



INAOE

Sonificación de EEG para la clasificación de habla imaginada

Por:

Erick Fernando González Castañeda

Tesis sometida como requisito parcial
para obtener el grado de:

**MAESTRÍA EN CIENCIAS EN LA ESPECIALIDAD DE
CIENCIAS COMPUTACIONALES**

en el

Instituto Nacional de Astrofísica, Óptica y Electrónica

Febrero, 2015
Tonantzintla, Puebla

Dirigida por:

**Dr. Luis Villaseñor Pineda
Dr. Carlos A. Reyes García**

©INAOE 2015

Derechos reservados

El autor otorga al INAOE el permiso de
reproducir esta tesis en su totalidad o en partes



TABLA DE CONTENIDOS

LISTA DE FIGURAS	VII
LISTA DE TABLAS	XI
1. INTRODUCCIÓN	1
1.1. Problemática	2
1.2. Solución propuesta	3
1.2.1. Preguntas de Investigación	4
1.2.2. Hipótesis de investigación	4
1.2.3. Objetivo general	4
1.2.4. Objetivos específicos	5
1.3. Contribuciones	5
1.4. Alcances y limitaciones	5
1.5. Estructura del documento	6
2. MARCO TEÓRICO	7
2.1. El cerebro	7
2.1.1. Procesamiento del lenguaje	8
2.2. Electroencefalograma (EEG)	9
2.2.1. Artefactos EEG	11
2.3. Brain Computer Interfaces (BCI)	12
2.4. Sonificación	13
2.4.1. Sonificación de EEG	14
2.5. Transformada de Fourier	15
2.6. Transformada Wavelet	16
2.6.1. Energía Relativa Wavelet (RWE)	16
2.7. Coeficientes cepstrales en la frecuencia Mel (MFCC)	17
2.8. Clasificación automática	19

2.8.1. Random Forest	19
2.9. Prueba estadística de resultados	19
3. TRABAJOS RELEVANTES RELACIONADOS	21
3.1. Clasificación de habla imaginada	21
3.1.1. Clasificación por vocales, sílabas o fonemas	22
3.1.2. Clasificación de palabras	24
3.2. Sonificación de EEG para clasificación	26
3.3. Análisis de los trabajos relacionados	30
4. METODOLOGÍA	33
4.1. Descripción General	33
4.2. Dispositivo EEG utilizado	33
4.3. Adquisición de la señal cerebral	34
4.4. Montaje de referencia	35
4.5. Sonificación de la señal EEG	36
4.6. Extracción de características	39
4.6.1. Usando DWT y Energía Relativa	39
4.6.2. Usando MFCC	41
4.7. Clasificación	42
5. EXPERIMENTACIÓN Y EVALUACIÓN	45
5.1. Descripción de los datos	45
5.2. Configuración de parámetros del método de Sonificación de EEG	45
5.3. Selección de Wavelet para EEG Sonificado	49
5.4. Configuración de parámetros de MFCC para sonificación de EEG	50
5.5. Configuración de parámetros de MFCC para EEG	52
5.6. Resumen de configuraciones utilizadas	54
5.7. Experimento de clasificación para los 27 sujetos	55
5.7.1. Resultados de clasificación en los 4 enfoques	55
5.7.2. Prueba estadística	58
5.7.3. Análisis de clasificación por palabra	60
5.7.4. Energía relativa por palabra	63
6. CONCLUSIONES Y TRABAJO FUTURO	69
6.1. Conclusiones	69
6.2. Trabajo futuro	70

A. PUBLICACIONES	73
A.1. Presentación	73
A.2. Artículos	73
A.3. Capítulo en libro	73
B. DETALLES DE LOS EXPERIMENTOS	75
B.1. Matrices de confusión de todos sujetos por palabra	75
C. RESULTADOS ADICIONALES	85
BIBLIOGRAFÍA	87

