

La investigación reportada en esta tesis es parte de los programas de investigación del CICESE (Centro de Investigación Científica y de Educación Superior de Ensenada, Baja California).

La investigación fue financiada por el CONAHCYT (Consejo Nacional de Humanidades, Ciencias y Tecnologías).

Todo el material contenido en esta tesis está protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México). El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo o titular de los Derechos Autor.

CICESE © 2023, Todos los Derechos Reservados, CICESE

Centro de Investigación Científica y de Educación Superior de Ensenada, Baja California



Maestría en Ciencias en Tecnologías Avanzadas e Integradas

Aproximación de señales acústicas a través de funciones polinomiales.

Tesis

para cubrir parcialmente los requisitos necesarios para obtener el grado de
Maestra en Ciencias

Presenta:

Andrea Cruz Rojas

Ensenada, Baja California, México

2023

Tesis defendida por

Andrea Cruz Rojas

y aprobada por el siguiente Comité

Dra. Ana Laura Padilla Ortiz

Codirectora de tesis

Dr. Jesús Emmanuel Gómez Correa

Codirector de tesis

Dr. Víctor Manuel Coello Cárdenas

Dra. Patricia Juárez Camacho



Dr. Víctor Manuel Coello Cárdenas

Coordinador de la Maestría en Tecnologías Avanzadas e Integradas

Dra. Ana Denise Re Araujo

Directora de Estudios de Posgrado

Resumen de la tesis que presenta Andrea Cruz Rojas como requisito parcial para la obtención del grado de Maestra en Ciencias en Tecnologías Avanzadas e Integradas.

Aproximación de señales acústicas a través de funciones polinomiales.

Resumen aprobado por:

Dra. Ana Laura Padilla Ortiz

Codirectora de tesis

Dr. Jesús Emmanuel Gómez Correa

Codirector de tesis

Las señales acústicas presentan diversas características, como la frecuencia, amplitud y duración. Existen varias formas de representar una señal acústica, siendo la representación en el dominio del tiempo la más común. Esta representación permite visualizar la variación de la amplitud a lo largo del tiempo, analizar la duración, la forma de onda y los cambios rápidos o lentos en la amplitud. Sin embargo, las señales acústicas pueden contener un gran número de elementos, lo que resulta en archivos digitales de gran tamaño que dificultan su envío y almacenamiento. En este trabajo, se llevó a cabo un estudio de aproximación de señales acústicas utilizando diversas funciones polinomiales, como las funciones Zernike, Bessel y Jinc para reducir la cantidad de datos necesarios para representar la señal acústica, lo que simplifica la manipulación de dichos datos. El objetivo es identificar qué polinomio y cuántos coeficientes ofrecen un mejor rendimiento, es decir, una mejor señal acústica. Este estudio se basa en una evaluación objetiva y subjetiva de la reconstrucción de las señales acústicas, lo cual, hasta donde tenemos conocimiento, no se ha estudiado utilizando ambas metodologías en conjunto.

Palabras clave: Calidad de audio, Compresión de señal acústica, Polinomios ortogonales, Pruebas objetivas y subjetivas.

Abstract of the thesis presented by Andrea Cruz Rojas as a partial requirement to obtain the Master of Science degree in Advanced and Integrated Technologies.

Approximation of acoustic signals through polynomial functions.

Abstract approved by:

Dra. Ana Laura Padilla Ortiz

Thesis Co-Director

Dr. Jesús Emmanuel Gómez Correa

Thesis Co-Director

Acoustic signals have various characteristics, such as frequency, amplitude and duration. There are several ways to represent an acoustic signal, with time domain representation being the most common. This representation allows you to visualize the variation of the amplitude over time, analyze the duration, the waveform and the fast or slow changes in the amplitude. However, acoustic signals can contain a large number of elements, resulting in large digital files that make them difficult to send and store. In this work, an approximation study of acoustic signals was carried out using various polynomial functions, such as the Zernike, Bessel and Jinc functions to reduce the amount of data necessary to represent the acoustic signal, which simplifies the manipulation of said data. The objective is to identify which polynomial and how many coefficients offer better performance, that is, a better acoustic signal. This study is based on an objective and subjective evaluation of the reconstruction of acoustic signals, which, to the best of our knowledge, has not been studied using both methodologies together.

Keywords: Audio quality, Acoustic signal compression, Orthogonal polynomials, Objective and Subjective Tests.

Dedicatoria

A mis padres Guillermina Rojas Aragón y Andrés Cruz Hernández por apoyar mis sueños y siempre confiar en mí incluso cuando yo no lo hacía. A mi hermana Tania Cruz Rojas, por su amor, apoyo y compañía.

Agradecimientos

Quiero expresar mi más sincero agradecimiento al Consejo Nacional de Humanidades, Ciencias y Tecnologías (CONAHCYT) por haberme otorgado el número de becario 1135453 y brindarme el apoyo económico que ha sido fundamental para realizar mis estudios de maestría.

Además, deseo agradecer de manera especial al Centro de Investigación Científica y de Educación Superior de Ensenada, Baja California (CICESE), Unidad Foránea Monterrey (UFM). Esta institución ha sido mi hogar desde el verano de investigación de 2018, mis prácticas profesionales en 2019 y durante todo mi programa de maestría. El CICESE-UFM ha jugado un papel esencial en mi desarrollo personal y académico, proporcionándome innumerables oportunidades y un valioso apoyo a lo largo de mi trayectoria.

Asimismo, me gustaría expresar mi más sincero agradecimiento a mis directores de tesis, la Dra. Ana Laura Padilla Ortiz y el Dr. Jesús Emmanuel Gómez Correa. Su paciencia, confianza y guía han sido fundamentales para resolver los desafíos que se han presentado tanto en el proyecto de tesis como en las materias que me han impartido.

Quiero agradecer también a los miembros de mi comité de tesis, la Dra. Patricia Juárez Camacho y el Dr. Víctor Manuel Coello Cárdenas, por sus valiosos consejos, comentarios y observaciones, los cuales han contribuido en gran medida al desarrollo y éxito de este trabajo de investigación.

Agradezco de corazón al Dr. Aldo Eleazar Pérez Ramos por su confianza y por ser un apoyo incondicional en mi introducción al mundo de la investigación. Agradezco su paciencia, tiempo y amistad sincera. Gracias por creer en mí.

Quiero expresar mi profundo agradecimiento a mis padres, hermana y familia. Su amor incondicional y constante apoyo han sido pilares fundamentales en mi vida. Nunca han dejado de creer en mí, y por eso les estoy eternamente agradecida. La persona que soy hoy en día es un reflejo de todo lo que he aprendido de ustedes. Su guía, sabiduría y ejemplo han dejado una huella

imborrable en mi camino. Desde lo más profundo de mi corazón, les doy las gracias por todo.

Agradezco también a mis compañeros, Fis. Fernando Serrano Monroy y Lic. Rodrigo Lecanda Rodríguez, por su amistad, apoyo, cariño y por compartir innumerables momentos. Siempre los llevaré en mi corazón.

A mis amigos que nunca me han dejado sola, gracias por estar siempre presentes con una llamada, un mensaje o algún gesto para animarme y hacerme sentir querida. Nuestra amistad es mutua y especial, y quiero agradecer especialmente a Angélica, Jassiel y Emmanuel.

Por último, pero no menos importante, me agradezco a mí misma por nunca rendirme, por enfrentar y superar todos los problemas, por tener la valentía y la fortaleza para seguir adelante. Agradezco mi perseverancia en los momentos más difíciles. Rendirse era una opción, pero seguir adelante fue la decisión.

Tabla de contenido

	Página
Resumen en español	ii
Resumen en inglés	iii
Dedicatoria	iv
Agradecimientos	v
Lista de figuras	ix
Lista de tablas	xi
Capítulo 1. Introducción	
1.1. Antecedentes	2
1.2. Justificación	4
1.3. Hipótesis	5
1.4. Objetivos de la investigación	5
1.4.1. Objetivos específicos	5
Capítulo 2. Señales	
2.1. Señal	6
2.2. Digitalización de una señal de voz	7
2.3. Señal de voz	12
Capítulo 3. Polinomios ortogonales	
3.1. Polinomios	14
3.2. Polinomios ortogonales	14
3.3. Series de Fourier	16
3.3.1. Coeficientes de Fourier	17
3.4. Funciones de Bessel	18
3.4.1. Coeficientes de Bessel	19
3.4.2. Función Airy	20
3.5. Polinomios de Zernike	21
3.5.1. Coeficientes de Zernike	22
3.6. Cálculo de coeficientes por mínimos cuadrados	23
Capítulo 4. Reconstrucción de una señal mediante polinomios ortogonales	
4.1. Procedimiento de reconstrucción	27
4.1.1. Leer archivo WAV	28
4.1.2. Canales de audio	29
4.1.3. Tamaño muestral de la señal	30
4.1.4. División del tamaño muestral en 4 secciones	30

4.1.5.	Cálculo del número de coeficientes	31
4.1.6.	Reconstrucción de la señal	32
4.1.7.	Unión de las 4 secciones reconstruidas	33
4.1.8.	Cálculo del Error Cuadrático Medio Normalizado	33
Capítulo 5. Pruebas objetivas y subjetivas		
5.1.	Pruebas objetivas	34
5.1.1.	Errores	34
5.2.	Pruebas subjetivas	35
5.2.1.	Inicio de la página web	36
5.2.2.	Cuestionario de audición	38
5.2.3.	Secciones de la página web	40
5.2.4.	Orden de los audios	40
Capítulo 6. Resultados		
6.1.	Información de los 20 audios	42
6.2.	Audios seleccionados para las pruebas objetivas y subjetivas	47
6.3.	Pruebas objetivas	48
6.3.1.	Error cuadrático medio normalizado	48
6.4.	Pruebas subjetivas	50
6.5.	Discusión	52
Capítulo 7. Conclusiones		
7.1.	Conclusiones	54
7.2.	Trabajo a futuro	55
Literatura citada		57
Apéndices		60

Lista de figuras

Figura	Página
1. Señal senoidal. Esta figura muestra algunas características de una señal como es la amplitud, la longitud de onda y el periodo. Las unidades de la amplitud dependerán del fenómeno físico que se este describiendo (voltaje, decibeles, etc.).	7
2. Muestreo de una señal. El muestreo periódico en una señal analógica.	9
3. Cuantificación de una señal. Cada muestra obtenida anteriormente se aproxima mediante un valor seleccionado.	10
4. Codificación de una señal. La representación de los valores será de forma binaria (0,1).	11
5. Ejemplo de una señal de voz. La palabra "Clavo" podemos observarla como señal de voz en esta figura, donde, en el eje horizontal tenemos el tiempo y en el eje vertical tenemos la amplitud de la señal.	12
6. Primeros 5 términos de las Series de Fourier, los cuales podremos observar cada uno con un color diferente. Esta gráfica puede extenderse desde 0 hasta el infinito, aunque en la representación actual abarca hasta el valor 6.	16
7. Funciones de Bessel para órdenes enteros $J_0(x)$, $J_1(x)$, $J_2(x)$, $J_3(x)$ y $J_4(x)$. . .	18
8. Función Airy para órdenes enteros $Airy_0(u)$, $Airy_1(u)$, $Airy_2(u)$ y $Airy_3(u)$. .	20
9. Polinomios de Zernike. Cuando $m = 0$, los polinomios de Zernike solo presentan simetría radial.	21
10. Diagrama a bloques de la reconstrucción de la señal de voz. Siguiendo este diagrama, se reconstruyó una señal de voz con los diferentes polinomios mencionados.	28
11. Ejemplificación de una señal completa y la señal separada en el canal derecho y canal izquierdo.	29
12. Ejemplo de una señal de 36,000 muestras divide en 4 secciones.	30
13. Comparación de la señal original y la señal dividida en 4 secciones. En la primera señal (a), se observa que todas las muestras de la señal están juntas y en la segunda señal (b) se observa que la señal esta dividida en 4 secciones.	31
14. Ejemplo del cálculo de los coeficientes para cada audio.	32
15. Esquema del canal derecho y canal izquierdo dividido en 4 secciones y los coeficientes divididos en estas 4 secciones.	32
16. Representación del canal derecho y del canal izquierdo después de su reconstrucción, culminando con la obtención de la señal completamente reconstruida. . . .	33
17. Presentación de la página web. Se muestra la forma en la que el participante tendrá el primer contacto con la página web, donde se le explica qué es una prueba subjetiva y el objetivo de esta prueba.	37

Figura	Página
18. Consentimiento informado. Se contestan las posibles preguntas que pueda tener el participante al desarrollar la prueba y en qué consiste la misma.	38
19. Error cuadrático medio normalizado. Se observa cada uno de los polinomios con sus respectivos porcentajes de coeficientes que están representados con un color diferente.	44
20. Tiempo promedio de la reconstrucción de los 20 audios elegidos al azar por cada polinomio. La desviación estándar de los audios son tan altas ya que los audios tienen diferentes números de muestras y esto hace que algunos audios se reconstruyan más rápido que otros.	46
21. El orden de los audios reconstruidos se encuentra estructurado de la siguiente manera: Figura a) En esta figura se presentan las cinco carpetas principales donde se almacenan los audios. Figura b) Cada una de las carpetas en esta figura lleva el nombre de una palabra específica.	47
22. Error cuadrático medio normalizado. Se observa cada uno de los polinomios con sus respectivos porcentajes de coeficientes que están representados con un color diferente.	48
23. Resultados de la puntuación otorgada por los participantes a cada reconstrucción de la señal de audio, en función del porcentaje de coeficientes utilizados. La línea roja destaca que los polinomios Airy obtuvieron una puntuación similar a las Series de Fourier, pero con un menor número de coeficientes.	51

Lista de tablas

Tabla		Página
1.	Total de audios reconstruidos.	43
2.	Total de audios reconstruidos.	43
3.	Error cuadrático medio normalizado de los 20 audios.	45
4.	Desviación estandar de los 20 audios.	45
5.	Tabla de palabras utilizadas.	48
6.	Tabla de valores del ECMN de las 5 palabras.	49
7.	Tabla de puntuación.	50
8.	Promedio de las pruebas subjetiva.	51
9.	Desviación estándar pruebas subjetivas.	52

Capítulo 1. Introducción

La compresión de voz consiste en la reducción del volumen de información tratable, es decir, que se puede procesar, transmitir o grabar. Se pretende transportar la misma información pero empleando una menor cantidad de espacio (Barzola Tobar & Cabrera Velasco, 2013).

El objetivo de la compresión de voz es representar la voz en forma digital con la menor cantidad de bits posibles manteniendo al mismo tiempo la inteligibilidad y la calidad requerida para una aplicación particular (Gibson, 2016).

La compresión de voz presenta una serie de ventajas significativas. En primer lugar, permite optimizar la velocidad necesaria para cumplir con las restricciones de distorsión o para minimizar directamente dicha distorsión. Además, posibilita un ahorro notable de espacio, una reducción en la demanda de ancho de banda y un aumento en la capacidad de almacenamiento disponible.

Sin embargo, como contrapartida, surgen algunas desventajas importantes. Una de ellas es la pérdida de calidad en la voz, que puede ser apreciable en ciertos casos, afectando la comprensión del contenido transmitido. Además, puede conllevar la degradación del audio, dificultando aún más la experiencia auditiva. Otro aspecto negativo se relaciona con la complejidad que presentan ciertos algoritmos de compresión, lo cual puede aumentar la carga de procesamiento y, en algunos escenarios, generar demoras en la reproducción debido a la necesidad de descomprimir el audio antes de su reproducción.

Las aplicaciones que hacen uso de la compresión de voz presentan un amplio espectro de alcance, abarcando desde el ámbito de la telefonía hasta el campo del reconocimiento automático de voz y asistentes de voz. Además, engloban áreas como la transmisión de voz sobre IP y encuentran aplicaciones en el ámbito médico (Romero & Mares, 2013).

Día tras día, el número de aplicaciones que requieren una mayor calidad va en aumento, según los requisitos de los usuarios. Por este motivo, se busca constantemente alternativas para comprimir los datos con el fin de facilitar su transmisión. A lo largo de las últimas décadas, diversos investigadores han desarrollado técnicas y algoritmos de compresión que hacen viable la transmisión de contenido multimedia (Kaschel C et al., 2005).

En su artículo, Dusan et al. (2007) presenta la codificación segmentaria como un método de

compresión que busca aprovechar las relaciones temporales en segmentos más largos del habla. Estos enfoques pueden implementarse en codificadores de voz basados en tramas existentes y, aunque pueden lograr una mayor compresión que el codificador original, esto a menudo implica la introducción de retardos de transmisión adicionales. Además, un tipo específico de codificación segmentaria utiliza funciones polinomiales para aproximar las trayectorias de las características del habla a lo largo del tiempo, habiendo sido aplicado previamente en el procesamiento de contornos de tono, donde las expresiones en inglés se descomponen en segmentos de tono y se aproximan mediante conjuntos de polinomios ortogonales.

1.1. Antecedentes

La compresión de voz es un proceso esencial para reducir la cantidad de datos necesarios al almacenar o transmitir señales de voz digitalizada. Los métodos de compresión del habla tienen como objetivo disminuir la tasa de bits de transmisión sin comprometer la calidad del habla. No obstante, es crucial lograr una compresión de voz que no afecte la calidad del audio resultante, ya que una compresión excesiva podría llevar a una degradación perceptible de la voz.

Bria et al. (1995) menciona que un tema obligado en el uso de multimedia es la compresión de datos, ya que la calidad de audio se necesita muestrear 44,100 veces por segundo, con una resolución de 16 bits por cada muestra. En stereo esto lleva a 1.4 Mbits por segundo, lo cual se vuelve una necesidad comprimir la señal.

Se presenta una Técnica de Compresión de Audio en el Espacio Transformado, donde la señal de entrada es dividida en bloques de longitud fija, los cuales son transformados en tres alternativas: La transformada de Karhunen-Loève (KLT), La transformada Discreta de Fourier (DFT), Transformada Discreta Coseno (DCT).

Una técnica utilizada es la cuantización vectorial no determinista para codificar y comprimir la información, pero es posible que al implementar este tipo de algoritmos se encuentren problemas en la calidad de la información, lo que nos lleva a que, al querer recuperar los datos originales, tengamos pérdidas en la inteligibilidad de la señal original.

Durante la cuantización vectorial no determinista, los valores de la señal de voz se asignan a un conjunto limitado de valores discretos. Debido a la naturaleza no determinista, pueden ocurrir pequeñas variaciones en los valores cuantizados. Estas variaciones pueden resultar en un ruido adicional en la señal de voz, lo que podría degradar la calidad y la claridad del habla, especialmente en partes más silenciosas o detalles sutiles (Dutra et al., 2009).

De igual forma, podría introducir errores en la representación de los fonemas, que son sonidos básicos de un idioma y esenciales para la comprensión del habla, lo que podría resultar en una pronunciación distorsionada y, en última instancia, una calidad de audio deficiente.

Con el fin de aprovechar al máximo estas señales, se recurre al algoritmo LSM (Least-Mean-Square Algorithm) de filtrado adaptativo (Arias, 2007). Su principal propósito radica en ajustar de forma iterativa los coeficientes de un filtro adaptativo para minimizar el error cuadrático medio entre la señal de entrada y la señal de salida filtrada. Esta técnica resulta especialmente beneficiosa en contextos en los que las características de la señal pueden cambiar con el tiempo, y la optimización del filtro es crucial ante tales variaciones.

Es imperativo hallar el equilibrio adecuado entre la compresión y la calidad del audio obtenido. Un ajuste meticuloso de los parámetros tanto del algoritmo LMS como del filtro adaptativo asegura una reducción de la redundancia y de los detalles de menor perceptibilidad, al mismo tiempo que se mantiene una calidad auditiva aceptable.

Dusan et al. (2007) habla de un método de compresión, llamado codificación segmentaria, tiene como objetivo explotar las relaciones temporales en segmentos del habla más largos. Dichos métodos se pueden aplicar encima de cualquier codificador de voz basado en tramas y pueden lograr valores de compresión mayores que el codificador original a expensas de introducir mayores retardos de transmisión.

Un tipo de codificación segmentaria utiliza funciones polinomiales, para aproximar las trayectorias de las características del habla presentes en sucesivos periodos de tiempo. Una de las primeras aplicaciones de la aproximación polinomial en el procesamiento del habla se aplicó en contornos de tono, donde algunas expresiones en inglés se dividieron en segmentos de tono y cada curva de tono se aproximó mediante un conjunto de polinomios ortogonales.

Strange (2005) nos explica que una técnica de análisis consiste en transformar los vectores de características en expresiones polinómicas. Los coeficientes de estos polinomios contienen información relevante acerca de las particularidades de la forma de onda que no resulta tan evidente en la representación original de la onda sonora.

1.2. Justificación

Como se menciona anteriormente, reconstrucción de señales por medio de funciones polinomiales nos permite aproximar las trayectorias de una señal, lo que simplifica su análisis y facilita la identificación de características importantes, como frecuencias dominantes y patrones temporales. Se han encontrado trabajos relacionados con los polinomios de Chebyshev, sin embargo, se desconoce cuál es el polinomio que puede aproximar mejor una característica de una señal.

En esta tesis se propone contribuir al campo con el desarrollo de un programa computacional capaz de realizar la reconstrucción de señales a través de la utilización de diversos polinomios. Para evaluar su desempeño de manera objetiva, se someterá el resultado de este programa a un análisis cuantitativo que empleará métricas específicas. Una de las métricas clave en esta evaluación es el error cuadrático medio normalizado, que proporcionará una evaluación global de la discrepancia entre la señal original y la señal reconstruida.

Además, se empleará una aproximación subjetiva para evaluar el resultado de este programa, la cual involucrará a un panel de oyentes que responderán a un conjunto de preguntas vinculadas con la señal reconstruida. Este cuestionario se administrará a través de una plataforma virtual alojada en Google Sites (<https://sites.google.com/cicese.edu.mx/prueba-subjetiva/inicio>), donde se registrarán y almacenarán las respuestas y opiniones de los oyentes.

