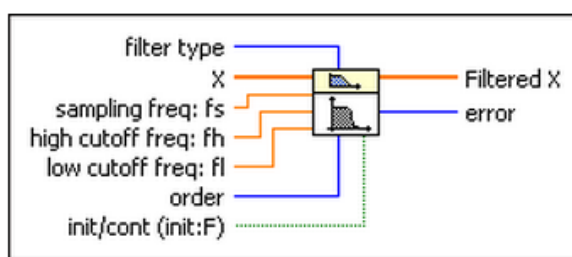


Figura 11. Butterworth Filter VI.

Fuente: zone.ni.com/reference/en-XX/help/371361J-01/1vanls/butterworth_filter

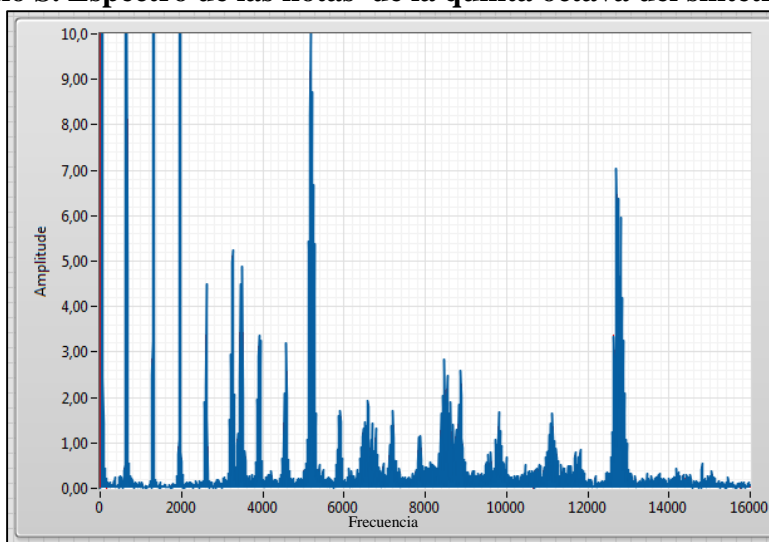
Anexo S. Espectro de las notas de la quinta octava del sintetizador

Figura. Espectro en frecuencia de la nota Mi.

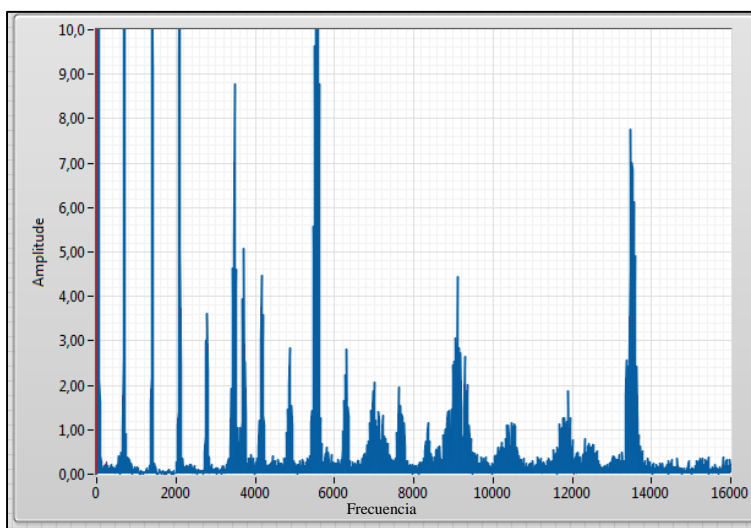


Figura. Espectro en frecuencia de la nota Fa.

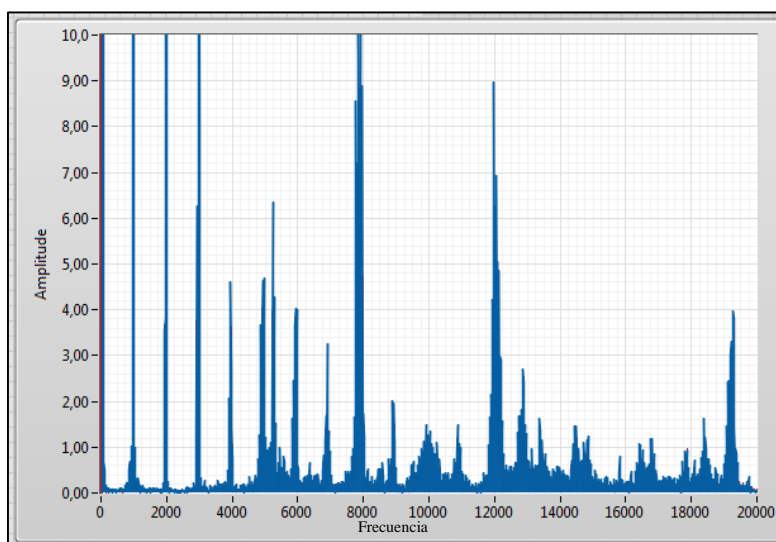


Figura. Espectro en frecuencia de la nota Si.