LISTA DE FIGURAS

2.1. División cerebral por lóbulos	7
2.2. Modelo Geschwind-Wernicke	8
2.3. Electrodo de EEG registrando actividades sinápticas de las neuronas.	9
2.4. Dispositivos para hacer lectura de EEG.	10
2.5. Interfaz Cerebro Computadora 'Intendix'	13
2.6. Ejemplo de espectrograma de una señal de EEG	16
2.7. Ejemplo de grupo de filtros triangulares usado para calcular los MFCCs	18
3.1. Esquema de grabación de una repetición para la investigación realizada	22
3.2. Esquema de grabación de la pronunciación imaginada de la sílaba /ba/.	23
3.3. Comparación de exactitudes de trabajos relacionados	28
3.4. Espectrogramas generados por el algoritmo de sonificación <i>EEG a tonos</i>	29
4.1. Metodología seguida para la clasificación de palabras no pronunciadas usando sonificación de EEG.	34
4.2. Localización de los electrodos en el kit EPOC de Emotiv	34
4.3. Señal de EEG del canal F7 mientras imagina la dicción de la palabra <i>Abajo</i>	35
4.4. Ejemplo de sonificación para configuraciones distintas.	39
4.5. Distribución e intensidad de la energía de los tonos de las señales de audio de la figura 4.4	40
4.6. Espectrogramas del resultado de la sonificación de EEG para los ejemplos del la figura 4.4.	41
4.7. Descomposición en niveles de la transformada Wavelet.	41
5.1. Promedios de exactitud en la clasificación con distinto número de tonos.	46
5.2. Promedios de exactitud en la clasificación con distinta duración (en segundos) de los tonos.	46

5.3. Promedios de exactitud en la clasificación con distinta frecuencia de muestreo para el audio de salida.	47
5.4. Promedios de exactitud en la clasificación con distinto rango de frecuencias en los audios	47
5.5. Promedios de exactitud en la clasificación con distinto número de muestras por ventana para formar el espectrograma.	48
5.6. Promedios de exactitud en la clasificación con las dos mejores muestras por ventana para el espectrograma.	48
5.7. Comparativa para la elección de wavelets.	49
5.8. Resultados de exactitud promedio en la clasificación usando Daubechies de orden 2 y de orden 20, con distintos niveles de descomposición. . . .	50
5.9. Promedios de exactitud en la clasificación con distinto número de coeficientes cepstrales, usando MFCC para sonificación de EEG.	51
5.10. Promedios de exactitud en la clasificación con distinto tamaño de ventana, usando MFCC para sonificación de EEG.	51
5.11. Promedios de exactitud en la clasificación con distinto traslape entre ventanas, usando MFCC para sonificación de EEG.	51
5.12. Promedios de exactitud en la clasificación con distinto tamaño de ventanas en delta y doble delta, usando MFCC para sonificación de EEG. . .	52
5.13. Promedios de exactitud en la clasificación con distinto número de coeficientes cepstrales, usando MFCC para EEG sin sonificar.	52
5.14. Promedios de exactitud en la clasificación con distinto tamaño de ventana y traslape que en combinación dan el número de muestras por segundo, usando MFCC para EEG sin sonificar.	53
5.15. Promedios de exactitud en la clasificación con distinto número de filtros en las frecuencias Mel, usando MFCC para EEG sin sonificar.	53
5.16. Promedios de exactitud en la clasificación con distinto tamaño de ventanas en delta y doble delta, usando MFCC para EEG sin sonificar.	54
5.17. Exactitud de la clasificación para los 27 sujetos usando 4 canales	58
5.18. Exactitud de la clasificación para los 27 sujetos usando 14 canales	59
5.19. Promedios de exactitud de los 4 enfoques en 4 y 14 canales.	59
5.20. Cantidad de sujetos que obtienen sus mejores resultados para cada enfoque.	60
5.21. Resultados de la prueba de normalidad Shapiro-Wilk por pares.	60
5.22. Resultados de la prueba de significancia estadística por pares, usando comparación por pares.	60