El resultado de estas pruebas tanto objetivas como subjetivas, nos darán una perspectiva más amplia sobre qué polinomios son los que obtienen mejores resultados. Los resultados de estas pruebas serán documentados en la tesis de este proyecto, los cuales, serán de gran ayuda para investigaciones futuras y en diferentes áreas de aplicación de las mismas.

1.3. Hipótesis

Las aproximaciones utilizando funciones polinomiales permiten reconstruir señales de voz con un error que influye en la preservación de la calidad acústica de la señal. Específicamente, se plantea que a menor error en la aproximación, se logra una mejor calidad acústica en la señal reconstruida.

1.4. Objetivos de la investigación

Desarrollar y evaluar, mediante técnicas objetivas y subjetivas, un sistema que aproxime y reconstruya señales acústicas con diferentes polinomios.

1.4.1. Objetivos específicos

- Seleccionar los polinomios ortogonales que serán implementados
- Desarrollar algoritmos computacionales en MATLAB de las funciones polinomiales seleccionadas
- Generación de señales reconstruidas utilizando diferentes funciones polinomiales
- Identificación de las funciones polinomiales que mejor realicen la reconstrucción de las señales acústicas, utilizando métricas objetivas y subjetivas
- Analizar y comparar los resultados obtenidos

Capítulo 2. Señales

Las señales analógicas se manifiestan en la naturaleza como representaciones dinámicas de información a lo largo del tiempo. Estas señales tienen la versatilidad de adoptar valores en cualquier punto dentro de un rango definido, experimentando cambios de forma continua y fluida. En contraste, una señal digital se presenta como una representación discreta de información, codificada en formato binario mediante los dígitos 0 y 1. A diferencia de las señales analógicas, las señales digitales se distinguen por niveles discretos y transiciones de valor en intervalos específicos.

La transformación de señales analógicas a formato digital, un proceso fundamental denominado conversión analógico-digital, se torna indispensable para permitir su correcto procesamiento. Información más detallada acerca de este proceso se proporcionará a lo largo de este capítulo.

2.1. Señal

Proakis & Manolakis (1996) nos menciona que una señal se define como cualquier magnitud física que varía con el tiempo, el espacio o cualquier otra variable o variables independientes. Matemáticamente, describimos una señal como una función de una o más variables independientes.

Algunas características de una señal las podemos observar en la figura 1:

- Amplitud: Máximo desplazamiento de la onda respecto a su posición de equilibrio
- Frecuencia: Es el número de ciclos que ocurren por unidad de tiempo. Se mide en Hertz (Hz)
- Periodo (T): Es el tiempo que se tarda en realizar una oscilación completa y se mide en segundos (s)
- Longitud de onda (λ): Es la distancia que hay entre crestas y/o valles consecutivos de la onda, podemos medirlo en metros (m)

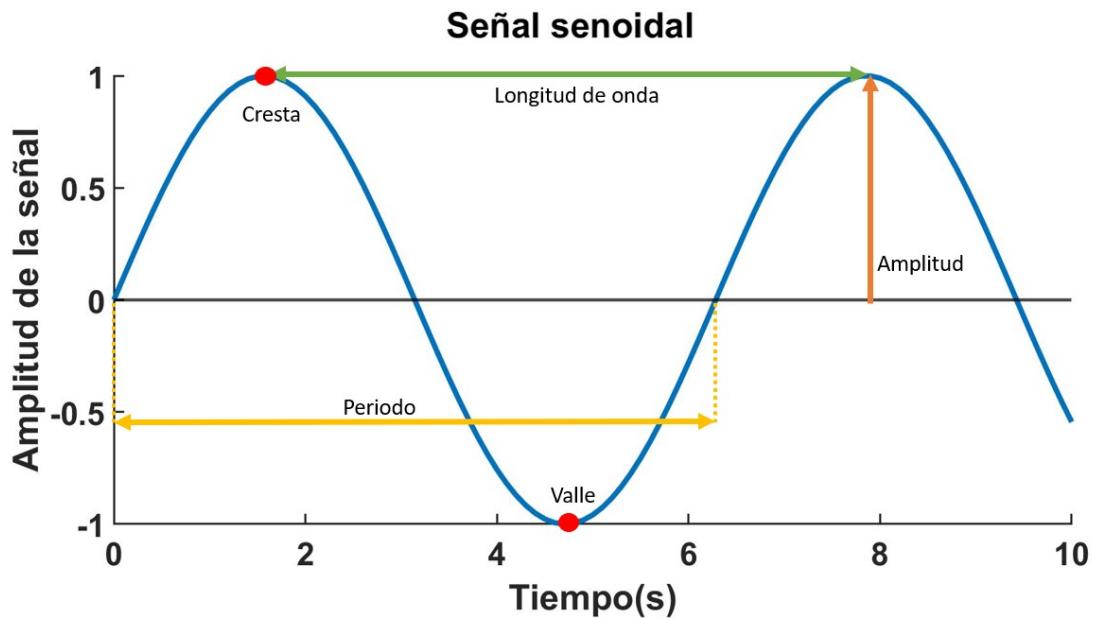


Figura 1. Señal senoidal. Esta figura muestra algunas características de una señal como es la amplitud, la longitud de onda y el periodo. Las unidades de la amplitud dependerán del fenómeno físico que se este describiendo (voltaje, decibeles, etc.).

2.2. Digitalización de una señal de voz

La gran mayoría de las señales que se estudian en el ámbito de la ciencia y la ingeniería poseen una naturaleza analógica. Estas señales se modelan como funciones de una variable continua, ya sea el tiempo o el espacio y típicamente adoptan valores en una escala ininterrumpida (Proakis & Manolakis, 1996).

Un ejemplo de señal analógica es el electrocardiograma (ECG), que brinda al facultativo datos fundamentales sobre la condición cardíaca del paciente. De manera similar, la señal captada por un electroencefalograma (EEG) proporciona valiosa información acerca de la actividad cerebral en cuestión.

El sonido ejemplifica de manera clásica una señal analógica. Las ondas sonoras que captamos en el medio ambiente experimentan variaciones constantes en amplitud y frecuencia, otorgándonos la capacidad de apreciar una extensa diversidad de tonos y matices tanto en la música como en el habla.

Otras ilustraciones abarcan el habla, señales biológicas, fenómenos sísmicos, datos de radar, lecturas de sonar, etc. Este conjunto denota que tanto sus variables independientes como sus respectivos valores se desenvuelven de manera continua (Moya, 2011).

Para procesar señales analógicas por medios digitales, primero es necesario convertirlos a forma digital, es decir, convertirlos a una secuencia de números con precisión finita. Este procedimiento se llama de analógico a digital (A/D) (Proakis & Manolakis, 1996).

Para la conversión de una señal analógica a digital, se tienen 3 pasos:

- Muestreo

El muestreo uniforme o periódico se describe mediante la siguiente relación:

$$x(n) = x_a(nT), -\infty < n < \infty. \quad (1)$$

En esta definición x_n representa la señal discreta en el tiempo obtenida "al muestrear" la señal analógica $x_a(t)$ cada T segundos. Un ejemplo de este proceso se representa en la figura 2. El intervalo de tiempo T entre muestras sucesivas define el periodo de muestreo y su inverso $1/T$ se denota como F_s que corresponde a la tasa de muestreo o frecuencia de muestreo.

El muestreo periódico establece una relación entre las variables temporales t y n correspondientes a las señales continuas y discretas en el tiempo, respectivamente. Estas variables se conectan de manera lineal a través del periodo de muestreo T , o de manera equivalente, mediante la tasa de muestreo $F_s = 1/T$. Esto se expresa de la siguiente manera:

$$t = nT = \frac{n}{F_s}. \quad (2)$$

Para cualquier señal analógica, es crucial seleccionar una frecuencia de muestreo conocida como F_{max} que sea lo suficientemente alta. En este sentido, se debe elegir $F_s/2$ de manera que sea mayor que F_{max} . Por lo tanto, se establece la siguiente ecuación:

$$F_s > 2F_{max}. \quad (3)$$

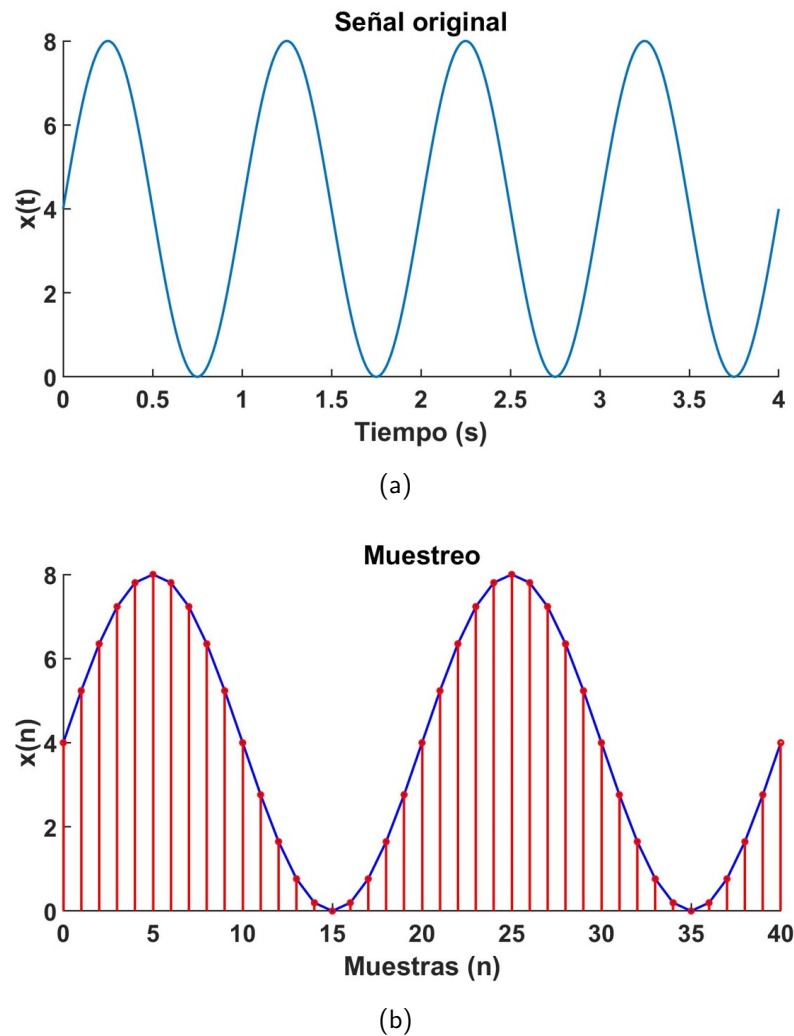


Figura 2. Muestreo de una señal. El muestreo periódico en una señal analógica.

- **Cuantificación**

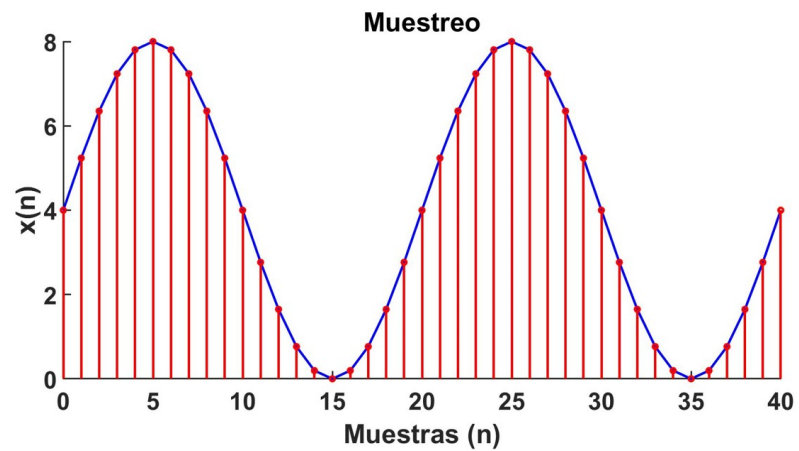
El proceso de transformar una señal discreta en el tiempo con amplitud continua en una señal digital, donde cada valor de muestra se expresa como un número finito de dígitos en lugar de una representación infinita, es conocido como "cuantificación". El error resultante al representar la señal continua mediante un conjunto limitado de niveles discretos se denomina "error de cuantificación".

La cuantificación de las muestras $x(n)$ se representa como $Q[x(n)]$, y utilizamos $x_q(n)$ para referirnos a la secuencia de muestras cuantificadas en la salida del cuantificador. Por lo tanto:

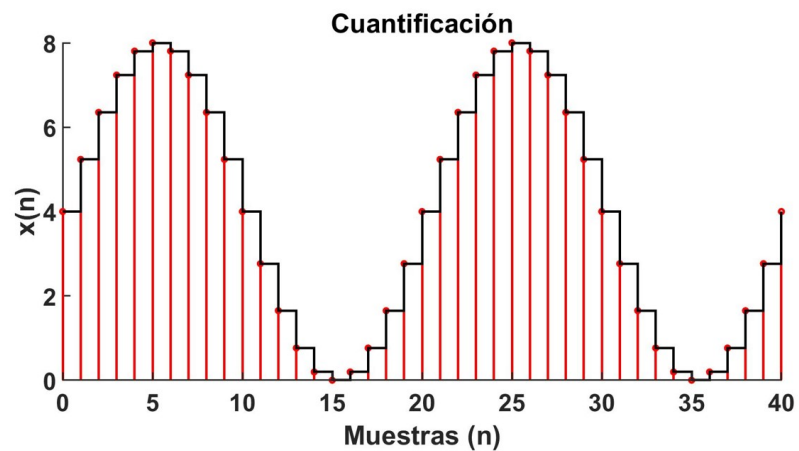
$$x_q(n) = Q[x(n)]. \quad (4)$$

El error de cuantificación se caracteriza por la secuencia $e_q(n)$, la cual se define como la diferencia entre el valor cuantificado y el valor real de la muestra. Por consiguiente:

$$e_q(n) = x_q(x) - x(n). \quad (5)$$



(a)



(b)

Figura 3. Cuantificación de una señal. Cada muestra obtenida anteriormente se aproxima mediante un valor seleccionado.

En teoría, la cuantificación de señales analógicas conlleva inevitablemente pérdida de información. Este resultado surge de la ambigüedad introducida por el proceso de cuantificación.

Además, la cuantificación es un proceso irreversible o no invertible, lo que significa que existe una relación de muchos a uno. En otras palabras, varias muestras distintas pueden mapear a un mismo nivel de cuantificación. Esta ambigüedad complica en gran medida la realización de un análisis cuantitativo preciso de la cuantificación.

- Codificación

El proceso de codificación asigna a cada nivel de cuantización un número binario singular. Si tenemos L niveles, se requerirá como mínimo L números binarios distintos. Con una longitud de palabra de b bits, podemos generar 2^b números binarios. En consecuencia, la cantidad de bits necesarios para el codificador $b \geq \log_2 L$. En términos generales, a medida que la frecuencia de muestreo sea mayor y la cuantización más precisa. Un ejemplo podemos observar en la figura 4.

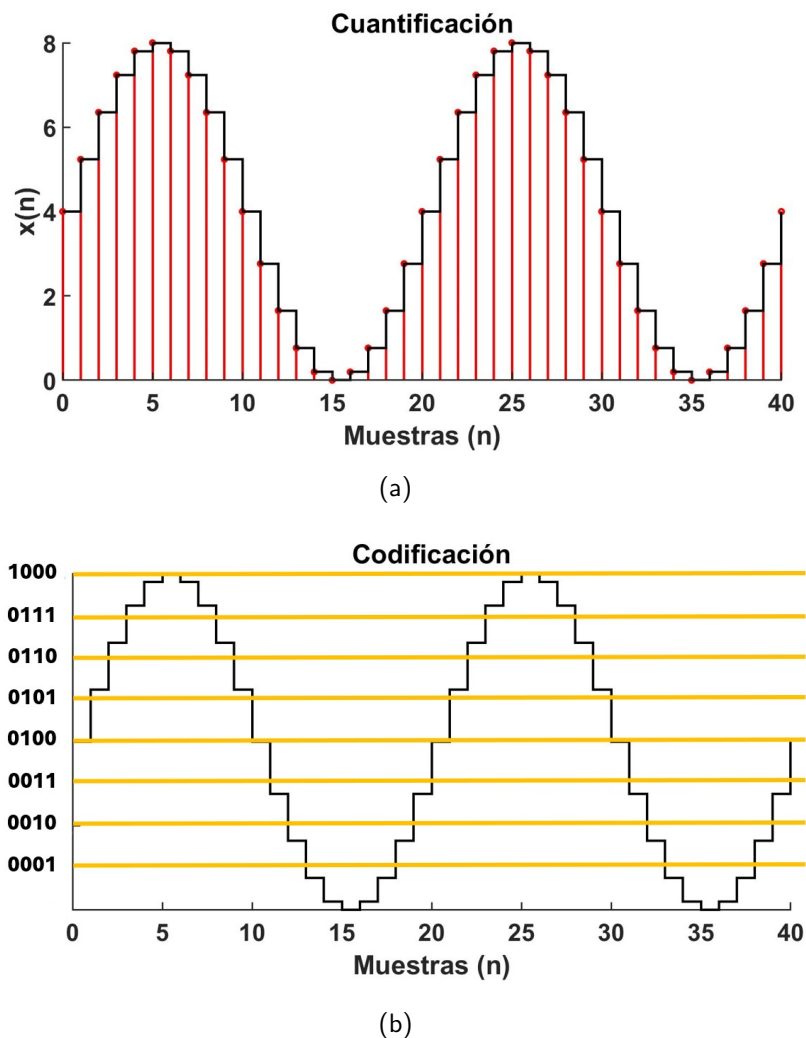


Figura 4. Codificación de una señal. La representación de los valores será de forma binaria (0,1).

Al codificar una palabra, en este caso particular utilizando palabras bisílabas, se generan numerosas muestras. En este contexto, la compresión de una señal de voz tiene como objetivo reducir la cantidad de datos requeridos para representarla, sin comprometer de manera significativa la calidad y la inteligibilidad del habla.

2.3. Señal de voz

Una señal de voz es una forma de señal acústica que resulta de la producción y transmisión de sonidos generados por las cuerdas vocales y modificados por las cavidades orales y nasales durante el habla o la vocalización (Romero & Mares, 2013).

La señal de voz captura las variaciones de presión del aire a medida que pasa a través del tracto vocal, creando un patrón característico de ondas sonoras que contiene información sobre el habla y otras características vocales, como el tono, la entonación, la intensidad y el timbre.

En general, una palabra puede ser representada como una suma de varias sinusoidales de diferentes amplitudes y frecuencias, como podemos observar en la imagen 5.

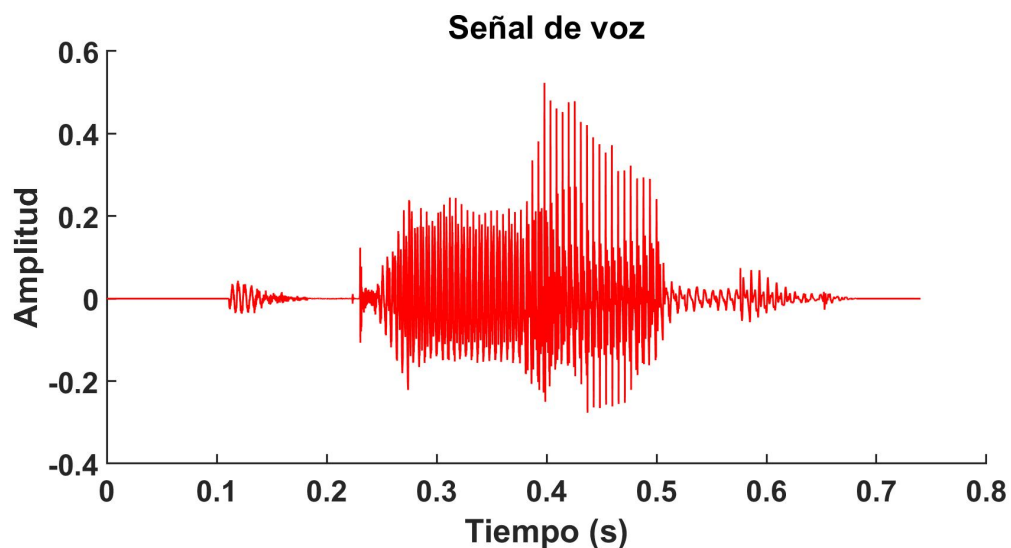


Figura 5. Ejemplo de una señal de voz. La palabra "Clavo" podemos observarla como señal de voz en esta figura, donde, en el eje horizontal tenemos el tiempo y en el eje vertical tenemos la amplitud de la señal.

La compresión de señales de voz mediante la aplicación de polinomios ortogonales es una estrategia

empleada con el propósito de reducir la cantidad de datos necesarios para representar con precisión la señal vocal. Esta técnica contribuye a la disminución del tamaño total de la señal y, por ende, a la optimización del espacio requerido para su almacenamiento. De manera similar, en el contexto de la transmisión de la señal, se logra una reducción en la tasa de datos necesarios para la transferencia de información. Esta aproximación se basa en la descomposición de la señal de voz en una serie de polinomios ortogonales. (Dusan et al., 2007).

En este enfoque metodológico, se seleccionan los polinomios ortogonales adecuados para obtener una representación de la señal vocal que sea eficaz en términos de eficiencia y calidad. La elección de estos polinomios en el contexto de la compresión de voz es un proceso que demanda una combinación de enfoque experimental y análisis detallado. La consecución de una compresión de alta calidad y rendimiento implica la identificación precisa del conjunto óptimo de polinomios que mejor se ajusten a las características únicas de la señal de voz en cuestión.

La cantidad de coeficientes empleados en el proceso de reconstrucción de una señal de voz mediante polinomios ortogonales tiene un impacto directo tanto en la calidad de la señal resultante como en la eficacia de la compresión. Al ajustar el número de coeficientes, se busca un equilibrio entre la fidelidad de la señal reconstruida y la cantidad de datos requeridos para almacenar o transmitir la señal comprimida.

La fase de compresión se ejecuta al almacenar o transmitir exclusivamente los coeficientes de expansión de mayor relevancia. Al eliminar o cuantificar los coeficientes de menor significado, se logra una reducción sustancial en la cantidad de datos necesarios para representar la señal.