Anexo T. Espectro de las vocales I, O, U

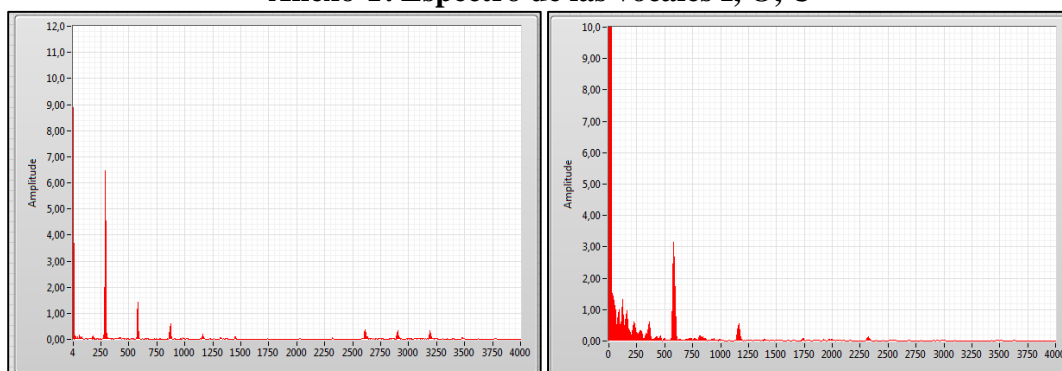
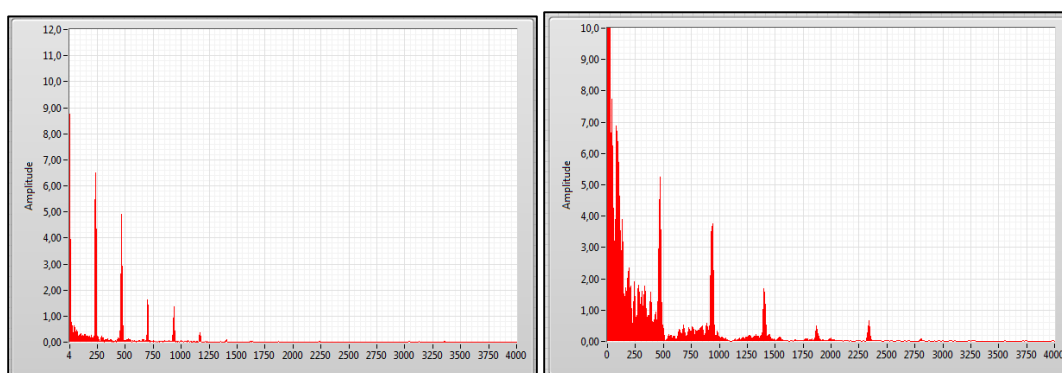
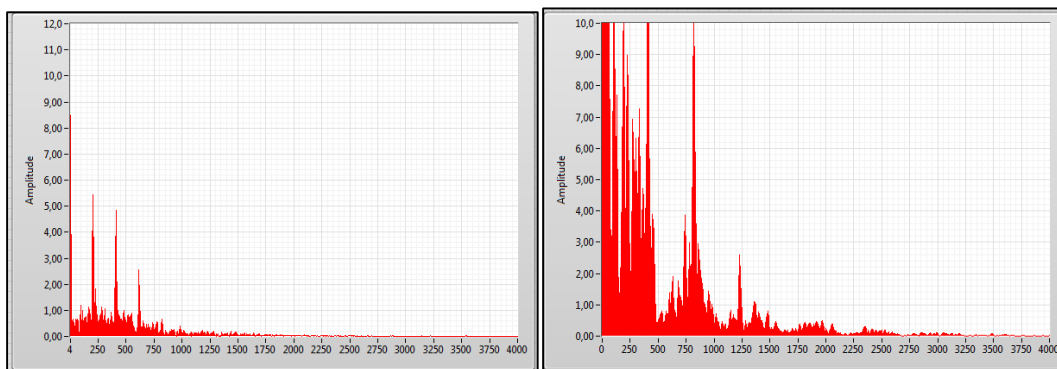


Figura 43. FFT de señal de entrada y su envolvente. La señal corresponde a la vocal "I".



FFT de señal de entrada y su envolvente. La señal corresponde a la vocal "O".



FFT de señal de entrada y su envolvente. La señal corresponde a la vocal "U".