5.23. Promedios de exactitud en la clasificación por palabra usando 4 canales.	61
5.24. Promedios de exactitud en la clasificación por palabra usando 14 canales.	61
5.25. Palabras ordenadas de acuerdo a la exactitud de clasificación promedio.	62
5.26. Mejores porcentajes de exactitud por método en las 5 palabras usando 4 canales.	63
5.27. Mejores porcentajes de exactitud por método en las 5 palabras usando 14 canales.	64
5.28. Porcentajes de exactitud más bajos por método en las 5 palabras usando 4 canales.	65
5.29. Porcentajes de exactitud más bajos por método en las 5 palabras usando 14 canales.	66
5.30. Energía relativa Wavelet promedio de los 27 sujeto usando 14 canales. Se muestra la distribución de la energía por niveles de descomposición para cada palabra para los métodos de EEG (arriba) y de sonificación de EEG (abajo).	67
B.1. Promedios de clasificación por palabra para el sujeto 1.	76
B.2. Promedios de clasificación por palabra para el sujeto 2.	76
B.3. Promedios de clasificación por palabra para el sujeto 3.	76
B.4. Promedios de clasificación por palabra para el sujeto 4.	77
B.5. Promedios de clasificación por palabra para el sujeto 5.	77
B.6. Promedios de clasificación por palabra para el sujeto 6.	77
B.7. Promedios de clasificación por palabra para el sujeto 7.	77
B.8. Promedios de clasificación por palabra para el sujeto 8.	78
B.9. Promedios de clasificación por palabra para el sujeto 9.	78
B.10. Promedios de clasificación por palabra para el sujeto 10.	78
B.11. Promedios de clasificación por palabra para el sujeto 11.	78
B.12. Promedios de clasificación por palabra para el sujeto 12.	79
B.13. Promedios de clasificación por palabra para el sujeto 13.	79
B.14. Promedios de clasificación por palabra para el sujeto 14.	79
B.15. Promedios de clasificación por palabra para el sujeto 15.	79
B.16. Promedios de clasificación por palabra para el sujeto 16.	80
B.17. Promedios de clasificación por palabra para el sujeto 17.	80
B.18. Promedios de clasificación por palabra para el sujeto 18.	81
B.19. Promedios de clasificación por palabra para el sujeto 19.	81
B.20. Promedios de clasificación por palabra para el sujeto 20.	81
B.21. Promedios de clasificación por palabra para el sujeto 21.	82

B.22. Promedios de clasificación por palabra para el sujeto 22.	82
B.23. Promedios de clasificación por palabra para el sujeto 23.	82
B.24. Promedios de clasificación por palabra para el sujeto 24.	82
B.25. Promedios de clasificación por palabra para el sujeto 25.	83
B.26. Promedios de clasificación por palabra para el sujeto 26.	83
B.27. Promedios de clasificación por palabra para el sujeto 27.	83
C.1. Promedios en la exactitud de clasificación de los 27 sujetos usando distintos clasificadores con 4 canales.	85
C.2. Promedios en la exactitud de clasificación de los 27 sujetos usando distintos clasificadores con 14 canales.	86

LISTA DE TABLAS

3.1. Conjunto de fonemas utilizados en el estudio.	24
3.2. Trabajos relacionados con la clasificación de habla imaginada.	27
3.3. Trabajos relacionados con la sonificación de EEG usada para clasificación.	30
5.1. Promedios de exactitud en la clasificación de los 4 enfoques usando 4 canales	56
5.2. Promedios de exactitud en la clasificación de los 4 enfoques usando 14 canales	57

AGRADECIMIENTOS

Agradezco al INAOE por el apoyo brindado para realizar este trabajo; al Consejo de Ciencia y Tecnología (CONACyT) por el apoyo económico para la investigación con la beca# 483056.