Capítulo 3. Polinomios ortogonales

En este capítulo, nos adentraremos en el mundo de los polinomios ortogonales y su distinción frente a los polinomios tradicionales. La diferencia principal reside en sus propiedades y enfoques de aplicaciones específicos. Los polinomios ortogonales desempeñan un rol fundamental en transformaciones, compresiones y resolución de ecuaciones diferenciales. En contraste, los polinomios tradicionales representan expresiones algebraicas más generalizadas, empleadas para modelar y describir una amplia gama de fenómenos.

3.1. Polinomios

Un monomio es una expresión algebraica que consta de un único término. Está compuesto por dos partes: el coeficiente y una variable. El coeficiente es un número que multiplica a la variable y representa su escala. Puede ser positivo o negativo. La variable está elevada a exponentes enteros no negativos, representa la incógnita o elemento variable de la expresión algebraica, mientras que los exponentes indican el grado o la potencia a la que se eleva la variable. Un ejemplo de monomio es $5y^3$ donde el coeficiente es 5, la variable es y y el exponente es 3.

Un polinomio es una expresión algebraica que consta de la suma o resta de monomios. Los monomios que conforman un polinomio se denominan términos del polinomio, es decir, si el polinomio tiene dos términos se le denomina binomio, si tiene tres términos se le denomina trinomio y cuando el polinomio tiene más de tres términos, solo es denominado polinomio. Por ejemplo, $7x^4 - 8x^3 + 3x^2 - 5x + 2$ es un polinomio donde su variable es x .

3.2. Polinomios ortogonales

Los polinomios ortogonales son secuencias de polinomios, tales que el producto interno entre dos polinomios de la secuencia es cero si tales polinomios son diferentes (Cítores, 2015).

Sea $f(x)$ un polinomio ortogonal con respecto a la función de peso $w(x)$ en el intervalo $[a, b]$. La función de peso $w(x)$ es una función continua y no negativa en el intervalo $[a, b]$ y se utiliza para

ponderar los términos del polinomio ortogonal durante la integración. Las funciones de peso son esenciales para determinar las propiedades de ortogonalidad de los polinomios con respecto a un intervalo dado (Morey, 2019).

La propiedad de ortogonalidad se define mediante el producto interno entre dos polinomios $f(x)$ y $g(x)$ y se expresa como:

$$\int_a^b f(x) * g(x) * w(x) dx = 0, \quad (6)$$

cuando $f(x)$ y $g(x)$ son distintos polinomios ortogonales,

$$\int_a^b f(x) * f(x) * w(x) dx \neq 0, \quad (7)$$

cuando $g(x) = f(x)$, es decir, $f(x)$ es el mismo polinomio ortogonal.

Los polinomios ortogonales desempeñan un papel fundamental en diversos campos de la física. Por ejemplo, los polinomios Airy encuentran aplicaciones relevantes en óptica visual y en el análisis de aberraciones de sistemas ópticos, como se evidencia en (Gómez-Correa et al., 2020).

Asimismo, en trabajos de investigación relacionados con la segmentación de imágenes médicas, se han empleado exitosamente los polinomios de Zernike para obtener imágenes filtradas que permiten una segmentación más eficiente (Yanez, 2018).

Por otro lado, los polinomios de Bessel y las Series de Fourier son herramientas ampliamente utilizadas en el análisis de señales y sistemas. Estas técnicas se revelan especialmente eficaces para describir señales que exhiben componentes cíclicos o armónicos, como es el caso de las señales periódicas y las oscilaciones sinusoidales, tal como se detalla en Heredia (2014).

Dadas las características previamente mencionadas de cada uno de los polinomios, se han seleccionado para llevar a cabo la reconstrucción de las señales de voz. Sin embargo, es importante destacar que existe una limitada información disponible sobre la implementación de estos polinomios en la reconstrucción de señales de voz, tanto en términos de resultados objetivos como subjetivos. Este hecho subraya que este trabajo se enfoca en un área poco explorada, lo que lo

convierte en un esfuerzo pionero en el campo de la reconstrucción de señales de voz utilizando estos polinomios.

3.3. Series de Fourier

Las series de Fourier son un conjunto de técnicas matemáticas que permiten descomponer una función periódica en una suma de componentes sinusoidales (senos y cosenos) con diferentes frecuencias y amplitudes.

Se le llama serie trigonométrica de periodo 2π a toda serie de funciones de la forma:

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx), \quad (8)$$

si uno de los coeficientes a_k y b_k es diferente a cero, el grado del polinomio es k .

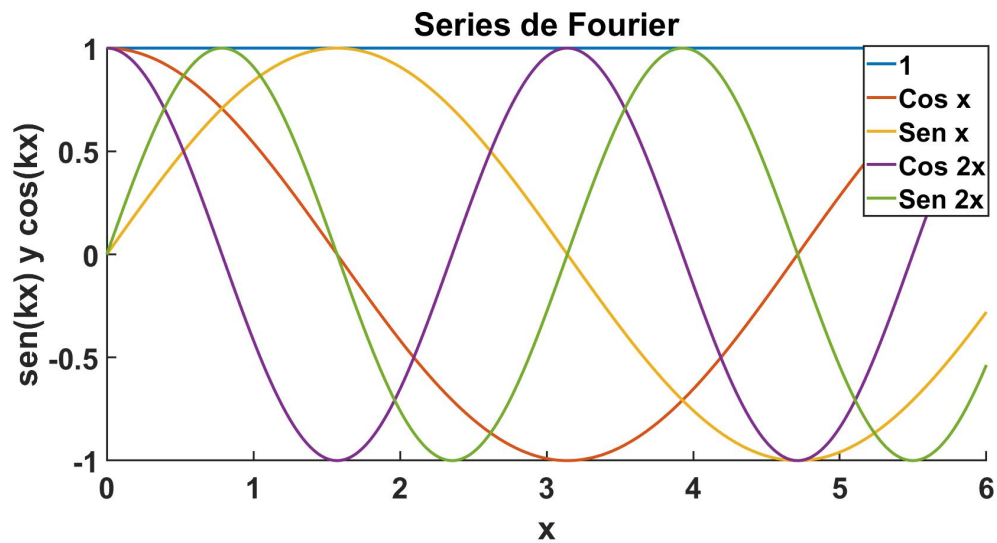


Figura 6. Primeros 5 términos de las Series de Fourier, los cuales podremos observar cada uno con un color diferente. Esta gráfica puede extenderse desde 0 hasta el infinito, aunque en la representación actual abarca hasta el valor 6.

En la figura 6, se ilustran los primeros 5 términos de las series de Fourier. El primer término corresponde a una constante con un valor de 1. El segundo término es una función coseno con una determinada frecuencia y amplitud. El tercer término es una función seno con otra frecuencia

y amplitud. A medida que se incrementan los términos, se agregan más funciones coseno y seno con diferentes frecuencias y amplitudes. En consecuencia, la suma de estos términos contribuye a la aproximación de la función original mediante la serie de Fourier.

La idea básica detrás de las series de Fourier es que cualquier función periódica puede ser descompuesta en una suma infinita de componentes sinusoidales de diferentes frecuencias que se combinan para formar la señal original. Esta descomposición se puede expresar la fórmula de la serie de Fourier, que involucra los coeficientes de Fourier, que son los pesos asignados a cada componente sinusoidal.

3.3.1. Coeficientes de Fourier

Los coeficientes de Fourier son los valores numéricos que representan las amplitudes de las componentes sinusoidales con la descomposición de una función periódica mediante la serie de Fourier.

En la descomposición de Fourier de la función f , se utilizan dos tipos de coeficientes de Fourier: los coeficientes a_k y los coeficientes b_k .

Los coeficientes a_k son los coeficientes de las componentes coseno en la serie de Fourier.

$$a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos kx dx. \quad (9)$$

Los coeficientes b_k son los coeficientes de las componentes seno en la serie de Fourier.

$$b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin kx dx. \quad (10)$$

La combinación de los coeficientes a_k y b_k en la serie de Fourier permite reconstruir la función periódica original con una aproximación adecuada dentro del intervalo de período dado.

3.4. Funciones de Bessel

Las funciones de Bessel se definen como soluciones de la ecuación diferencial de Bessel, que es una ecuación de segundo orden y lineal con coeficientes constantes. Esta ecuación está dada por (Acebrás, 2020):

$$x^2 y'' + xy' + (x^2 - p^2)y = 0. \quad (11)$$

La solución a la ecuación diferencial de Bessel es la función de Bessel que esta dada por:

$$J_p(x) = \sum_{n=0}^{\infty} (-1)^n \frac{(x/2)^{2n+p}}{n!(p+n)!}. \quad (12)$$

Los primeros 5 órdenes de las funciones de Bessel las podemos observar en la figura 7. Una descripción general de las gráficas es que en el orden $J_0(x)$ es un orden cosenoidal modificado y que oscila alrededor de cero. La amplitud de la función va disminuyendo a medida que se aleja del origen.

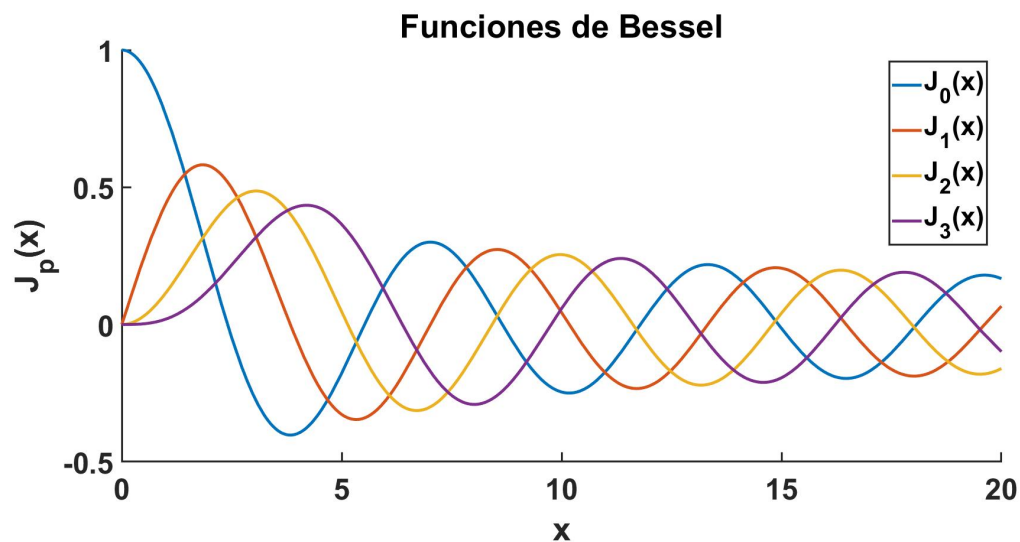


Figura 7. Funciones de Bessel para órdenes enteros $J_0(x)$, $J_1(x)$, $J_2(x)$, $J_3(x)$ y $J_4(x)$

A partir del orden $J_1(x)$, el origen de la señal es en punto cero y a medida que x aumenta, la

función oscila cerca de cero y disminuye a medida que x se aleja del origen. En el orden $J_2(x)$ tiene mayor número de oscilaciones en comparación del orden $J_1(x)$ y de igual manera, la amplitud de las oscilaciones disminuye a medida que x se aleja del origen. En general, en la figura 7 observamos los patrones oscilatorios de las funciones de Bessel y la cantidad de oscilaciones va aumentando con el orden de la función.

3.4.1. Coeficientes de Bessel

La forma en la expresamos los polinomios de Bessel es la siguiente:

$$f(x) = \sum_{n=1}^{\infty} a_n J_p(\lambda_n x) = a_1 J_p(\lambda_1 x) + a_2 J_p(\lambda_2 x) + \dots, \quad (13)$$

donde $f(x)$ está definida en el intervalo $[0, 1]$ y los λ_n son los ceros de la función de Bessel. Los ceros de algunas funciones de Bessel se encuentran en tablas matemáticas. Los ceros de los órdenes de $J_0(x)$ y $J_1(x)$ se conocen con mucha precisión.

Para mayor comodidad, nos enfocaremos en el intervalo $[0, 1]$ y, mediante un cambio de variable, podremos ajustar las fórmulas para intervalos de la forma $[0, a]$.

Supongamos que es posible escribir la función $f(x)$ como un desarrollo de la forma 13. Multiplicando por $x J_p(\lambda_m x) dx$ tenemos:

$$x f(x) J_p(\lambda_m x) = \sum_{n=0}^{\infty} x a_n J_p(\lambda_m x) J_p(\lambda_n x). \quad (14)$$

Integrando término por término entre 0 y 1 llegamos a

$$\int_0^1 x f(x) J_p(\lambda_m x) dx = \frac{1}{2} a_m J_{p+1}(\lambda_m)^2, \quad (15)$$

sustituimos m por n y obtenemos los coeficientes a_n ,

$$a_n = \frac{2}{J_{p+1}(\lambda_n)^2} \int_0^1 x f(x) J_p(\lambda_n x) dx. \quad (16)$$

3.4.2. Función Airy

La función Airy tiene un comportamiento oscilatorio similar a las funciones de Bessel, pero tienen una forma más suave y bien definida en el origen. Por esta razón, la función Airy es útil para describir la propagación de ondas en sistemas ópticos como lentes y aberraciones en el ojo humano (Gómez-Correa et al., 2020).

La función Airy es representada por:

$$\text{Airy}(u) = \frac{J_1(u)}{u}, \quad (17)$$

donde $J_1(u)$ es la función de Bessel de primer orden.

Como se mencionó anteriormente, la función Airy exhibe un comportamiento similar a la función de Bessel, como se puede apreciar en la Figura 8.

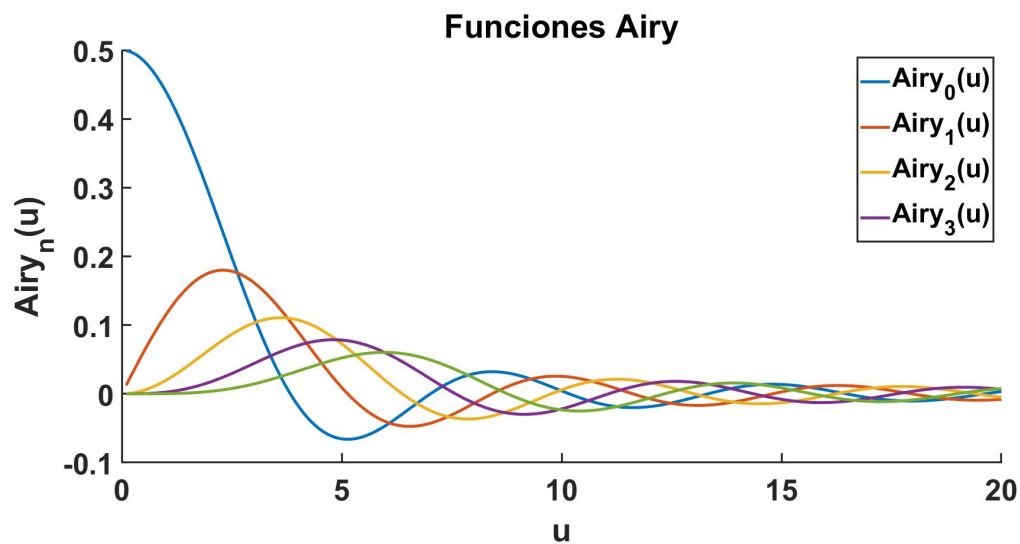


Figura 8. Función Airy para órdenes enteros $\text{Airy}_0(u)$, $\text{Airy}_1(u)$, $\text{Airy}_2(u)$ y $\text{Airy}_3(u)$

No obstante, existe una distinción significativa: en la función Airy, el valor de $Airy_0(u)$ no es igual a 1 cuando $u = 0$, sino que se establece en 0.5. A medida que se incrementa el orden de la función Airy, estos comienzan en 0, al igual que las funciones de Bessel, pero con la diferencia de que las oscilaciones son de menor amplitud.

Los polinomios Airy en u para un número de puntos enteros de n :

$$Airy_n(u) = \sum_{m=0}^n (-2)^m \frac{n!}{(n-m)!} \frac{J_{m+1}(u)}{u^{m+1}}. \quad (18)$$

3.5. Polinomios de Zernike

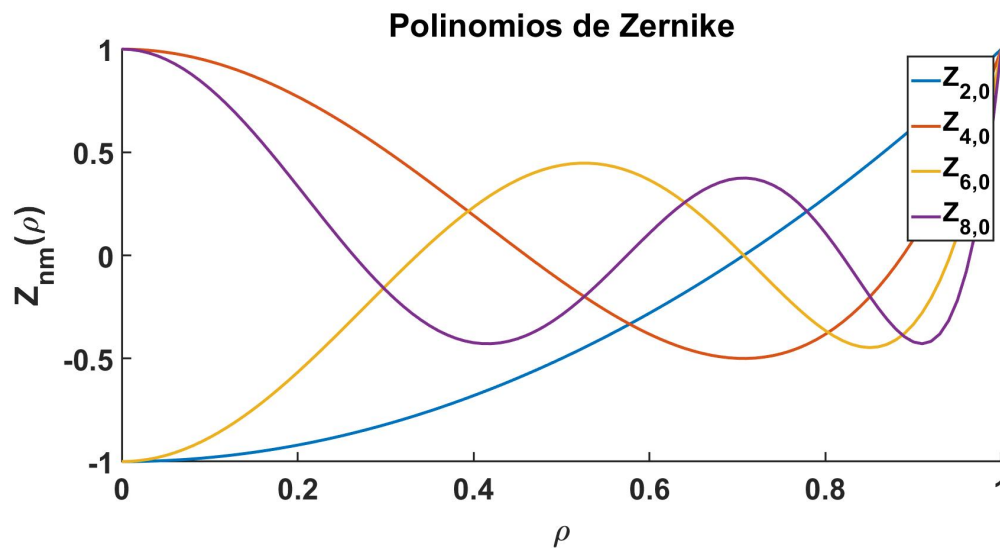


Figura 9. Polinomios de Zernike. Cuando $m = 0$, los polinomios de Zernike solo presentan simetría radial.

Los polinomios de Zernike constituyen un conjunto completo y ortogonal de polinomios complejos definidos en el interior del círculo unitario, es decir, para $x^2 + y^2 \leq 1$. Este conjunto de polinomios se denota como $Z_{nm}(\rho)$ está definido por:

$$Z_{nm}(\rho) = \sum_{s=0}^{n-|m|} \frac{(-1)^s (2n+1-s)! \rho^{n-s}}{s!(n+|m|+1-s)!(n-|m|-s)!}, \quad (19)$$

donde Z_{nm} es el polinomio de Zernike, $\rho = \sqrt{x^2 + y^2}$, m es la frecuencia angular y n el orden radial.

Los polinomios de Zernike pueden ser visualizados originalmente en coordenadas polares, pero también se pueden representar en coordenadas cartesianas mediante la correspondiente transformación. Cada polinomio tiene una forma específica, como se puede apreciar en la figura 9.

3.5.1. Coeficientes de Zernike

Los polinomios de Zernike pueden utilizarse para aproximar una señal $f(x)$. Esto significa que podemos descomponer la señal en una combinación lineal de los polinomios de Zernike, lo que nos permite aproximarla de manera eficiente y precisa. Al utilizar estos polinomios como base, podemos ajustar los coeficientes de la combinación lineal para obtener una representación que se acerque lo más posible a la señal original. Podemos obtener el valor de esos coeficientes a partir de:

$$f(x, y) = \sum_{n=0}^{\infty} \sum_{\{m: |m| \leq n\}} A_{nm} Z_{nm}(x, y), \quad (20)$$

donde A_{nm} representa los coeficientes de Zernike de orden n y repetición m . Estos coeficientes se definen mediante la siguiente expresión:

$$A_{nm} = \frac{n+1}{\pi} \int \int_{x^2+y^2 \leq 1} f(x, y) Z_{nm}(x, y) dx dy. \quad (21)$$

La ortogonalidad y la integridad de los polinomios de Zernike dan lugar a la siguiente fórmula para la reconstrucción de una señal:

$$f(x, y) = \sum_{n=0}^{n_{max}} \sum_{m=-n}^n A_{nm} Z_{nm}(x, y). \quad (22)$$

3.6. Cálculo de coeficientes por mínimos cuadrados

Para aproximar una función utilizando un polinomio tenemos la siguiente función:

$$f(x) = a_1P_1 + a_2P_2 + a_3P_3 + \cdots + a_nP_n. \quad (23)$$

El método de mínimos cuadrados es comúnmente empleado para encontrar los coeficientes de los polinomios debido a varias razones fundamentales. En primer lugar, el método busca minimizar la suma de los cuadrados de las diferencias entre los valores reales y los valores aproximados por el polinomio, lo cual resulta en un ajuste óptimo de los coeficientes que se adaptan mejor a los datos disponibles.

Además, el método de mínimos cuadrados es reconocido por su capacidad de ofrecer un ajuste robusto frente al ruido presente en los datos. Al minimizar el impacto de las perturbaciones y dar mayor peso a los puntos con menor error cuadrático, se obtiene una solución que es más resistente a las fluctuaciones o errores de medición.

Otra ventaja es que el método de mínimos cuadrados proporciona una solución única para los coeficientes del polinomio, lo cual garantiza resultados consistentes y determinísticos en el proceso de ajuste.

Asimismo, este método ofrece una gran flexibilidad, ya que puede aplicarse a polinomios de diferentes órdenes. Esto permite adaptar el grado del polinomio, a las características específicas de los datos, logrando un equilibrio adecuado entre la complejidad del modelo y la precisión del ajuste.

En resumen, el método de mínimos cuadrados se destaca como una herramienta confiable y efectiva para calcular los coeficientes de los polinomios debido a su capacidad de ofrecer un ajuste óptimo, tolerancia al ruido, solución única, flexibilidad y estabilidad numérica. Estas ventajas fundamentales respaldan la amplia utilización del método de mínimos cuadrados en el cálculo de coeficientes de polinomios.

Para comenzar, expandimos la función objetivo, denotada como 23, en términos de cada uno de

los puntos que la componen:

$$a_0 \begin{bmatrix} P0_0 \\ P0_1 \\ P0_2 \\ \vdots \\ P0_n \end{bmatrix} + a_1 \begin{bmatrix} P1_0 \\ P1_1 \\ P1_2 \\ \vdots \\ P1_n \end{bmatrix} + a_2 \begin{bmatrix} P2_0 \\ P2_1 \\ P2_2 \\ \vdots \\ P2_n \end{bmatrix} + \dots + a_n \begin{bmatrix} Pm_0 \\ Pm_1 \\ Pm_2 \\ \vdots \\ Pm_n \end{bmatrix} = \begin{bmatrix} f(0) \\ f(1) \\ f(2) \\ \vdots \\ f(n) \end{bmatrix}. \quad (24)$$

La forma matricial del sistema es $Ax = b$:

$$\begin{bmatrix} P0_0 & P1_0 & P2_0 & \cdots & Pm_0 \\ P0_1 & P1_1 & P2_1 & \cdots & Pm_1 \\ P0_2 & P1_2 & P2_2 & \cdots & Pm_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ P0_n & P1_n & P2_n & \cdots & Pm_n \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} f(0) \\ f(1) \\ f(2) \\ \vdots \\ f(n) \end{bmatrix}. \quad (25)$$

Aprovecharemos este hecho para encontrar una fórmula para x , en este caso los coeficientes, la “solución” de mínimos cuadrados.

A la matriz donde tenemos los polinomios la nombraremos A :

$$A = \begin{bmatrix} P0_0 & P1_0 & P2_0 & \cdots & Pm_0 \\ P0_1 & P1_1 & P2_1 & \cdots & Pm_1 \\ P0_2 & P1_2 & P2_2 & \cdots & Pm_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ P0_n & P1_n & P2_n & \cdots & Pm_n \end{bmatrix}. \quad (26)$$

Y B como la función a aproximar:

$$B = \begin{bmatrix} f(0) \\ f(1) \\ f(2) \\ \vdots \\ f(n) \end{bmatrix}. \quad (27)$$

La definición de la transpuesta A^T de una matriz A de tamaño $m \times n$, es una matriz de tamaño $n \times m$ donde las filas son las columnas de A y las columnas son las filas de A , en el mismo orden. En otras palabras, se intercambian las filas por las columnas.

Una propiedad importante es que la transpuesta de la suma de dos matrices es igual a la suma de las transpuestas de esas matrices, es decir, $(A + B)^T = A^T + B^T$. Esto implica que al transponer una suma de matrices, puedes transponer cada matriz por separado y luego sumarlas.

Además, la transpuesta de un producto de dos matrices sigue una regla conocida como regla del orden inverso, que establece que $(AB)^T = B^T A^T$. Esto significa que al transponer el producto de dos matrices, debes transponer cada matriz por separado en orden inverso y luego multiplicarlas.

Esto da un sistema de ecuaciones que define la solución de mínimos cuadrados:

$$A^T A x = A^T B. \quad (28)$$

Entonces tenemos la transpuesta de A por A que llamaremos D :

$$D = A^T A = \begin{bmatrix} P_{0_0} & P_{0_1} & P_{0_2} & \cdots & P_{0_n} \\ P_{1_0} & P_{1_1} & P_{1_2} & \cdots & P_{1_n} \\ P_{2_0} & P_{2_1} & P_{2_2} & \cdots & P_{2_n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ P_{m_0} & P_{m_1} & P_{m_2} & \cdots & P_{m_n} \end{bmatrix} \begin{bmatrix} P_{0_0} & P_{1_0} & P_{2_0} & \cdots & P_{m_0} \\ P_{0_1} & P_{1_1} & P_{2_1} & \cdots & P_{m_1} \\ P_{0_2} & P_{1_2} & P_{2_2} & \cdots & P_{m_2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ P_{0_n} & P_{1_n} & P_{2_n} & \cdots & P_{m_n} \end{bmatrix}, \quad (29)$$

y a la transpuesta de A por B , llamaremos E :

$$E = A^T B = \begin{bmatrix} P0_0 & P0_1 & P0_2 & \cdots & P0_n \\ P1_0 & P1_1 & P1_2 & \cdots & P1_n \\ P2_0 & P2_1 & P2_2 & \cdots & P2_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ Pm_0 & Pm_1 & Pm_2 & \cdots & Pm_n \end{bmatrix} \begin{bmatrix} f(0) \\ f(1) \\ f(2) \\ \vdots \\ f(n) \end{bmatrix}. \quad (30)$$

Colocando lo anterior de la forma $Ax = B$:

$$D \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = E. \quad (31)$$

Ahora que hemos reorganizado el sistema en esta forma, podemos determinar los valores de x , es decir, los coeficientes. Existen diversos métodos para resolver sistemas de ecuaciones lineales representados por matrices. Algunos de los métodos más utilizados son:

- Método de eliminación de Gauss-Jordan
- Método de Jacobi
- Método de Gauss-Seidel

La elección del método dependerá de las características específicas del sistema, como la estructura de la matriz, el tamaño del sistema y la precisión requerida. Cada uno de estos métodos tiene sus propias ventajas y consideraciones, por lo que es importante seleccionar el más adecuado para cada situación.

Capítulo 4. Reconstrucción de una señal mediante polinomios ortogonales

Al realizar la compresión de la voz, se tiene pérdida de calidad en la voz, lo cual puede afectar en la calidad del audio. Por ello, el objetivo principal del procedimiento de reconstrucción de una señal de voz es recuperar la información de la señal con la máxima fidelidad posible.

Cabe destacar que el proceso de reconstrucción de señales de voz puede variar según la aplicación, los métodos específicos utilizados y las características de la señal de voz original. Es fundamental considerar estos factores para obtener resultados óptimos en la reconstrucción de la señal de voz.

4.1. Procedimiento de reconstrucción

El procedimiento empleado para la reconstrucción de las señales de voz en este proyecto se llevó a cabo de la siguiente manera:

Se utilizó un programa desarrollado en Matlab R2019a para la lectura de archivos de audio almacenados en el formato WAV. Cada archivo de audio contenía dos canales: uno derecho y otro izquierdo. Se procedió a separar estos dos canales para su posterior procesamiento. Para cada uno de los canales obtenidos, se determinó el tamaño muestral de la señal. Posteriormente, se dividió la señal en cuatro secciones equidistantes.

Basado en el porcentaje elegido y el tamaño muestral de la señal, se calculó el número de coeficientes necesarios. Este número de coeficientes se distribuyó de manera uniforme entre las cuatro secciones previamente calculadas. Se procedió a la reconstrucción de las señales utilizando los polinomios seleccionados. Esto resultó en la obtención de cuatro secciones reconstruidas para cada canal (ocho en total).

Las secciones reconstruidas de ambos canales se fusionaron para obtener la señal de audio completamente reconstruida. Con la señal de audio reconstruida en manos, se calculó el error cuadrático medio normalizado.

A continuación, se proporciona una descripción detallada de este proceso, el cual también podemos observarlo en la Figura 10.

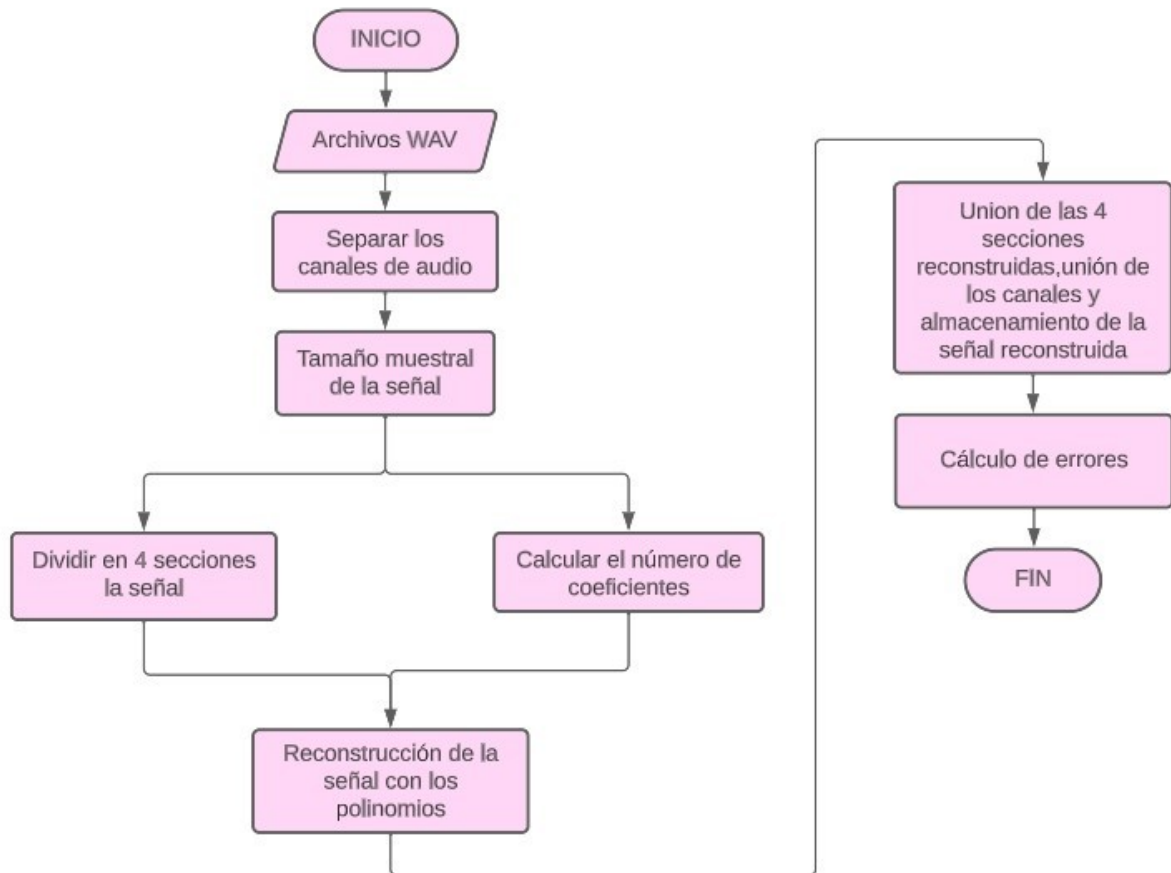


Figura 10. Diagrama a bloques de la reconstrucción de la señal de voz. Siguiendo este diagrama, se reconstruyó una señal de voz con los diferentes polinomios mencionados.

4.1.1. Leer archivo WAV

Según las características del audio que se desea almacenar, existen diferentes formatos para su almacenamiento digital. Los formatos de sonido digital sin comprimir conservan todos los datos y no eliminan, potencian ni agrupan rangos de frecuencias, lo que resulta en un tamaño considerable.

Un ejemplo de formato de audio sin comprimir es el archivo WAV o Wave Audio Format, el cual ocupa mucho espacio en el almacenamiento. Este formato es compatible con archivos tanto estéreo como mono y admite diversas velocidades de muestreo y resoluciones (García & Martínez, 2014).

El material de voz utilizado fue construido considerando las condiciones fonéticas del idioma

español (Castañeda & Pérez, 1991). Posteriormente este material fue digitalizado en formato WAV.

4.1.2. Canales de audio

Básicamente, podemos distinguir cuatro sistemas de sonido según el número de canales utilizados: mono (sistema 1.0, donde la grabación y reproducción del sonido se realiza mediante un solo canal), estéreo (sistema 2.0, donde el audio se graba y reproduce en dos canales), multicanal (donde el audio consta de tres o más canales) y sonido envolvente (surround) (a partir de 4 canales 2/2) (Prácticas, 2023).

En este proyecto, se utilizaron audios estéreo que consisten en 2 canales. Para llevar a cabo la reconstrucción, se procedió a separar cada canal y luego se reconstruyó individualmente. Posteriormente, se unieron nuevamente los canales para obtener el audio completo.

A continuación, se presenta un ejemplo que ilustra este concepto mediante la Figura 11. En esta representación gráfica, se distingue la señal de voz en formato WAV en verde, mientras que en azul y amarillo se muestran los dos canales resultantes al dividirla. Es importante destacar que, al separar estas señales, mantienen un tamaño muestral idéntico.

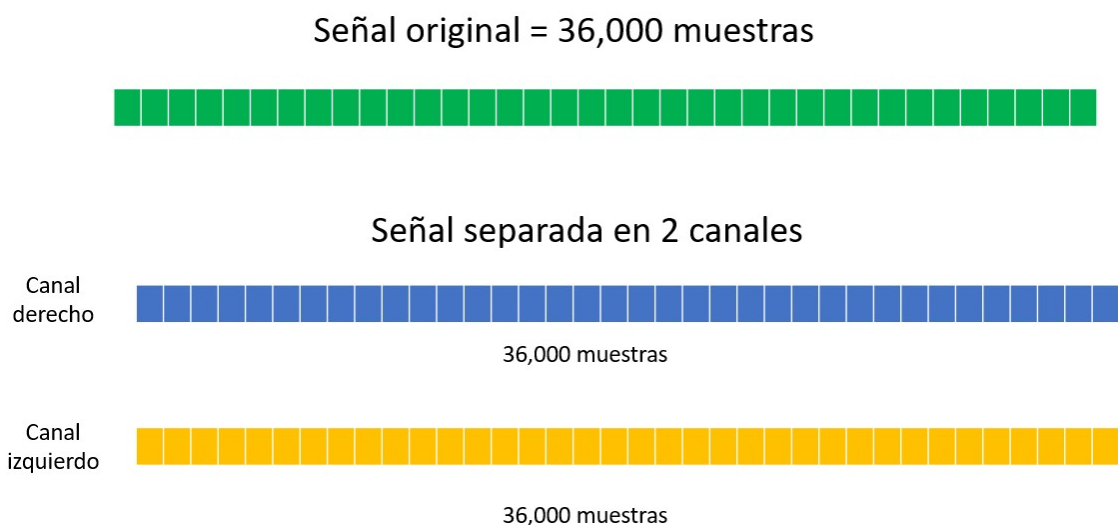


Figura 11. Ejemplificación de una señal completa y la señal separada en el canal derecho y canal izquierdo.

4.1.3. Tamaño muestral de la señal

Para este proyecto, disponemos de un conjunto de 20 grabaciones de palabras bisílabas en español, como se documenta en Castañeda & Pérez (1991). A pesar de su naturaleza bisílaba, estas palabras presentan variaciones en la cantidad de muestras, por lo que es esencial determinar el tamaño muestral de cada una de ellas. Esta información resultará fundamental para los pasos subsiguientes en el proceso de reconstrucción de la señal.

4.1.4. División del tamaño muestral en 4 secciones

Basándonos en la investigación realizada por Dusan et al. (2007), la cual sugiere la segmentación de señales de voz, se decidió dividir nuestra señal en cuatro secciones, como se muestra en la Figura 12. Para ofrecer un ejemplo concreto, hemos tomado una señal de voz compuesta por 36,000 muestras y la hemos dividido sistemáticamente en cuatro secciones, cada una de ellas conteniendo 9,000 muestras. Este mismo proceso se ha aplicado a los 20 conjuntos de datos de audio con palabras bisílabas en español.

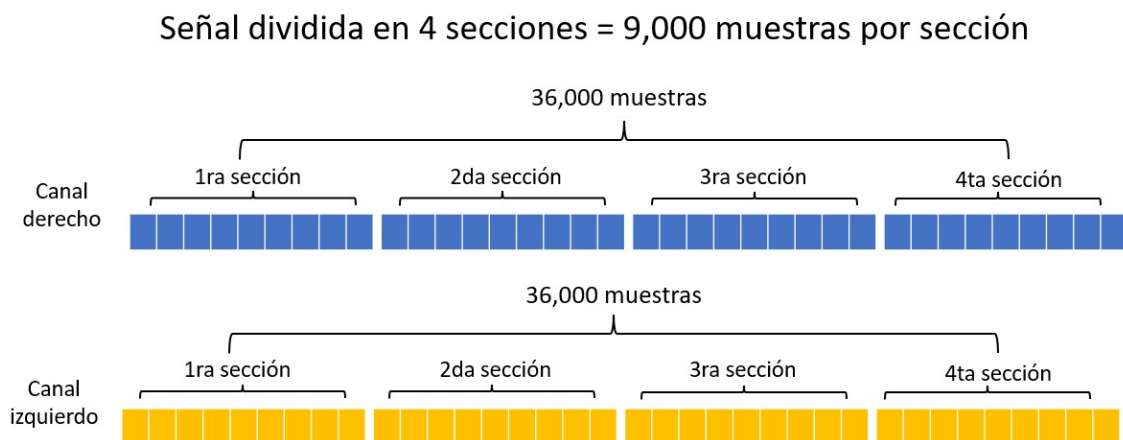


Figura 12. Ejemplo de una señal de 36,000 muestras dividida en 4 secciones.

Con el objetivo de ofrecer una representación visual de cómo se percibe la señal en la realidad, en la Figura 13 mostramos tanto la señal original como la misma señal dividida en cuatro secciones.

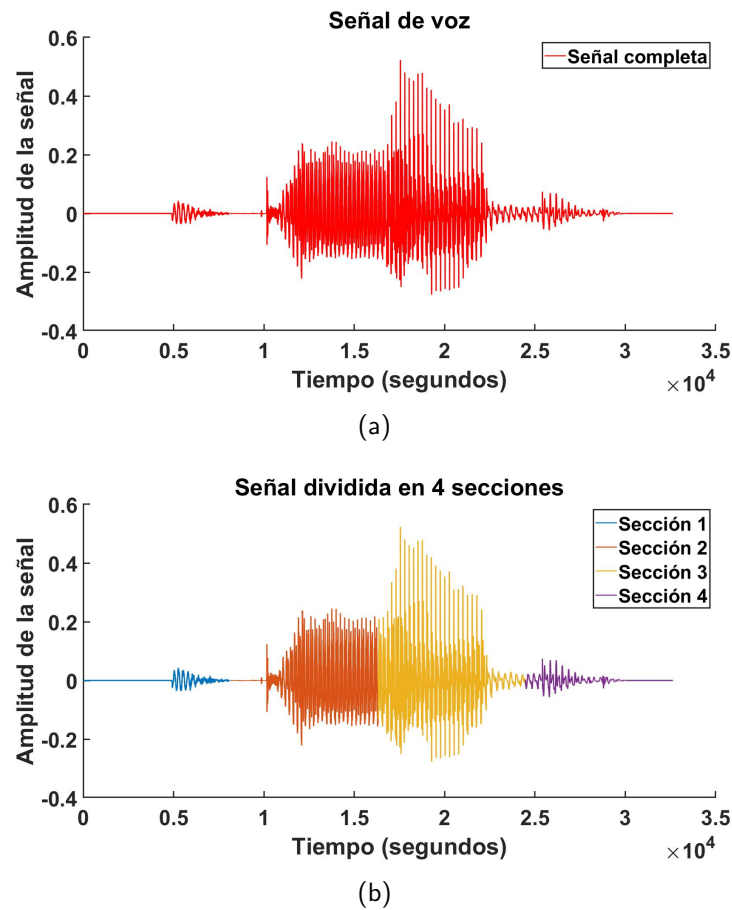


Figura 13. Comparación de la señal original y la señal dividida en 4 secciones. En la primera señal (a), se observa que todas las muestras de la señal están juntas y en la segunda señal (b) se observa que la señal está dividida en 4 secciones.

4.1.5. Cálculo del número de coeficientes

Dado que los audios varían en tamaño muestral, hemos adoptado un enfoque equitativo para calcular el número de coeficientes en todos ellos. Esto se logra calculando un porcentaje específico de las muestras totales, siendo estos porcentajes: 3 %, 4 %, 5 % y 10 %. Por ejemplo, si consideramos un audio con 36,000 muestras, el 10 % de estas correspondería a 3,600 coeficientes, como se ilustra en la Figura 14. Estos coeficientes se distribuirán en cuatro secciones, similar a lo mostrado en la Figura 12. Este procedimiento se aplicará de manera consistente para los diferentes porcentajes en todos los audios, asegurando un enfoque uniforme en el análisis de las señales de voz.

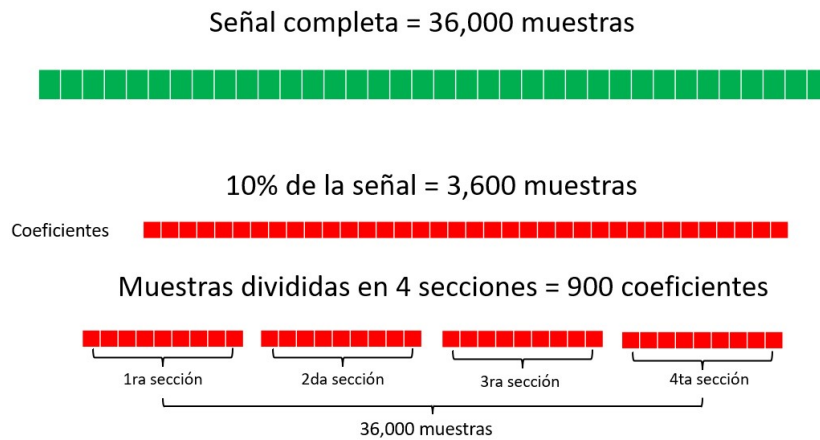


Figura 14. Ejemplo del cálculo de los coeficientes para cada audio.

4.1.6. Reconstrucción de la señal

En la Figura 15, se puede apreciar cómo el canal derecho y el canal izquierdo han sido dividido en cuatro secciones y los coeficientes también se han dividido en estas mismas cuatro secciones. A continuación, llevamos a cabo la reconstrucción de la señal utilizando los polinomios previamente estudiados, que son Airy, Bessel, Fourier y Zernike. Es importante destacar que esta reconstrucción se ejecuta por separado para cada uno de los canales de la señal.



Figura 15. Esquema del canal derecho y canal izquierdo dividido en 4 secciones y los coeficientes divididos en estas 4 secciones.

4.1.7. Unión de las 4 secciones reconstruidas

Una vez que hemos completado la reconstrucción de las cuatro secciones tanto del canal derecho como del canal izquierdo, como se muestra en la Figura 16, procedemos a unir las cuatro secciones de cada canal. De esta manera, obtenemos los canales derecho e izquierdo completos. Luego, combinamos estos canales para generar una señal de audio estéreo que puede ser guardada en un archivo en formato WAV.