Agradezco a mis asesores el Dr. Luis Villaseñor Pineda y el Dr. Carlos Reyes García por sus conocimientos y apoyo para la realización de esta tesis; a Alejandro Torres García por su ayuda y consejos, a Alejandro Rosales Pérez por su ayuda y por crear la plantilla de estilo para la tesis. Agradezco a mis revisores de tesis, cuyos consejos ayudaron a mejorar la calidad del documento.

A la Sra. Martha Olmos Flores por brindarme su apoyo aún sin ser todavía alumno del instituto.

Agradezco a mi Padre Juan González y Lilia Castañeda, pues sin ellos yo no estaría aquí, sin su apoyo y educación hubiera sido muy difícil lo logrado hasta el momento. Agradezco la amistad y el apoyo de todos los amigos que conocí durante la maestría.

ABSTRACT

Currently brain-computer interfaces (BCI) based on EEG are mechanisms to, among other objectives, help people with severe motor disabilities to communicate with their environment. However, these are not used in everyday life because that they are not intuitive to control electrophysiological sources. To address this problem, it explored using the electrophysiological source known as imagined speech or unspoken speech. This process can help to generate neural signals patterns more naturally.

The Sonification of EEG signals was used in this work, which allows us to characterize the EEG signal as an audio signal. The objective is to analyze whether the application process sonification of EEG signal can be discriminated or highlight patterns in the signal to improve the results of classification of unspoken words. The signals with and without sonification were processed. We extracted features with two different methods DWT and MFCC which are commonly used in speech recognition.

The results were obtained from the 4 closest to the language areas of Broca and Wernicke and 14 channels of EEG device. The mean percentage accuracy for the 27 subjects in the 2 groups, with 4 and 14 channels of EEG using sonification improved 5.41 % and 2.25 %, respectively. With the results it was found that applying sonification of EEG can better characterize the EEG signal, with respect to work in the state art that does not apply sonification technique. By means of the selection of the dominant frequencies and better distribution of signal energy on a broader spectrum of frequencies, we obtain discriminate patterns that help tp improve slightly better accuracy percentages for the classification of unspoken words.

RESUMEN

Actualmente las interfaces cerebro-computadora (BCI) basadas en EEG son mecanismos que buscan, entre otros objetivos, ayudar a personas con discapacidad motriz severa a comunicarse con su entorno. Sin embargo, éstas aún no son utilizadas en la vida cotidiana por lo poco intuitivas que son las fuentes electrofisiológicas para controlarlas. Para afrontar dicho problema se ha explorado utilizar la fuente electrofisiológica conocida como habla imaginada con el objetivo de generar patrones de neuroseñales que ayuden a la clasificación ocurra de forma más natural.

En este trabajo se utilizó la técnica de sonificación de señales de EEG, la cual nos permite caracterizar la señal de EEG como una señal de audio. El objetivo es analizar si al aplicar el proceso de sonificación de la señal de EEG se puede discriminar o resaltar patrones en la señal que mejoren los resultados de clasificación de palabras no pronunciadas. Para ello se procesó la señal con y sin sonificación. En esta investigación se extrajeron características con dos diferentes métodos DWT y MFCC, éste último comúnmente utilizado en tareas de reconocimiento de voz.

Se obtuvieron los resultados de los 4 canales más cercanos a las áreas de lenguaje de Broca y Wernicke y los 14 canales del dispositivo EEG. Los porcentajes de exactitud promedio para los 27 sujetos en los 2 conjuntos con 4 y 14 canales usando sonificación de EEG mejoran 7.72 % y 5.41 % respectivamente. Con los resultados obtenidos se pudo constatar que al aplicar la sonificación de EEG se puede caracterizar mejor la señal de EEG, con respecto al trabajo en el estado del arte que no aplica dicha técnica. Por medio de la selección de las frecuencias dominantes y la mejor distribución de la energía de la señal en un espectro de frecuencias más amplio se logró discriminar los patrones que apoyan a mejorar ligeramente los porcentajes de clasificación de las palabras imaginadas.