Este proceso se repite para los 20 audios originales utilizando los cuatro polinomios mencionados anteriormente. Dado que realizamos la reconstrucción de las señales con cuatro porcentajes diferentes, terminamos con un total de 380 audios reconstruidos, lo que resulta de multiplicar los 20 audios iniciales por los cuatro polinomios y los cuatro diferentes porcentajes mencionados.

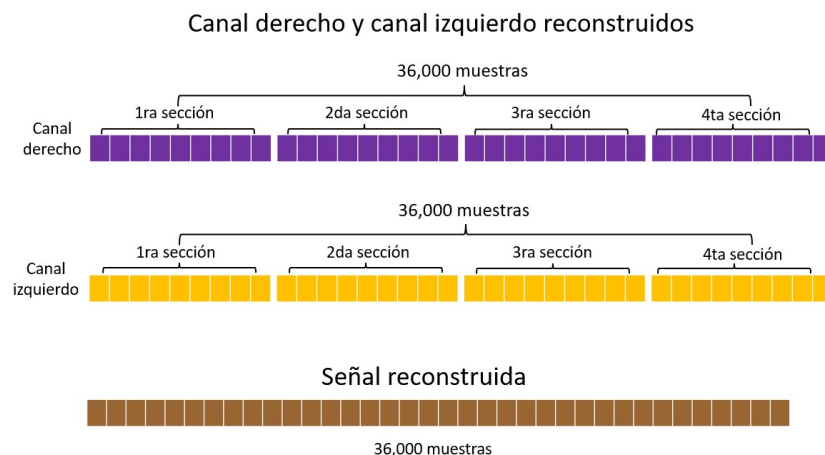


Figura 16. Representación del canal derecho y del canal izquierdo después de su reconstrucción, culminando con la obtención de la señal completamente reconstruida.

4.1.8. Cálculo del Error Cuadrático Medio Normalizado

Con el propósito de llevar a cabo una evaluación objetiva de la similitud entre la señal reconstruida y la señal original, se utilizó el Error Cuadrático Medio Normalizado (NRMSE), representado por sus siglas en inglés. El NRMSE evalúa la diferencia promedio entre los valores de la señal original y la señal reconstruida (Gómez-Correa et al., 2020). Estas métricas son esenciales para cuantificar la precisión de la reconstrucción y evaluación objetiva del proceso.

Capítulo 5. Pruebas objetivas y subjetivas

La evaluación de las señales reconstruidas en este proyecto se llevó a cabo desde dos perspectivas distintas: objetiva y subjetiva. En el análisis objetivo, se utilizaron métricas como el error cuadrático medio normalizado para obtener una evaluación cuantitativa del desempeño. Mientras que desde una perspectiva subjetiva, se recopilaron respuestas de los participantes para evaluar la calidad y la aceptación de las señales reconstruidas en términos de la percepción de los usuarios. Esta combinación de enfoques nos proporcionó una comprensión completa de las características de rendimiento y la percepción de calidad de las señales reconstruidas.

5.1. Pruebas objetivas

Las pruebas objetivas son métodos o técnicas utilizadas para evaluar de manera cuantitativa y precisa la calidad o rendimiento de un sistema, producto o proceso. Estas pruebas se basan en medidas objetivas, generalmente utilizando métricas o algoritmos predefinidos, para realizar una evaluación imparcial y reproducible, en la cuál no afecta el juicio u opinión del participante o de la persona que coordina la prueba (Vazquez, 2017).

En el contexto de señales de audio, las pruebas objetivas se utilizan para medir la calidad perceptiva de la señal reconstruida o comprimida. Las pruebas objetivas proporcionan una forma cuantitativa de medir el rendimiento y la calidad de una señal o un sistema, y son especialmente útiles cuando se desea una evaluación objetiva y comparativa de diferentes técnicas o en este caso, de los diferentes polinomios.

5.1.1. Errores

El error cuadrático medio normalizado (en inglés, Normalized root mean square error(NRMSE)), se expresa en la siguiente ecuación donde Z_j es la señal original y \hat{Z}_j es la señal reconstruida (Gómez-Correa et al., 2020).

$$NRMSE = \sqrt{\frac{\sum |Z_j - \hat{Z}_j|^2}{\sum |Z_j|^2}}. \quad (32)$$

El NRMSE se basa en el cálculo del error cuadrático medio (RMSE) normalizado por la escala de los datos de referencia. Proporciona una medida de la discrepancia promedio entre los valores estimados y los valores de referencia, teniendo en cuenta la variabilidad y amplitud de los datos.

5.2. Pruebas subjetivas

A diferencia de las pruebas objetivas, las pruebas subjetivas se ven afectadas por el juicio y la opinión tanto del participante como de la persona que coordina la prueba. Esto se debe a que las preguntas u opiniones son abiertas, abarcando soluciones de problemas, análisis de situaciones, interpretación de casos, entre otros aspectos (Vazquez, 2017).

En las pruebas subjetivas, los resultados pueden variar significativamente entre las personas debido a una serie de factores, como sesgos personales, preferencias individuales, expectativas personales y la interpretación subjetiva de los estímulos o situaciones evaluadas.

Un claro ejemplo de esta variabilidad se observa en la diferencia de opiniones entre un ingeniero de audio con experiencia en el campo y una persona sin experiencia en el ámbito de audio. Dependiendo de las características de los audios, algunas personas podrían calificar el ruido de fondo de una señal de manera muy distinta; algunos ni siquiera lo notarían, mientras que otros podrían considerarlo sumamente molesto.

Estas pruebas subjetivas se emplean con el propósito de recopilar información y opiniones de los usuarios o participantes, centrándose en aspectos subjetivos como la satisfacción, preferencias y percepciones individuales.

Este tipo de mediciones tiene diversas aplicaciones en ámbitos sociales como la educación, el entretenimiento y la ingeniería. Normalmente, se utiliza para evaluar la acústica natural de una sala de conferencias o aulas, evaluar un sistema de intercomunicación a distancia o evaluar cómo contribuye un sistema electroacústico a mejorar la inteligibilidad del habla en una sala.

La calidad del audio puede experimentar variaciones debido a múltiples factores, entre los que se incluyen la calidad del sonido, la presencia de ruido de fondo, la pronunciación del hablante, la calidad del equipo utilizado para la grabación o reproducción y la existencia de interferencias u otros elementos que puedan dificultar la audición clara.

Para llevar a cabo la prueba subjetiva en este proyecto, se ha desarrollado una página web utilizando la plataforma Google Sites (<https://sites.google.com/cicese.edu.mx/prueba-subjetiva/inicio>). En esta página, se proporciona al participante información clara sobre el concepto de una prueba subjetiva y se ofrecen respuestas a preguntas frecuentes relacionadas. Además, se incluye un breve cuestionario de audición que el participante deberá completar.

La parte principal de la prueba consiste en cinco secciones diferentes, cada una con diecisiete audios para evaluar. El participante tendrá la tarea de calificar estos audios en una escala que va desde “pésimo” hasta “excelente”. Además, se han incorporado los audios originales entre las opciones para evaluar, lo que permitirá observar la respuesta del participante con los audios originales.

Este enfoque proporcionará una visión amplia y detallada de la opinión subjetiva de los participantes sobre los diferentes audios, lo que resultará en una valiosa retroalimentación para el proyecto. A continuación, se detallará cada una de las secciones que componen la página web diseñada.

5.2.1. Inicio de la página web

Se desarrolló una página web en la plataforma de Google Sites para facilitar la evaluación de los audios. Se eligió esta plataforma debido a su facilidad de uso y a la amplia gama de herramientas disponibles, lo que permite que cualquier participante pueda acceder sin dificultad. Además, se implementaron formularios en los cuales los participantes podían ingresar la información requerida. Estos formularios se integraron en la página web junto con los audios reconstruidos.

Al acceder a la página web, como se muestra en la figura 17, se dará la bienvenida al participante con una breve y clara explicación sobre el propósito de la prueba auditiva. A continuación, se presentará un resumen completo de la prueba, donde se informará al participante sobre los objetivos

que se pretenden alcanzar y se destacará la presencia de instrucciones en cada sección, con el fin de evitar confusiones y garantizar una interacción fácil en la página.



Figura 17. Presentación de la página web. Se muestra la forma en la que el participante tendrá el primer contacto con la página web, donde se le explica qué es una prueba subjetiva y el objetivo de esta prueba.

Es fundamental mencionar desde el inicio de la prueba que los datos personales serán tratados de manera confidencial y no serán divulgados. Esta medida se implementa para brindar seguridad al participante y permitirle realizar la prueba en un ambiente tranquilo y confiable.

Se tiene un consentimiento informado como se observa en la figura 18 donde se explica que esta es una prueba subjetiva de audio, nos interesa tu opinión acerca de diferentes audios y decirnos cómo calificaría cada audio y conocer cómo se perciben y entienden los audios que han sido reconstruidos por polinomios ortogonales.

Además, se enfatiza que la prueba es no invasiva, lo que significa que el participante no experimentará ningún daño físico durante ni después de la misma. Asimismo, se aclara que no se ofrecerá ninguna compensación económica por realizar la prueba, ni se aplicará ningún costo económico al participante.

Por último, se proporcionan los datos de contacto de la investigadora principal, con el objetivo de resolver cualquier duda o escuchar cualquier queja que los participantes puedan tener. Esto se realiza para asegurar una comunicación efectiva y una respuesta rápida a las inquietudes de los participantes.

50 ANIVERSARIO
CICESE

Inicio Test de audición 1ra sección 2da sección 3ra sección 4ta sección 5ta sección Q

Consentimiento informado

1. ¿Dónde se llevará a cabo esta investigación?

Esta investigación se llevará a cabo de manera virtual, a través de la plataforma de Google Sites y Google Forms.

2. ¿Cuál es el objetivo de esta investigación?

Conocer cómo se perciben y entienden los audios que han sido reconstruidos por polinomios ortogonales.

3. ¿Estoy obligado a participar?

Su participación es voluntaria, anónima y confidencial; no tiene que participar forzosamente. No habrá impacto negativo alguno si decide no participar en la investigación.

4. ¿En qué consistirá mi participación y cuánto durará?

Su participación consistirá en lo siguiente:

- Escuchar 17 diferentes audios de cada sección, las cuales serán 5.
- Evaluar la calidad de cada uno de los audios .
- Solo se puede participar una vez .
- La participación tendrá una duración aproximada de 5 minutos.
- Si está de acuerdo en participar, se le pide rellenar el test de audición de la siguiente pestaña de la página web.

5. ¿Cuáles son los posibles riesgos de formar parte de esta investigación?

No existe riesgo alguno al formar parte de esta investigación.

Figura 18. Consentimiento informado. Se contestan las posibles preguntas que pueda tener el participante al desarrollar la prueba y en qué consiste la misma.

Al finalizar esta sección de la página web, se incluyen las aclaraciones generales sobre la prueba. Se destaca que la participación en la prueba es completamente voluntaria y que en cualquier momento se puede solicitar más información acerca de la misma. Además, se garantiza la confidencialidad de los datos de identificación de cada participante.

Si no surgen más dudas, se anima al participante a proceder a la siguiente pestaña, donde encontrará la carta de consentimiento informado. Esta sección es fundamental para que el participante pueda revisar y dar su consentimiento formal antes de continuar con la prueba.

5.2.2. Cuestionario de audición

En la segunda pestaña de la página web, el participante tiene la oportunidad de aceptar que ha sido informado sobre la prueba de audición y que participa de forma voluntaria. Asimismo, se le proporciona información sobre los posibles riesgos asociados, aunque se destaca que el procedimiento se considera de bajo riesgo. Se asegura al participante que todas sus preguntas han sido respondidas de manera clara y satisfactoria.

Es importante destacar que la prueba subjetiva se llevó a cabo con un grupo de 20 personas voluntarias. No se cuenta con información sobre la salud auditiva de dichas personas, es decir, no

se tienen resultados de audiometrías, que son mediciones de la agudeza auditiva en relación con las diferentes frecuencias de sonido (Olvera & Queb, 2017).

Por lo tanto, se lleva a cabo un cuestionario de audición en el que se solicita el nombre completo del participante para llevar un registro de sus respuestas. También se requiere la edad del participante, ya que la prueba está dirigida a personas de entre 18 y 45 años. Esta selección de edad se debe principalmente a la intención de evitar problemas relacionados con la presbiacusia o la pérdida auditiva (Varela-Nieto & Rivera, 2010).

El cuestionario inicial consta de una serie de preguntas de respuesta binaria (sí o no). A continuación se presenta el cuestionario:

1. ¿Le resulta fácil comunicarse en conversaciones cara a cara cuando no hay ruido?
2. ¿Le resulta fácil comprender el habla cuando hay mucho ruido de fondo (por ejemplo, en restaurantes, aeropuertos o reuniones)?
3. ¿Le resulta fácil oír a su familia, amigos y vecinos?
4. ¿Le resulta fácil oír a través del teléfono?
5. ¿Le resulta fácil escuchar películas, conferencias, etc.?
6. ¿Pide que le repitan las palabras frecuentemente?
7. ¿Escucha menos por un oído?
8. ¿Usas frecuentemente audífonos?
9. Si tu respuesta anterior fue Sí, ¿Cuánto tiempo aproximadamente utilizas audífonos en el día?

Además, se pide al participante que informe si experimenta una disminución de la audición en alguno de sus oídos. Dado que la prueba se enfoca en jóvenes adultos, se reconoce que el uso frecuente de audífonos es común en este grupo. Por lo tanto, se busca conocer si el participante utiliza audífonos regularmente y durante cuánto tiempo al día los utiliza aproximadamente.

5.2.3. Secciones de la página web

En la primera sección, se llevarán a cabo las pruebas iniciales. Para garantizar una adecuada realización de las pruebas en todas las secciones, se proporcionan instrucciones generales que deben seguirse. Además, se ha incluido un audio de prueba con el fin de permitir al participante ajustar el volumen de su dispositivo a un nivel cómodo. De esta manera, no será necesario realizar cambios en el volumen durante la prueba.

Se solicita al participante que reproduzca el audio de prueba una sola vez y que lo califique en términos de calidad, pudiendo elegir entre las siguientes categorías: excelente, bueno, regular, malo y pésimo. Una vez que haya terminado de calificar, simplemente debe presionar el botón de enviar. Cada una de las 5 secciones consta de 17 audios.

5.2.4. Orden de los audios

Se realizaron pruebas de reconstrucción utilizando cuatro polinomios diferentes y con coeficientes del 3 %, 4 %, 5 % y 10 %. Esto resultó en un total de 16 audios reconstruidos para cada palabra, además del audio original, lo que suma un total de 17 archivos de audio por palabra. Multiplicado por los cinco audios utilizados, se obtuvo un total de 85 audios para la prueba subjetiva.

Los audios fueron organizados siguiendo un diseño de cuadrado latino, el cual es una técnica particular para estructurar información utilizando números enteros, letras del alfabeto u otros símbolos. Este enfoque, propuesto por Fisher (1992), también encuentra utilidad en la investigación experimental para controlar los efectos de variables no deseadas.

Dentro del ámbito de la experimentación subjetiva, el uso de cuadrados latinos se ha convertido en una herramienta estadística altamente eficaz. Esto se debe a que cuando los sujetos se exponen sucesivamente a varios estímulos, en este caso estímulos acústicos, es lógico que la percepción del sujeto se vea influenciada tanto por experiencias previas como por la cantidad de procesos históricos que ha experimentado. Cuando un estímulo llega al cerebro, deja una huella que afecta la percepción de los sonidos subsiguientes, ya sea enmascarándolos o resaltándolos, como mencionado por (Keedwell & Dénes, 2015).

El diseño de cuadrado latino utilizado se basó en la distribución aleatoria de los tratamientos, con restricciones en la agrupación tanto en columnas como en filas (bloques). Esto permite eliminar la variabilidad del error experimental asociada con ambos efectos, ya que cada tratamiento ocurre el mismo número de veces (una sola vez) en cada fila y columna, como se describe en (Osorio et al., 1987).

Capítulo 6. Resultados

En este capítulo, se presentarán a detalle los resultados obtenidos tanto de las pruebas objetivas como de las pruebas subjetivas. En primer lugar, se describirán los resultados de las pruebas objetivas realizadas en un conjunto de 20 audios seleccionados al azar. Estos audios fueron reconstruidos utilizando diferentes números de coeficientes, lo que permitió obtener el registro del tiempo promedio requerido para la reconstrucción de cada audio.

Dado el tamaño significativo del grupo de audios reconstruidos, se optó por realizar pruebas subjetivas con un conjunto de 5 audios seleccionados al azar. Esto se hizo para evitar la fatiga de los participantes, que podría conducir a respuestas sesgadas o pruebas inconclusas.

6.1. Información de los 20 audios

En el primer anexo del artículo de Padilla-Ortiz & Orduña-Bustamante (2021), se presenta un conjunto de 191 palabras en español obtenidas de la investigación realizada por Castañeda & Pérez (1991). De este conjunto de palabras, se seleccionaron aleatoriamente 20 palabras para llevar a cabo la reconstrucción de las señales. Las palabras elegidas son las siguientes: Asno, cero, cielo, deme, diego, disco, fierro, freno, niquel, norte, pera, piano, pleno, reto, saco, suela, tape, tapia, tigre y tina.

Como se puede apreciar en la Tabla 1, se presentan los 20 archivos de audio correspondientes a las palabras mencionadas anteriormente. Cada uno de estos archivos se ha sometido a un proceso de reconstrucción utilizando distintos polinomios y porcentajes de coeficientes. Esto ha dado como resultado un total de 560 archivos de audio. Este número se obtiene al multiplicar los 20 archivos de audio por los 4 polinomios mencionados y por los 7 porcentajes de coeficientes utilizados en el proceso de reconstrucción.

Como se ha mencionado previamente, cada una de las palabras en los 20 audios presenta un número distinto de muestras. La Tabla 2 proporciona información sobre cada palabra en estos 20 audios, indicando el número de muestras correspondiente.

Tabla 1. Total de audios reconstruidos.

Audios	Polinomios	Porcentaje de Coeficientes	Total de audios reconstruidos
20	Airy Bessel Fourier Zernike	1 %	560
		2 %	
		3 %	
		4 %	
		5 %	
		10 %	
		15 %	

Es importante notar que al seleccionar un determinado porcentaje de coeficientes para cada audio, este número de muestras variará, ya que afecta la cantidad de información retenida en la señal reconstruida.

Tabla 2. Total de audios reconstruidos.

Audio	N.Muestras	Numero de muestras por porcentaje						
		1 %	2 %	3 %	4 %	5 %	10 %	15 %
Asno	31,312	313	626	939	1,252	1,565	3,131	4,696
Cero	34,840	348	696	1,045	1,393	1,742	3,484	5,226
Cielo	42,337	423	846	1,270	1,693	2,116	4,233	6,350
Deme	31,753	317	635	952	1,270	1,587	3,175	4,762
Diego	36,604	366	732	1,098	1,464	1,830	3,660	5,490
Disco	36,163	361	723	1,084	1,446	1,808	3,616	5,424
Fierro	32,635	326	652	979	1,305	1,631	3,263	4,895
Freno	31,312	313	626	939	1,252	1,565	3,131	4,696
Niquel	35,281	352	705	1,058	1,411	1,764	3,528	5,292
Norte	36,163	361	723	1,084	1,446	1,808	3,616	5,424
Pera	24,256	242	485	727	970	1,212	2,425	3,638
Piano	28,225	282	564	846	1,129	1,411	2,822	4,233
Pleno	26,020	260	520	780	1,040	1,301	2,602	3,903
Reto	32,635	326	652	979	1,305	1,631	3,263	4,895
Saco	37,486	374	749	1,124	1,499	1,874	3,748	5,622
Suela	42,337	423	846	1,270	1,693	2,116	4,233	6,350
Tape	25,579	255	511	767	1,023	1,278	2,557	3,836
Tapia	29,548	295	590	886	1,181	1,477	2,954	4,432
Tigre	26,902	269	538	807	1,076	1,345	2,690	4,035
Tina	26,020	260	520	780	1,040	1,301	2,602	3,903

En la Figura 19, se presenta el gráfico del error cuadrático medio normalizado (ECMN) de los 20

audios. Esta gráfica está dividida en cuatro secciones, que corresponden a los cuatro polinomios implementados: Airy, Bessel, Fourier y Zernike. Cada color en la gráfica representa un porcentaje específico de coeficientes, como se mencionó anteriormente.

Es importante destacar que a medida que el número de coeficientes utilizados en la reconstrucción aumenta, el valor del error cuadrático medio normalizado tiende a disminuir. Este comportamiento sugiere que el aumento en la cantidad de coeficientes contribuye a una mejora en la precisión de la reconstrucción de las señales de audio.

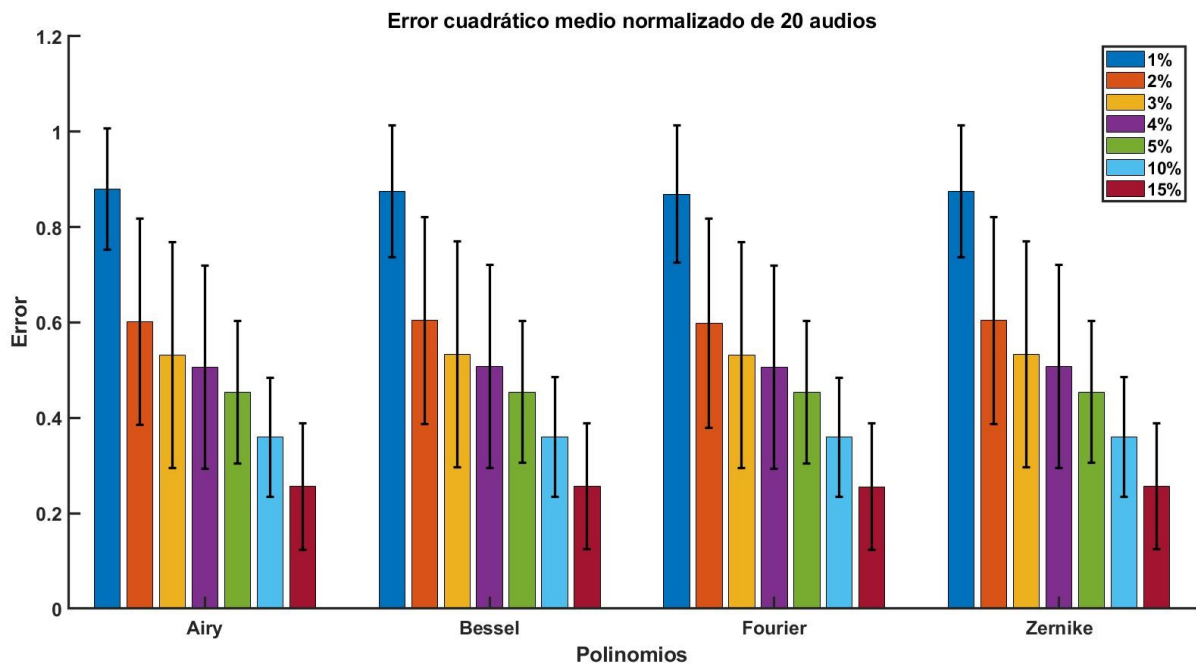


Figura 19. Error cuadrático medio normalizado. Se observa cada uno de los polinomios con sus respectivos porcentajes de coeficientes que están representados con un color diferente.

Es cierto que en el contexto de errores, utilizar más decimales no siempre es necesario ni significativo. En muchos casos, los resultados pueden expresarse con la precisión adecuada sin la necesidad de incluir un número excesivo de decimales.

Por otro lado, en situaciones donde se requiere una alta precisión o cuando el error tiene un impacto significativo en los resultados y conclusiones, puede ser apropiado utilizar un mayor número de decimales para expresarlo con mayor exactitud.

En el caso específico mencionado, se tomó la decisión de expresar los resultados con una precisión

de 5 decimales, con el fin de proporcionar detalles y precisión adicionales. Esto se puede apreciar claramente en la Tabla 3, donde se muestran los resultados de los errores cuadráticos medios normalizados (ECMN) para los diferentes polinomios utilizados. Es notable que las Series de Fourier exhiben un error menor en comparación con los demás polinomios.

Tabla 3. Error cuadrático medio normalizado de los 20 audios.

ECMN de 20 audios				
Porcentaje	Airy	Bessel	Fourier	Zernike
1 %	0.87955	0.87534	0.86939	0.87534
2 %	0.60201	0.60441	0.59849	0.60441
3 %	0.53188	0.53272	0.53175	0.53272
4 %	0.50672	0.50759	0.50657	0.50759
5 %	0.45428	0.45435	0.45367	0.45435
10 %	0.35962	0.35974	0.35939	0.35974
15 %	0.25599	0.25636	0.25579	0.25636

En la Tabla 4, se presenta la desviación estándar de los 20 audios de una manera que facilita la visualización y comprensión de los datos obtenidos.

Tabla 4. Desviación estandar de los 20 audios.

Desviación estándar de 20 audios				
Porcentaje	Airy	Bessel	Fourier	Zernike
1 %	0.12758	0.13832	0.14316	0.13832
2 %	0.21585	0.21706	0.21888	0.21706
3 %	0.23727	0.23693	0.23691	0.23693
4 %	0.21307	0.21321	0.21265	0.21321
5 %	0.14966	0.14884	0.14882	0.14884
10 %	0.12513	0.12512	0.12521	0.12512
15 %	0.13221	0.13208	0.13229	0.13208

Para la reconstrucción de los 20 audios, teniendo en cuenta las diferencias en el tamaño de muestras entre ellos, se calculó el tiempo promedio de reconstrucción, como se muestra en la Figura 20. Los resultados indican que el tiempo promedio de reconstrucción para los polinomios Airy es de 73.8 segundos, mientras que para los polinomios Bessel es de 86.75 segundos. Las Series de Fourier muestran el menor tiempo promedio, con un valor de 16.5 segundos, mientras que los polinomios Zernike requieren un tiempo promedio mayor, llegando a 142.65 segundos, siendo los que demandan más tiempo en la reconstrucción.

Además, se calculó la desviación estándar, que refleja la variabilidad en los tiempos de reconstrucción. Para los polinomios Airy, la desviación estándar es de 32.44, para los polinomios Bessel es de 40.0025, para las Series de Fourier es de 7.1708 y para los polinomios de Zernike es de 74.2615. Estos datos proporcionan una visión detallada del tiempo requerido para la reconstrucción de los audios y su variabilidad asociada.

Si bien el tiempo de reconstrucción de una señal puede ser crucial en diversas aplicaciones, en este proyecto no se consideró como el factor principal. El enfoque principal se centró en determinar el menor error cuadrático medio normalizado y evaluar la calidad de las reconstrucciones a través de la puntuación proporcionada por los participantes.

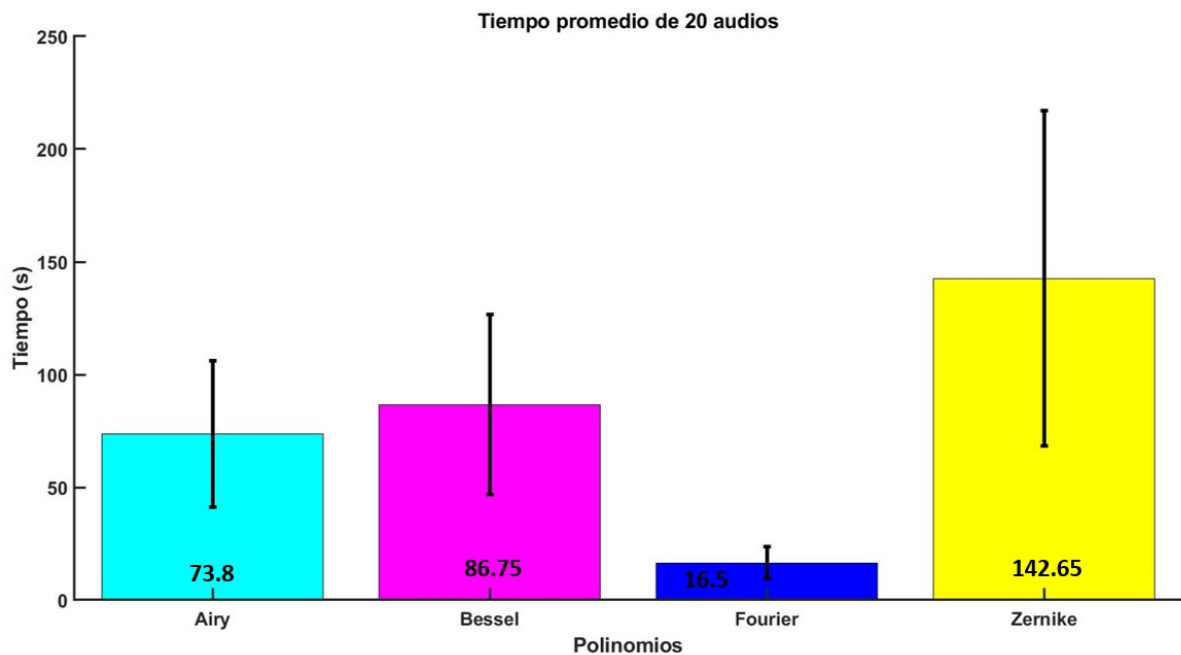
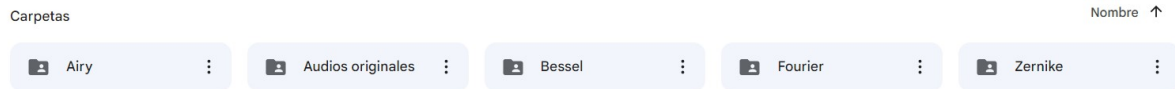


Figura 20. Tiempo promedio de la reconstrucción de los 20 audios elegidos al azar por cada polinomio. La desviación estándar de los audios son tan altas ya que los audios tienen diferentes números de muestras y esto hace que algunos audios se reconstruyan más rápido que otros.

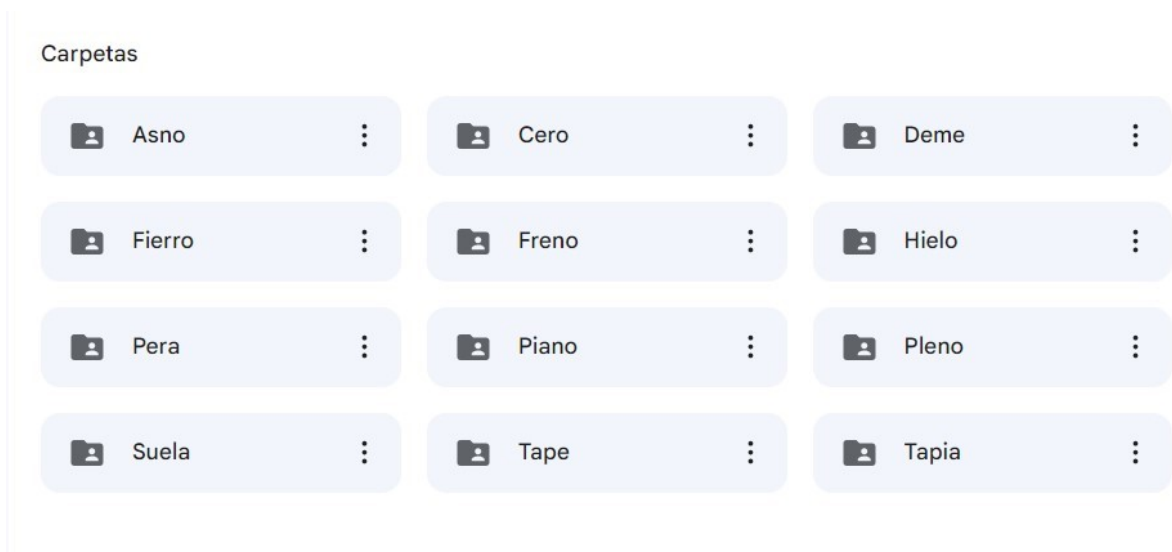
Los archivos de audio se ubican en un directorio de Google Drive, al cual se puede acceder mediante el siguiente enlace: https://n9.c1/audios_reconstruidos Como se ilustra en la Figura 21 (a), esta carpeta está organizada en cinco subdirectorios: audios originales y audios reconstruidos utilizando los polinomios Airy, Bessel, Fourier y Zernike.

Cada uno de estos subdirectorios incluye carpetas individuales correspondientes a las palabras, tal

como se muestra en la Figura (b). Dentro de cada carpeta de palabras, se pueden encontrar los archivos de audio correspondientes a los diferentes porcentajes de reconstrucción. Este esquema de organización permite una fácil identificación y acceso a los audios reconstruidos, garantizando una estructura ordenada y clara.



(a)



(b)

Figura 21. El orden de los audios reconstruidos se encuentra estructurado de la siguiente manera: Figura a) En esta figura se presentan las cinco carpetas principales donde se almacenan los audios. Figura b) Cada una de las carpetas en esta figura lleva el nombre de una palabra específica.

6.2. Audios seleccionados para las pruebas objetivas y subjetivas

Al inicio de este capítulo, se menciona que se dispone de un total de 560 audios reconstruidos. Sin embargo, se redujo el número de coeficientes utilizados en la reconstrucción de estas señales, con el propósito de garantizar una variedad de reconstrucciones sin que los participantes se familiarizaran en exceso con los audios. En total, se trabajó con un conjunto de 80 audios para llevar a cabo las pruebas objetivas y subjetivas. Los audios seleccionados al azar para las pruebas fueron: Pleno, Piano, Pera, Deme y Fierro, como se detalla en la Tabla 5.

Tabla 5. Tabla de palabras utilizadas.

Audios	Polinomios	Porcentaje de Coeficientes	Total de audios reconstruidos
Pleno	Airy	3 %	80
Piano	Bessel	4 %	
Pera	Fourier	5 %	
Deme	Zernike	10 %	
Fierro			

6.3. Pruebas objetivas

6.3.1. Error cuadrático medio normalizado

El error cuadrático medio normalizado ahora se presenta exclusivamente para las 5 palabras previamente seleccionadas, como se mencionó anteriormente. Además, se redujo el número de coeficientes utilizados, lo que se refleja en la Figura 22.

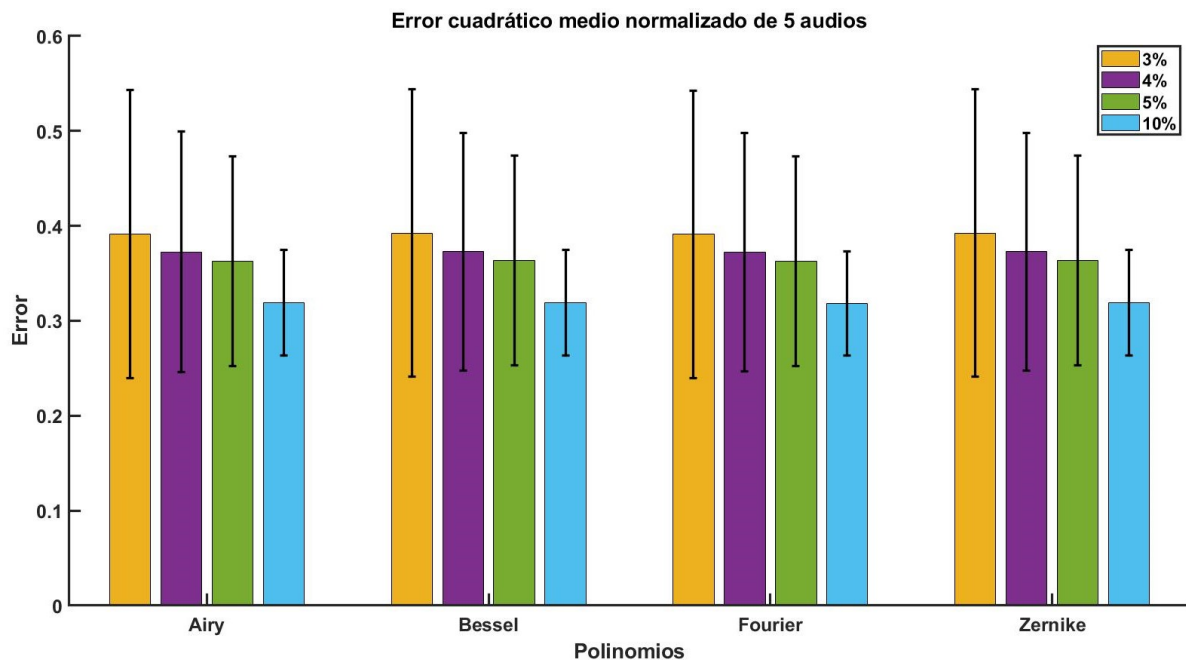


Figura 22. Error cuadrático medio normalizado. Se observa cada uno de los polinomios con sus respectivos porcentajes de coeficientes que están representados con un color diferente.

En esta figura, se utilizan diferentes colores para representar distintos porcentajes de coeficientes:

el 3 % se muestra en amarillo, el 4 % en morado, el 5 % en verde y el 10 % en azul claro. Esta representación visual permite apreciar la variación del ECMN en función del número de coeficientes utilizados para la reconstrucción de las señales de audio.

Con respecto al 3 % de los coeficientes, los polinomios Bessel y Zernike presentaron un error igual de 0.39243. Los polinomios Airy tuvieron un error ligeramente menor de 0.39128, seguidos de las series de Fourier con un error de 0.39122.

En el caso del 4 % de los coeficientes, los polinomios Bessel y Zernike exhibieron un error de 0.37302. Los polinomios Airy tuvieron un error ligeramente inferior de 0.37270, mientras que las series de Fourier presentaron un error de 0.37233. Se observa una ligera disminución en el error en comparación con el 3 % de los coeficientes.

Con el 5 % de los coeficientes, los polinomios Bessel y Zernike obtuvieron un error de 0.36345, seguidos de las series de Fourier con un error de 0.36276. Los polinomios Airy tuvieron el error más bajo de 0.36258.

Finalmente, en el caso del 10 % de los coeficientes, los polinomios Bessel y Zernike mostraron un error igual de 0.31906. Los polinomios Airy tuvieron un error ligeramente más bajo de 0.31905, mientras que las series de Fourier presentaron el error más pequeño de 0.31838. Es importante destacar que este error es significativamente más bajo en comparación con el mencionado anteriormente para el 3 % de los coeficientes.

Todos los datos previamente mencionados se encuentran resumidos de manera completa en la Tabla 6.

Tabla 6. Tabla de valores del ECMN de las 5 palabras.

ECMN de 5 audios				
Porcentaje	Airy	Bessel	Fourier	Zernike
3 %	0.39128	0.39243	0.39122	0.39243
4 %	0.37270	0.37302	0.37233	0.37302
5 %	0.36258	0.36345	0.36276	0.36345
10 %	0.31905	0.31906	0.31838	0.31906

6.4. Pruebas subjetivas

La prueba subjetiva se llevó a cabo con la participación de un grupo de 20 individuos, conformado por 5 mujeres y 15 hombres, cuyas edades oscilaron entre los 22 y 39 años con una desviación estandar de 5.254. Las marcas de los audífonos utilizados fueron: Apple, Samsung, Razer, Beats, JBL entre otras.

Para evaluar la calidad del audio, se implementó la escala de Likert, una herramienta común para medir la percepción de variables cualitativas. Ampliamente utilizada en estudios de percepción, esta escala proporciona niveles de medición con un punto neutral para aquellos participantes que no expresan ni acuerdo ni desacuerdo (Canto de Gante et al., 2020).

Las respuestas ofrecidas a los participantes se categorizan en diferentes niveles de la escala de medición, con un elemento neutral para aquellos que no tienen una posición definida. En este trabajo, se emplearon 5 niveles en la escala de medición para calificar los audios, asignando la puntuación más baja a "pésimo" cuando el participante considera que la calidad del audio es deficiente y la más alta a "excelente" cuando el audio tiene una calidad sobresaliente (Carter et al., 1999). Dichas puntuaciones se detallan en la tabla 7.

Tabla 7. Tabla de puntuación.

Calificación	Puntuación
Pésimo	20
Malo	40
Regular	60
Bien	80
Excelente	100

La figura 23 también se divide en 4 secciones que son los 4 polinomios implementados que son: Airy, Bessel, Fourier y Zernike. Cada color representa el porcentaje de coeficientes implementados; 3% es de color amarillo, 4% es de color morado, 5% es color verde y el 10% es color azul claro.

Los resultados detallados en la Tabla 8 revelan algunas tendencias interesantes en función del porcentaje de coeficientes utilizados en la reconstrucción:

Para el 3% de los coeficientes, los polinomios Zernike obtienen el menor promedio con 41.6,

seguidos por los polinomios Airy y Bessel con una puntuación de 44.8, mientras que las Series de Fourier tienen la puntuación más alta de 47.2.

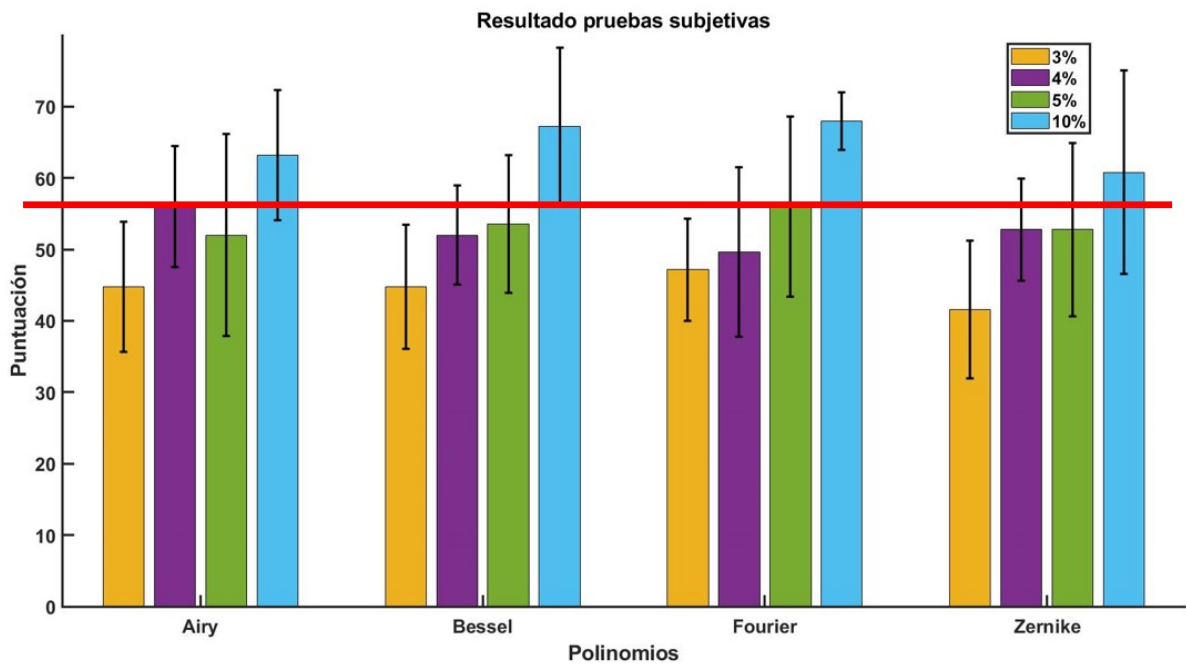


Figura 23. Resultados de la puntuación otorgada por los participantes a cada reconstrucción de la señal de audio, en función del porcentaje de coeficientes utilizados. La línea roja destaca que los polinomios Airy obtuvieron una puntuación similar a las Series de Fourier, pero con un menor número de coeficientes.

En el caso del 4% de los coeficientes, las Series de Fourier muestran la puntuación más baja de 49.6, seguidas por los polinomios de Bessel con 52, los polinomios de Zernike con 52.8 y los polinomios Airy obtienen la mejor puntuación con 56.

Tabla 8. Promedio de las pruebas subjetiva.

Promedio pruebas subjetivas				
Porcentaje	Airy	Bessel	Fourier	Zernike
3%	44.8	44.8	47.2	41.6
4%	56	52	49.6	52.8
5%	52	53.6	56.0	52.8
10%	63.2	67.2	68.0	60.8

Para el 5% de los coeficientes, se observa un aumento en las puntuaciones, aunque no significativo. Los polinomios Airy tienen una puntuación de 52, seguidos por los polinomios de Zernike con 52.8, los polinomios de Bessel con 53.6, y las Series de Fourier obtienen una puntuación de 56.

Finalmente, al considerar el 10 % de los coeficientes, se observa un aumento significativo en las puntuaciones en comparación con los porcentajes anteriores. Los polinomios de Zernike tienen la puntuación más alta con 60.8, seguidos por los polinomios Airy con 63.2, los polinomios de Bessel con 67.2 y las Series de Fourier obtienen la mejor puntuación con 68. Estos resultados proporcionan una visión detallada de cómo varía el rendimiento de los polinomios en función del porcentaje de coeficientes utilizados.

Finalmente, en la Tabla 9 se presenta la desviación estándar correspondiente a los valores resumidos en la Tabla 8. Esta desviación estándar proporciona información sobre la variabilidad de las puntuaciones obtenidas para cada conjunto de coeficientes y polinomios, complementando así la evaluación de su rendimiento.

Es posible que con un mayor número de participantes, las desviaciones estándar se reduzcan.

Tabla 9. Desviación estándar pruebas subjetivas.

Desviación estándar pruebas subjetivas				
Porcentaje	Airy	Bessel	Fourier	Zernike
3 %	9.1	8.7	7.2	9.6
4 %	8.5	6.9	11.9	7.2
5 %	14.1	9.6	12.6	12.1
10 %	9.1	11.1	4.0	14.3

6.5. Discusión

La eficacia de las aproximaciones polinomiales en la representación de señales acústicas destaca su capacidad para capturar patrones de manera compacta, resultando en una mejora en la eficiencia al reducir la cantidad de datos necesarios para representar la información. Esta reducción en la dimensionalidad no solo optimiza el proceso de compresión, sino que también simplifica la manipulación y transmisión de datos.

En los objetivos de esta investigación, se propuso la utilización de polinomios ortogonales, como los polinomios Airy, Bessel y Zernike, para la reconstrucción de señales acústicas. El valor significativo de los resultados obtenidos con los polinomios Airy, en comparación con otros polinomios, contribuye notablemente al avance en este ámbito, especialmente considerando que hasta la fecha

de nuestro conocimiento, estos polinomios no han sido ampliamente explorados en este contexto.

Aunque la literatura existente proporciona una base sólida, la presente investigación se enfoca en mejorar la eficiencia de los algoritmos de compresión y abordar desafíos específicos relacionados con la compresión de señales acústicas. La capacidad para mejorar la compresión sin pérdida continúa siendo áreas cruciales de interés y exploración.

La introducción de pruebas subjetivas en este proyecto aporta una dimensión crucial a la evaluación de las reconstrucciones. La combinación de pruebas subjetivas con las pruebas objetivas tradicionales representa un enfoque innovador, ya que no hay antecedentes claros sobre el uso conjunto de ambas en una sola investigación.

Tradicionalmente, la evaluación se ha centrado en métricas numéricas, como el error cuadrático medio normalizado, que capturan aspectos específicos de la señal pero pueden pasar por alto matices perceptuales. La inclusión de la opinión de los participantes permitió capturar elementos subjetivos esenciales para comprender la calidad percibida de las reconstrucciones polinomiales.

En términos de resultados, la combinación de pruebas objetivas y subjetivas reveló información valiosa. Aunque inicialmente se anticipaba que un aumento en el número de coeficientes polinomiales reduciría el error cuadrático medio normalizado, las opiniones de los participantes ofrecieron una perspectiva crítica adicional sobre la calidad percibida. Este hallazgo sugiere que la evaluación subjetiva desempeña un papel crucial al ajustar los modelos polinomiales para lograr una reproducción más fiel y satisfactoria de las señales acústicas.

En conclusión, la inclusión de pruebas subjetivas en el estudio de aproximación de señales acústicas a través de funciones polinomiales enriquece la evaluación al considerar la respuesta humana, contribuyendo así a una comprensión más completa y aplicable de la eficacia de esta aproximación en situaciones del mundo real.

Capítulo 7. Conclusiones

7.1. Conclusiones

La principal contribución de este trabajo de tesis reside en la evaluación, tanto objetiva como subjetiva, de la aproximación de señales de voz mediante el uso de polinomios de Airy, Bessel, Zernike y Series de Fourier. Al evaluar objetivamente cada uno de estos polinomios con diferentes porcentajes de coeficientes implementados, se observa una disminución en el error. Las Series de Fourier muestran el menor error, seguidas por los polinomios Airy, los polinomios de Bessel y, en última instancia, los polinomios de Zernike.

La mayoría de los participantes calificaron los audios como regulares, con una puntuación promedio de 64.8. Entre los polinomios, las Series de Fourier obtuvieron los mejores resultados, seguidas de los polinomios de Bessel, los polinomios Airy y en última instancia, los polinomios de Zernike, que obtuvieron una puntuación de 60.8.

El reducido número de participantes en las pruebas se debió a la dificultad de encontrar voluntarios dispuestos a invertir tiempo en este tipo de evaluaciones, ya que no se ofrecía ninguna compensación por su participación. Además, este desequilibrio de género en los participantes se debe a la falta de disponibilidad de individuos de ambos géneros dispuestos a realizar estas pruebas de manera voluntaria.

A pesar de ello, se tuvo un resultado favorecedor ya que, en las pruebas subjetivas, podemos observar que los audios reconstruidos con los polinomios Airy con un 4% de los coeficientes, tuvieron la misma puntuación que los audios reconstruidos con las Series de Fourier con el 5% de los coeficientes.

El uso de la plataforma web diseñada para este proyecto resultó de gran utilidad, ya que permitió a los participantes realizar la prueba de manera remota desde la comodidad de sus hogares, utilizando sus propios audífonos. Además, esta plataforma facilitó un control más efectivo de los datos, ya que se almacenaron directamente en un archivo tipo Excel, proporcionando una organización ordenada de los resultados obtenidos de los participantes.

También se presentaron desafíos técnicos relacionados con la plataforma web utilizada para llevar a cabo las pruebas auditivas. Debido al formato de los archivos de audio y al gran número de audios involucrados, algunos de ellos experimentaron retrasos en la reproducción. Además, la limitación de la plataforma para funcionar correctamente solo en computadoras y no en dispositivos móviles, generó algunas dificultades adicionales en la realización de las pruebas. Estos problemas técnicos pueden haber influido en la cantidad de participantes y en la calidad de los resultados obtenidos.

En resumen, objetivamente resulta difícil determinar cuál polinomio es superior, ya que los errores cuadráticos medios normalizados son bastante similares entre sí. Sin embargo, en la Figura 23 se observa un cambio significativo en los resultados de las pruebas subjetivas. Esto sugiere que el polinomio Airy es particularmente relevante, ya que las pruebas subjetivas indican que puede ser utilizado con solo el 4% de los coeficientes y aún así obtiene una puntuación positiva a comparación de los demás polinomios.

Estos resultados resaltan la relevancia de las pruebas subjetivas en este estudio, ya que indican que, con el número de participantes utilizado, los audios reconstruidos con polinomios Airy utilizando el 4% de los coeficientes obtuvieron resultados favorables.

7.2. Trabajo a futuro

Aunque este proyecto de investigación ha alcanzado sus objetivos, presenta un gran potencial para futuras investigaciones, tanto en el ámbito de las pruebas subjetivas como en las objetivas, con el fin de lograr una comprensión aún más profunda de las señales de voz.

En las pruebas objetivas, una dirección futura interesante podría ser aumentar el número de coeficientes utilizados en la reconstrucción de las señales de audio, por ejemplo, considerando porcentajes más altos, como del 11% al 15%. A pesar de que el error tiende a mantenerse constante cuando se utilizan solo dos cifras significativas después del punto decimal, esto permitiría disponer de un mayor número de audios para llevar a cabo pruebas subjetivas más exhaustivas. Además, se podría explorar la incorporación de un mayor número de polinomios ortogonales en el proceso de reconstrucción.

En el ámbito de las pruebas subjetivas, sería beneficioso aumentar la participación de oyentes en futuros estudios. Mantener una distribución equitativa de género entre los participantes podría ser útil para comprender posibles diferencias en la percepción de las señales de voz. Asimismo, se podría considerar agrupar a los participantes por edades, como crear grupos de oyentes de 20 años, 30 años y 40 años, con el propósito de investigar posibles relaciones entre la puntuación otorgada a los audios y la edad de los oyentes. Estos enfoques adicionales podrían arrojar información sobre cómo diferentes grupos demográficos perciben las señales de voz y proporcionar una comprensión más completa de los resultados.

Literatura citada

- Acebrás, A. (2020). Las funciones de Bessel y la función gamma. [Tesis de Licenciatura, Universidad de Santiago de Compostela]. Repositorio Institucional DA USC <https://minerva.usc.es/xmlui/bitstream/handle/10347/26007/Acebr%C3%A1s%20Bouza%2C%20Ana.pdf?sequence=1&isAllowed=y>,.
- Arias, A. H. (2007). Una variación del algoritmo lsm: discretización asociadas a sistemas haar. [Tesis de Licenciatura, Universidad Nacional de Mar del Plata]. Repositorio Digital de la FCEyS UNMDP <https://nulan.mdp.edu.ar/id/eprint/625/>.
- Barzola Tobar, F. D. & Cabrera Velasco, R. D. (2013). Comparación entre compresión de audio en diferentes formatos de imágenes equivalentes y el formato de compresión mp3. [Tesis de Licenciatura, Escuela Superior Politécnica de Litoral]. Repositorio de ESPOL <http://www.dspace.espol.edu.ec/handle/123456789/21613>,.
- Bria, O. N., Sosa, R. F., & Thomas, P. J. (Octubre de 1995). Compresión de audio en el espacio transformado. [Objeto de conferencia]. Congreso Argentino de Ciencias de la Computación. <http://sedici.unlp.edu.ar/handle/10915/24287>.
- Canalis, I. & Basso, G. (noviembre 2008). Evaluación de la calidad acústica del teatro Colón de Buenos Aires. [Objeto de conferencia]. VI Congreso Iberoamericano de Acústica-FIA 2008. <http://sedici.unlp.edu.ar/handle/10915/74408>.
- Canto de Gante, A. G., Sosa González, W. E., Bautista Ortega, J., Escobar Castillo, J., & Santillán Fernández, A. (2020). Escala de Likert: Una alternativa para elaborar e interpretar un instrumento de percepción social. *Revista de la Alta Tecnología y Sociedad*, 12(1), 38 – 45. <https://static1.squarespace.com/static/55564587e4b0d1d3fb1eda6b/t/5ffe0063b15beb25b917bec1/1610481763900/06+CantodeGante+ATS+V12N1+38-45.pdf>.
- Carter, S. J., Stewart, B. K., Kushmerick, M. J., Langer, S. G., Schmiedl, U. P., Winter, T. C., Conley, K. E., & Jubrias, S. A. (1999). Diagnostic ultrasound and telemedicine utilization in the international space station. *AIP Conference Proceedings*, 458(1), 409–414. <https://doi.org/10.1063/1.57603>.
- Castañeda, G. & Pérez, R. (1991). Análisis fonético de las listas de palabras de uso más extendido en logaudiometría. *Anales de la Sociedad Mexicana de Otorrinolaringología*, 36(1), 23–30. <https://biblat.unam.mx/en/revista/anales-de-otorrinolaringologia-mexicana/articulo/analisis-fonetico-de-las-listas-de-palabras-de-uso-mas-extendido-en-logoaudiometria>.
- Costas, R. (2007). Polinomios ortogonales respecto a medidas q-discretas. aplicaciones. [Tesis de Doctorado, Universidad Carlos III de Madrid]. Biblioteca Carlos III de Madrid <http://hdl.handle.net/10016/837>,.
- Cítores, J. (2015). Estudio de las propiedades de simetría de los polinomios ortogonales. [Tesis de Licenciatura, Universidad de Valladolid]. Repositorio Documental Universidad de Valladolid <http://uvadoc.uva.es/handle/10324/13486>,.
- Dusan, S., Flanagan, J. L., Karve, A., & Balaraman, M. (2007). Speech compression by polynomial approximation. *IEEE Transactions on Audio, Speech, and Language Processing*, 15(2), 387–395. <https://doi.org/10.1109/TASL.2006.881705>.

- Dutra, M. S., Niño, C. H. V., & García, R. S. (2009). Codificación y compresión de señales de voz con cuantización vectorial no determinística. *ITECKNE: Innovación e Investigación en Ingeniería*, 6(1), 14–19. <https://dialnet.unirioja.es/servlet/articulo?codigo=4991530>.
- Fisher, R. A. (1992). *Statistical Methods for Research Workers*. Springer New York. https://doi.org/10.1007/978-1-4612-4380-9_6.
- García, T. & Martínez, F. d. J. (2014). Diseño de practicas de la sonoridad en formatos de audio digital. [Tesis de Licenciatura, Instituto Politécnico Nacional]. Repositorio Dspace <https://tesis.ipn.mx/handle/123456789/14882>,.
- Gibson, J. D. (2016). Speech compression. *Information*, 7(2), 32. <https://doi.org/10.3390/info7020032>.
- Gómez-Correa, J. E., Platas-Garza, M. A., Trevino, J. P., Jaimes-Nájera, A., Padilla-Ortiz, A. L., Galaviz-Mosqueda, A., Coello, V., Rogel-Salazar, J., & Chávez-Cerda, S. (2020). Topographic synthesis of arbitrary surfaces with vortex jinc functions. *Appl. Opt.*, 59(13), 95–103. <https://doi.org/10.1364/AO.383157>.
- Heredia, J. (2014). Generalización de la teoría de señales y sistemas para su aplicación a la descripción, resolución y parametrización de problemas electromagnéticos. [Tesis de Doctorado, Universidad de Oviedo]. Repositorio Institucional de la Universidad de Oviedo <http://hdl.handle.net/10651/25579>,.
- Kaschel C, H., Watkins, F., & San Juan U, E. (2005). Compresión de voz mediante técnicas digitales para el procesamiento de señales y aplicaciones de formatos de compresión de imágenes. *Revista Facultad de Ingeniería - Universidad de Tarapacá*, 13(3), 4–10. <https://doi.org/10.4067/S0718-13372005000300002>.
- Keedwell, A. & Dénes, J. (2015). *Latin Squares and Their Applications: Latin Squares and Their Applications*. Elsevier Science. <https://books.google.com.mx/books?id=hsxLCgAAQBAJ>.
- McAulay, R. & Quatieri, T. (1984). Magnitude-only reconstruction using a sinusoidal speech model. *ICASSP '84. IEEE International Conference on Acoustics, Speech, and Signal Processing*, 9(1), 441–444. <https://doi.org/10.1109/ICASSP.1984.1172492>.
- Morey, L. (2019). Propiedades analíticas y estructurales de polinomios ortogonales matriciales. [Tesis de Licenciatura, Universidad Nacional De Córdoba]. Repositorio Digital UNC <http://hdl.handle.net/11086/12696>,.
- Moya, J. (2011). Procesamiento digital de señales. [Notas de clase de Ingeniería Electrónica, Instituto Tecnológico de Costa Rica], <https://palvarado.ietec.org/PDS/pds.pdf>,.
- Olvera, D. & Queb, M. (2017). Audiómetro de tonos puros. [Tesis de Licenciatura, Instituto Politécnico Nacional]. Repositorio Dspace <http://tesis.ipn.mx/handle/123456789/21963>,.
- Osorio, O., Ibarra, E. L., et al. (1987). *Métodos estadísticos aplicados a la investigación agrícola*. IICA. <https://repositorio.iica.int/handle/11324/11308>.

- Padilla-Ortiz, A. & Orduña-Bustamante, F. (2021). Binaural speech intelligibility tests conducted remotely over the internet compared with tests under controlled laboratory conditions. *Applied Acoustics*, 172, 107574. <https://doi.org/https://doi.org/10.1016/j.apacoust.2020.107574>.
- Proakis, J. G. & Manolakis, D. G. (1996). *Digital Signal Processing*, (3a ed.). Prentice-Hall, Inc. https://uvceee.files.wordpress.com/2016/09/digital_signal_processing_principles_algorithms_and_applications_third_edition.pdf.
- Prácticas, G. (2023). Sistemas de sonido. *Guías Prácticas* <https://www.guiaspracticas.com/altavoces/sistemas-de-sonido>.
- Romero, C. & Mares, D. (2013). Compresión de voz. [Tesis de Licenciatura, Instituto Tecnológico y de Estudios Superiores de Monterrey]. Repositorio del Instituto Tecnológico y de Estudios Superiores de Monterrey <http://hdl.handle.net/11285/629637>,.
- Sommerhoff, J. & Rosas, C. (2007). Evaluación de la inteligibilidad del habla en español. *Estudios filológicos*, (42), 215–225. <https://doi.org/10.4067/S0071-17132007000100014>.
- Strange, J. (2005). Voice authenticationa study of polynomial representation of speech signals. [Tesis de Maestría, University of Central Florida]. Repositorio Stars <https://stars.library.ucf.edu/etd/399>,.
- Tobón, L. & Posada, C. (2015). Descomposición de señales bioacústicas usando familias de polinomios ortogonales. *Revista de Investigaciones Universidad del Quindío.*, 27(1), 39–45. <https://doi.org/10.33975/riuv.vol127n1.22>.
- Varela-Nieto, I. & Rivera, T. (2010). *Presbiacusia*. Consejo Superior de Investigaciones Científicas (España). <https://digital.csic.es/handle/10261/77678>.
- Vazquez, R. L. (2017). Las pruebas objetivas y subjetivas. *Centro de Excelencia Académica*. <https://cea.uprrp.edu/las-pruebas-objetivas-y-subjetivas/>.
- Yanez, M. (2018). Delimitación de gliomas mediante la segmentación de imágenes de rm y pet basada en polinomios de zernike para la evaluación de factores influyentes en radioterapias. [Tesis de Maestría, Universidad Central de Venezuela.] Repositorio Institucional de la Universidad Central de Venezuela <http://hdl.handle.net/10872/17662>,.

Apéndices

Cálculo de los ceros de Bessel

```

function x=besselzero(n,k,kind)
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% n es el orden de la Bessel
% k son los ceros que aparecieron
% kind encuentra los primeros k ceros positivos de
% la fn de Bessel usando el metodo de Halley
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
k3=3*k;
x=zeros(k3,1);

for j=1:k3

    % Suposicion inicial de ceros
    x0=1+sqrt(2)+(j-1)*pi+n+n^0.4;
    % Metodo de Halley
    x(j)=findzero(n,x0,kind);
    if x(j)==inf
        error('Mala suposicion. ');
    end
end

x=sort(x);
dx=[1;abs(diff(x))];
x=x(dx>1e-8);

x=x(1:k);

function x=findzero(n,x0,kind)
n1=n+1;      n2=n*n;

```

```

% Tolerancia
tol=1e-12;

% Numero maximo de veces para iterar
MAXIT=100;
% Error inicial
err=1;
iter=0;

while abs(err)>tol && iter<MAXIT

    switch kind
        case 1
            a=besselj(n,x0);
            b=besselj(n1,x0);
        case 2
            a=bessely(n,x0);
            b=bessely(n1,x0);
    end
    x02=x0*x0;
    err=2*a*x0*(n*a-b*x0)/(2*b*b*x02-a*b*x0*(4*n+1)+(n*n1+x02)*
        a*a);
    x=x0-err;
    x0=x;
    iter=iter+1;
end
if iter>MAXIT-1
    warning('No se pudo converger dentro de la tolerancia. ',...
        'Pruebe con una suposicion inicial diferente');
    x=inf;
end

```

Polinomios de Bessel

```

%funcion para obtener los polinomios de Bessel
function Bess=besselpolinomios(r,ren,num_mod)
m=0;
%obtener los ceros de bessel
bn=besselzero(m,num_mod,1);
Bess=zeros(ren,num_mod);

%un ciclo for desde 1 hasta el numero de coeficientes definidos
for k=1:num_mod
    %se calcula rho
    rho=bn(k)*r;
    %obtenemos los polinomios de Bessel
    Bess(:,k)=besselj(m+1,rho);
end
end

```

Reconstrucción de la señal de voz por medio de los polinomios de Bessel

```

clear all; close all ;clc;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%Se declaran las variables del tiempo
contador=1;
%crea la tabla para almacenar el tiempo
Time=table('size',[20 2],'VariableTypes',['double','double'],
    ...
    'VariableNames',['audio' 'tiempo']);
%los audios estan almacenados por medio de numeros entonces se

```

```

coloca
%el numero de audio que se desea reconstruir
for numaudio =[8 23 38 44 57 63 70 84 95 107 111 135 138 141
158 162 165 179 186 199]
%empieza el contador del tiempo
tic;
straudio = num2str(numaudio);
%leer el audio
[yy,Fs] = audioread(strcat(straudio, '.wav'));
sonido1=transpose(yy(:,1));
sonido2=transpose(yy(:,2));

%Tamano del audio
Tamano=size(sonido1);
ren=Tamano(1,2);
r=linspace(0,1,ren);

%canal 1
R1=sonido1;
%canal 2
R2=sonido2;

%con la funcion dividir, se divide en 4 secciones el audio
[n11,n12,n13,n14] = dividir(R1,ren);
[n21,n22,n23,n24] = dividir(R2,ren);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%numero de coeficientes
num_mod=[1 2 3 4 5 10 15];
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%se almacena todo en una carpeta con el mismo numero de audio

```

```

output_folder = straudio;
if ~exist(output_folder, 'dir')
    mkdir(output_folder)
end

%crea un ciclo para cada uno de los coeficientes
for ii = 1:length(num_mod)

    %calcula el numero de coeficientes
    porcen1 = porcentaje(R1,ren,num_mod(ii));
    porcen2(ii) = floor((ren*num_mod(ii))/100);

    %se hace la aproximacion de cada una de las secciones y de
    cada canal
    %canal 1
    Tam11 = size(n11);
    ren11 = Tam11(1,2);
    r11 = linspace(0,1,ren11);
    Bess11 = besselpolinomios(r11,ren11,porcen1);
    AB11 = Bess11\n11';
    ABesesel11 = Bess11*AB11;

    Tam12 = size(n12);
    ren12 = Tam12(1,2);
    r12 = linspace(0,1,ren12);
    Bess12 = besselpolinomios(r12,ren12,porcen1);
    AB12 = Bess12\n12';
    ABesesel12 = Bess12*AB12;

    Tam13 = size(n13);

```

```
ren13 = Tam13(1,2);
r13 = linspace(0,1,ren13);
Bess13 = besselpolinomios(r13,ren13,porcen1);
AB13 = Bess13\n13';
ABesesel13 = Bess13*AB13;

Tam14 = size(n14);
ren14 = Tam14(1,2);
r14 = linspace(0,1,ren14);
Bess14 = besselpolinomios(r14,ren14,porcen1);
AB14 = Bess14\n14';
ABesesel14 = Bess14*AB14;

%Se crea el primer canal reconstruido
ARbrBess1= [ABesesel11;ABesesel12;ABesesel13;ABesesel14];

%Canal 2
Tam21 = size(n21);
ren21 = Tam21(1,2);
r21 = linspace(0,1,ren21);
Bess21 = besselpolinomios(r21,ren21,porcen1);
AB21 = Bess21\n21';
ABesesel21 = Bess21*AB21;

Tam22 = size(n22);
ren22 = Tam22(1,2);
r22 = linspace(0,1,ren22);
Bess22 = besselpolinomios(r22,ren22,porcen1);
AB22 = Bess22\n22';
ABesesel22 = Bess22*AB22;
```

```

Tam23 = size(n23);
ren23 = Tam23(1,2);
r23 = linspace(0,1,ren23);
Bess23 = besselpolinomios(r23,ren23,porcen1);
AB23 = Bess23\n23';
ABesesel23 = Bess23*AB23;

Tam24 = size(n24);
ren24 = Tam24(1,2);
r24 = linspace(0,1,ren24);
Bess24 = besselpolinomios(r24,ren24,porcen1);
AB24 = Bess24\n24';
ABesesel24 = Bess24*AB24;

%Se crea el segundo canal reconstruido
ARbrBess2= [ABesesel21;ABesesel22;ABesesel23;ABesesel24];
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%guardar el audio en .wav
reconstruida=[ARbrBess1';ARbrBess2']';
audioname = fullfile(output_folder,...
    ['Bessel_',num2str(numaudio),'_coe_',num2str(num_mod(ii
    )),'.wav']);
audiowrite(audioname , reconstruida , Fs);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%Crea una tabla donde se almacenaran el ECMN de cada canal y
de cada
%seccion reconstruida
Tn = table('Size',[length(num_mod) 4],...
    'VariableTypes',['double', "double","double","double

```

```

        ],...
        'VariableNames',["Porcentaje","NCoeficientes","Canal1
        ","Canal2"]);

    %Calcula el ECMN de cada canal y de cada seccion de cada
    canal
    NRMSE1(ii)=sqrt(sum((abs(R1'-ARbrBess1).^2)/sum((abs(R1')
        .^2))));
    NRMSE2(ii)=sqrt(sum((abs(R2'-ARbrBess2).^2)/sum((abs(R2')
        .^2))));

    %calcula el error absoluto y lo almacena en un .mat
    Abs=abs(R1'-ARbrBess1);
    narch=strcat('Abs_Bessel_',num2str(numaudio),'_coe_',
        num2str(num_mod(ii)));
    absname = fullfile(output_folder,narch);
    save(absname,'Abs');

end

%guardar los errores en un archivo excel y lo almacena en
    la carpeta
    Tn(:, :) = num2cell([num_mod',porcen2',NRMSE1',NRMSE2']);
    errorname = fullfile(output_folder,...
        strcat('NRMSE_Bessel',num2str(numaudio),'.csv'));
    writetable(Tn,errorname);

    %Toma el tiempo de cada reconstruccion
    tiempo=toc;
    Time.audio(contador)=numaudio;

```



```

    Time.tiempo(contador)=tiempo;
    contador=contador+1;

disp('terminado')
end
    %almacena el tiempo de cada uno de los audios en un excel
tiempototal =('Tiempo_Bessel.csv');
writetable(Time,tiempototal);

```

Polinomios Airy

```

%funcion para obtener los polinomios Airy
function Airy=airypolinomios(r,ren,num_mod)
m=0;
%vector de ceros de bessel
bn=besselzero(m,num_mod,1);
%vector de ceros
Airy=zeros(ren,num_mod);
%eps cero numerico el cual su valor es 2.2204e-16
r=r+eps;
%ciclo for de 1 hasta el numero de coeficientes
for k=1:num_mod
    %calculamos rho
    rho=bn(k)*r;
    %se obtienen los polinomios Airy
    Airy(:,k)=besselj(m+1,rho)./rho;

end
r=linspace(0,1,ren);
end

```

Reconstrucción de la señal de voz por medio de los polinomios Airy

```

clear all; close all ;clc;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%Se declaran las variables del tiempo
tiempo=0;
tiempopre=0;
contador=1;
%crea la tabla para almacenar el tiempo
Time=table('size',[20 2],'VariableTypes',['double',"double"],
...
'VariableNames',['audio' "tiempo"]);

%los audios estan almacenados por medio de numeros entonces se
coloca
%el numero de audio que se desea reconstruir
for numaudio =[8 23 38 44 57 63 70 84 95 107 111 135 138 141
158 162 165 179 186 199]
%empieza el contador del tiempo
tic;
straudio = num2str(numaudio);
%leer el audio
[yy,Fs] = audioread(strcat(straudio, '.wav'));
%se dividen los dos canales de audio
sonido1=transpose(yy(:,1));
sonido2=transpose(yy(:,2));

%calcula el tamaño del audio
Tamano=size(sonido1);
ren=Tamano(1,2);
r=linspace(0,1,ren);

```

```

%canal 1
R1=sonido1;

%canal 2
R2=sonido2;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%numero de coeficientes en porcentaje
num_mod=[1 2 3 4 5 10 15];
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%con la funcion dividir, se divide en 4 secciones el audio
[n11,n12,n13,n14] = dividir(R1,ren);
[n21,n22,n23,n24] = dividir(R2,ren);

%se almacena todo en una carpeta con el mismo numero de audio
output_folder = straudio;
if ~exist(output_folder, 'dir')
    mkdir(output_folder)
end

%crea un ciclo para cada uno de los coeficientes
for ii = 1:length(num_mod)

    %calcula el numero de coeficientes
    porcen2(ii) = floor((ren*num_mod(ii))/100);
    porcen1 = porcentaje(R1,ren,num_mod(ii));

    %se hace la aproximacion de cada una de las secciones y de
    cada canal
    Tampe11 = size(n11);
    renpe11 = Tampe11(1,2);
    rpe11 = linspace(0,1,renpe11);

```

```
Airy11 = airypolinomios(rpe11,renpe11,porcen1);
Abu11 = Airy11\n11';
ARbr11 = Airy11*Abu11;

Tampe12 = size(n12);
renpe12 = Tampe12(1,2);
rpe12 = linspace(0,1,renpe12);
Airy12 = airypolinomios(rpe12,renpe12,porcen1);
Abu12 = Airy12\n12';
ARbr12 = Airy12*Abu12;

Tampe13 = size(n13);
renpe13 = Tampe13(1,2);
rpe13 = linspace(0,1,renpe13);
Airy13 = airypolinomios(rpe13,renpe13,porcen1);
Abu13 = Airy13\n13';
ARbr13 = Airy13*Abu13;

Tampe14 = size(n14);
renpe14 = Tampe14(1,2);
rpe14 = linspace(0,1,renpe14);
Airy14 = airypolinomios(rpe14,renpe14,porcen1);
Abu14 = Airy14\n14';
ARbr14 = Airy14*Abu14;

%Se crea el primer canal reconstruido
ARbr1 = [ARbr11;ARbr12;ARbr13;ARbr14];

%canal 2
Tampe21 = size(n21);
renpe21 = Tampe21(1,2);
```



```

%%guardar el audio en .wav y lo almacena en la carpeta
    correspondiente
reconstruida=[ARbr1';ARbr2'];
audioname = fullfile(output_folder,...
['Airy_',num2str(numaudio),'_coe_',num2str(num_mod(ii)),'.
    wav']);
audiowrite(audioname, reconstruida, Fs);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%Crea una tabla donde se almacenaran el ECMN de cada canal
    y de cada
%seccion reconstruida
T = table('Size',[length(num_mod) 4],...
    'VariableTypes',"double", "double","double","double",
    ...
    'VariableNames',"Porcentaje","NCoeficientes","Canal1
    ","Canal2"]);

%Calcula el ECMN de cada canal y de cada seccion de cada
    canal
NRMSE1(ii)=sqrt(sum((abs(R1'-ARbr1).^2)/sum((abs(R1').^2)))
    );
NRMSE2(ii)=sqrt(sum((abs(R2'-ARbr2).^2)/sum((abs(R2').^2)))
    );

%calcula el error absoluto y lo almacena en un .mat
Abs=abs(R1'-ARbr1);
narch=strcat('Abs_Airy_',num2str(numaudio),'_coe_',num2str(
    num_mod(ii)));

```

```

absname = fullfile(output_folder,narch);
save(absname,'Abs');

end

%guardar los errores en un archivo excel y lo almacena en
    la carpeta
    T(:, :) = num2cell([num_mod',porcen2',NRMSE1',NRMSE2']);
errorname = fullfile(output_folder,...
    strcat('NRMSE_Airy',num2str(numaudio),'.csv'));
writetable(T,errorname);

%Toma el tiempo de cada reconstruccion
tiempo=toc;
Time.audio(contador)=numaudio;
Time.tiempo(contador)=tiempo-tiempopre;
contador=contador+1;
tiempopre=tiempo;
disp('terminado')

end

%almacena el tiempo de cada uno de los audios en un excel
tiempototal = fullfile(output_folder,'Tiempo_Airy.csv');
writetable(Time,tiempototal);

```

Series de Fourier

```

function [serie] = FourierApproximation(y,n,v1,v2,x,k)
%intervalo de la funcion
L=v2-v1;
%delta equis

```

```

dx=x(2)-x(1);
%es un producto que aparece constantemente en las ecuaciones
%y no cambia y por lo tanto es una constante
w=pi/L;

%arreglo de los puntos que quiero conocer su valor
g=linspace(v1,v2,n);

sum=0;
a0=0;
    %empieza desde 1 hasta la cantidad de terminos que quiero
    encontrar en
    %la serie

for N=1:k
    a(N)=0;
    b(N)=0;
    a0=0;
    %desde 1 hasta los puntos que tenemos en la funcion que se
    desea aproximar, se calculan los coheficientes Ao, Ak y
    Bk como lo marca la serie de fourier

    for t=1:n
        a0=a0+dx*(1/L)*y(t);
        a(N)=a(N)+dx*(1/L)*y(t)*cos(N*w*x(t));
        b(N)=b(N)+dx*(1/L)*y(t)*sin(N*w*x(t));
    end
end

%se realiza la serie de fourier con los coheficientes
calculados

```



```

%anteriormente
for N=1:k
    sum=sum+(a(N).*cos(N*w*g)+b(N).*sin(N*w*g));
end

%valor final
serie=a0/2+sum;

end

```

Reconstrucción de la señal de voz por medio de las Series de Fourier

```

clear all; close all; clc
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%Se declaran las variables del tiempo
contador=1;
%crea la tabla para almacenar el tiempo
Time=table('size',[20 2],'VariableTypes',['double', 'double'],
    ...
    'VariableNames',['audio' 'tiempo']);

%los audios estan almacenados por medio de numeros entonces se
coloca
%el numero de audio que se desea reconstruir
for numaudio = [8 23 38 44 57 63 70 84 95 107 111 135 138 141
    158 162 165 179 186 199]
%empieza el contador del tiempo
tic;
straudio = num2str(numaudio);
%leer el audio
[yy,Fs] = audioread(strcat(straudio, '.wav'));

```

```
%Divide los audios en dos canales
sonido1=yy(:,1);
sonido2=yy(:,2);
y1 = sonido1';
y2 = sonido2';

%-----Fourier-----
%se almacenara todo en una carpeta con el mismo numero de audio
output_folder = straudio;
if ~exist(output_folder, 'dir')
    mkdir(output_folder)
end

%calcula el tamaño del audio
Tamano2=size(y1);
ren=Tamano2(1,2);

%con la función dividir, se divide en 4 secciones el audio
[n11,n12,n13,n14] = dividir(y1,ren);
[n21,n22,n23,n24] = dividir(y2,ren);

%Obtiene el tamaño de cada una de las secciones del canal 1
Tam11 = size(n11);
nn11 = Tam11(1,2);
Tam12 = size(n12);
nn12 = Tam12(1,2);
Tam13 = size(n13);
nn13 = Tam13(1,2);
Tam14 = size(n14);
nn14 = Tam14(1,2);
```

```

%obtiene el tamaño de cada una de las secciones del canal 1
Tam21 = size(n21);
nn21 = Tam21(1,2);
Tam22 = size(n22);
nn22 = Tam22(1,2);
Tam23 = size(n23);
nn23 = Tam23(1,2);
Tam24 = size(n24);
nn24 = Tam24(1,2);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%numero de coeficientes
num_mod = [1 2 3 4 5 10 15];
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

v1=0;%inicio del intervalo
v2=1;%fin del intervalo
%se crea un vector para cada seccion de cada canal
x11 = linspace(v1,v2,nn11);
x12 = linspace(v1,v2,nn12);
x13 = linspace(v1,v2,nn13);
x14 = linspace(v1,v2,nn14);

x21 = linspace(v1,v2,nn21);
x22 = linspace(v1,v2,nn22);
x23 = linspace(v1,v2,nn23);
x24 = linspace(v1,v2,nn24);

%crea un ciclo para cada uno de los coeficientes
for ii = 1:length(num_mod)

```

```

%calcula el numero de coeficientes
porcen1 = porcentaje(y1,ren,num_mod(ii));
porcen2(ii) = floor((ren*num_mod(ii))/100);

%se hace la aproximacion de cada una de las secciones y de
    cada canal
F11 = FourierApproximation(n11,nn11,v1,v2,x11,porcen1);
F12 = FourierApproximation(n12,nn12,v1,v2,x12,porcen1);
F13 = FourierApproximation(n13,nn13,v1,v2,x13,porcen1);
F14 = FourierApproximation(n14,nn14,v1,v2,x14,porcen1);

F21 = FourierApproximation(n21,nn21,v1,v2,x21,porcen1);
F22 = FourierApproximation(n22,nn22,v1,v2,x22,porcen1);
F23 = FourierApproximation(n23,nn23,v1,v2,x23,porcen1);
F24 = FourierApproximation(n24,nn24,v1,v2,x24,porcen1);
%-----guardar el audio-----

%En un vector se juntan las 4 partes de cada canal
F1 =[F11';F12';F13';F14'];
F2 =[F21';F22';F23';F24'];

%Juntamos los dos vectores
reconstruida=[F1,F2];
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%guardar el audio en .wav en la carpeta correspondiente con
    el
%numero de coeficiente

    audioname = fullfile(output_folder,...
['Fourier_',num2str(numaudio),'_coe_',num2str(num_mod(ii))],

```

```

        '.wav']));
    audiowrite(audioname, reconstruida, Fs)
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% errores

%Crea una tabla para guardar el ECMN de cada porcentaje, con
    el numero de
%coeficientes, de cada canal y de cada seccion
    Tn = table('Size',[length(num_mod) 4],...
        'VariableTypes',["double", "double","double","double"],
        ...
        'VariableNames',["Porcentaje","NCoeficientes","Canal1
            ","Canal2"]);

%calcula el ECMN para cada canal reconstruido y para cada
    una de las
%4 secciones de de cada canal

NRMSE1(ii)=sqrt(sum((abs(sonido1-F1).^2)/sum((abs(sonido1)
    .^2))));
NRMSE2(ii)=sqrt(sum((abs(sonido2-F2).^2)/sum((abs(sonido2)
    .^2))));

%Obtiene el valor absoluto de un canal y lo almacena en un
    archivo .mat
Abs=abs(sonido1-F1);
narch=strcat('Abs_Fourier_',num2str(numaudio),'_coe_',
    num2str(num_mod(ii)));
absname = fullfile(output_folder,narch);
save(absname,'Abs');

```

```

end

%se almacenan los ECMN en un documento de excel que estara
    en la misma
%carpeta del audio
Tn(:, :) = num2cell([num_mod', porcen2', NRMSE1', NRMSE2']);
errorname = fullfile(output_folder, ...
    strcat('NRMSE_Furier', num2str(numaudio), '.csv'));
writetable(Tn, errorname);

    %Toma el tiempo de cada reconstruccion
tiempo=toc;
Time.audio(contador)=numaudio;
Time.tiempo(contador)=tiempo;
contador=contador+1;

disp('terminado')

end

tiempototal = fullfile('Tiempo_Fourier.csv');
writetable(Time, tiempototal);

```

Polinomios de Zernike

```

function Rmn=zernikepolinomios(ren, num_mod);

r=linspace(0,1,ren)';
u=r.^2;
Rmn=zeros(ren, num_mod);

```

```

m=0;

Rmn(:,1)=u.^0;
Rmn(:,2)=-(m+1)+(m+2)*u;
for n=1:num_mod-1
    s=m+2*n;
    a=-(s+1)*((s-n)^2+n^2+s)/(n+1)/(s-n+1)/s;
    b=(s+2)*(s+1)/(n+1)/(s-n+1);
    c=(s+2)*(s-n)*n/(n+1)/(s-n+1)/s;

Rmn(:,n+2)=(a+b*u).*Rmn(:,n+1)-c*Rmn(:,n);

end

rr=r.^m;
rr= repmat(rr,1,num_mod+1);

%polinomios de zernike
Rmn=rr.*Rmn;

```

Reconstrucción de la señal de voz por medio de los polinomios de Zernike

```

clear all; close all ;clc;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%Se declaran las variables del tiempo
contador=1;
%crea la tabla para almacenar el tiempo
Time=table('size',[20 2],'VariableTypes',['double','double'],
    ...

```



```

%numero de coeficientes
num_mod=[1 2 3 4 5 10 15];
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%se almacenara todo en una carpeta con el mismo numero de audio
output_folder = straudio;
if ~exist(output_folder, 'dir')
    mkdir(output_folder)
end

%crea un ciclo para cada uno de los coeficientes
for ii = 1:length(num_mod)

    %calcula el numero de coeficientes
    porcen1 = porcentaje(R1,ren,num_mod(ii));
    porcen2(ii) = floor((ren*num_mod(ii))/100);

    %se hace la aproximacion de cada una de las secciones y de
    cada canal
    %canal 1

    Tam11 = size(n11);
    ren11 = Tam11(1,2);
    r11 = linspace(0,1,ren11);
    Rmn11 = besselpolinomios(r11,ren11,porcen1);
    mu11 = Rmn11\n11';
    Rzr11 = Rmn11*mu11;

    Tam12 = size(n12);
    ren12 = Tam12(1,2);
    r12 = linspace(0,1,ren12);

```

```
Rmn12 = besselpolinomios(r12,ren12,porcen1);
mu12 = Rmn12\n12';
Rzr12 = Rmn12*mu12;

Tam13 = size(n13);
ren13 = Tam13(1,2);
r13 = linspace(0,1,ren13);
Rmn13 = besselpolinomios(r13,ren13,porcen1);
mu13 = Rmn13\n13';
Rzr13 = Rmn13*mu13;

Tam14 = size(n14);
ren14 = Tam14(1,2);
r14 = linspace(0,1,ren14);
Rmn14 = besselpolinomios(r14,ren14,porcen1);
mu14 = Rmn14\n14';
Rzr14 = Rmn14*mu14;

%Se crea el primer canal reconstruido
Rzr1= [Rzr11;Rzr12;Rzr13;Rzr14];

%canal 2
Tam21 = size(n21);
ren21 = Tam21(1,2);
r21 = linspace(0,1,ren21);
Rmn21 = besselpolinomios(r21,ren21,porcen1);
mu21 = Rmn21\n21';
Rzr21 = Rmn21*mu21;

Tam22 = size(n22);
ren22 = Tam22(1,2);
```



```

%      %errores

%Crea una tabla para guardar el ECMN de cada porcentaje, con
    el numero de
%coeficientes, de cada canal y de cada seccion
    Tn = table('Size',[length(num_mod) 4],...
        'VariableTypes',["double", "double","double","double"],
        ...
        'VariableNames',["Porcentaje","NCoeficientes","Canal1
            ","Canal2"]);

%calcula el ECMN para cada canal reconstruido y para cada una
    de las
        %4 secciones de de cada canal
NRMSE1(ii)=sqrt(sum((abs(R1'-Rzr1).^2)/sum((abs(R1').^2))));
NRMSE2(ii)=sqrt(sum((abs(R2'-Rzr2).^2)/sum((abs(R2').^2))));

%Obtiene el valor absoluto de un canal y lo almacena en un
    archivo .mat
Abs=abs(R1'-Rzr1);
narch=strcat('Abs_Zernike_',num2str(numaudio),'_coe_',
    num2str(num_mod(ii)));
absname = fullfile(output_folder,narch);
save(absname,'Abs');

end

%se almacenan los ECMN en un documento de excel que estara
    en la misma
%carpeta del audio

```

```
Tn(:, :) = num2cell([num_mod', porcen2', NRMSE1', NRMSE2']);
errorname = fullfile(output_folder, ...
    strcat('NRMSE_Zernike_', num2str(numaudio), '.csv'));
writetable(Tn, errorname);

    %Toma el tiempo de cada reconstruccion
tiempo=toc;
Time.audio(contador)=numaudio;
Time.tiempo(contador)=tiempo;
contador=contador+1;

disp('terminado')
end

tiempototal = fullfile('Tiempo_Zernike.csv');
writetable(Time, tiempototal);
```