INTRODUCCIÓN

En nuestro país y el resto del mundo existe un gran número de personas con discapacidades. Tan solo en México existen al menos 3.3 millones de personas con discapacidades motrices [1], entre ellas se encuentran personas con una discapacidad cuya gravedad hace que no puedan comunicarse con su entorno. Es por ello que una interfaz cerebro-computadora (Brain Computer Interfaces-BCI) podría ser una vía de comunicación útil para personas con trastornos neuro-musculares graves tales como: lesión de la médula espinal, esclerosis lateral amiotrófica, accidente cerebro-vascular o parálisis cerebral parcial [2].

Actualmente las BCI son una alternativa que provee al cerebro de un nuevo canal de comunicación y control, transmitiendo mensajes y comandos al mundo exterior [3]. Las BCI, en personas con discapacidades (como las mencionadas anteriormente) ayudan a obtener la integración a su entorno y mejorar su calidad de vida, mientras que en personas sanas se utilizan como un medio alternativo de comunicación con dispositivos para aplicaciones enfocadas a la educación o el entretenimiento.

Para adquirir la señal cerebral en una BCI existen diversos métodos como lo son: Imágenes por resonancia magnética funcional (fMRI), magnetoencefalografía (MEG), Electroencefalograma (EEG) y Electrocorticograma (ECoG). De entre estos métodos se decidió utilizar el EEG como método para adquirir la señal cerebral, dado que no es invasivo, es económico y requiere poco tiempo de calibración. Además que este método es factible poder utilizarlo en una aplicación en la vida cotidiana.

Una BCI requiere que se generen señales específicas para hacer efectiva la comunicación cerebro-computadora. Estas señales se denominan fuentes electrofisiológicas las cuales son mecanismos neurológicos que producen un patrón reconocible en cada ser humano. Algunas fuentes electrofisiológicas utilizadas en las BCI comúnmente son: los potenciales evocados (ERP) los cuales incluye a las señales P300 o N100, los potenciales corticales lentos (SCP), las movimientos imaginados (motor imagery), los

potenciales evocados visuales (VEP), el ritmo sensorial mu y más recientemente el habla imaginada, habla no pronunciada o habla silente (Imagined speech, Unspoken speech, Silent speech), la cual refiere a la acción de imaginar la dicción de una palabra pero sin emitir ni articular sonidos [4]. En el habla imaginada no se requiere de un estímulo externo para generar la señal cerebral y el entrenamiento previo es corto.

1.1. Problemática

Actualmente las BCI no tienen un uso extendido fuera de las aplicaciones de laboratorio. Esto se debe a que las tasas de comunicación aún son bajas [5], por lo que no se efectúa una comunicación fluida y natural. Por otra parte las BCI basadas en fuentes electrofisiológicas como ERP, VEP o SCP, requieren que el usuario entrene por un lapso prolongado para poder generar las señales cerebrales voluntariamente. Estos inconvenientes han originado que en años recientes se busque explorar el uso del habla no pronunciada como fuente electrofisiológica para una BCI, la cual pertenece al grupo denominado interfaces de habla silenciosa (silent speech interfaces) [6]. El habla no pronunciada se genera de manera natural y a diferencia de las fuentes electrofisiológicas más usadas en BCIs (VEP, P300, ritmos mu y beta, SCP) cada persona es consciente del momento en que la genera. Por lo que el entrenamiento para generar palabras no pronunciadas por parte del usuario es mínimo. Además, no requiere de traducción al dominio del lenguaje para permitir comunicación [7].

La señal de EEG se ha estudiado ampliamente, pero por sus características y complejidad aún sigue siendo un tema de interés. Está vigente sobre todo en investigaciones con aplicaciones médicas. En muchas de estas investigaciones se requiere tener una retroalimentación de apoyo cuando un especialista revisa la lectura de EEG. En consecuencia, en años recientes se ha investigado la alternativa de usar la sonificación de EEG para dar dicha retroalimentación en forma de sonidos, complementando así a la lectura visual que convencionalmente se realiza, pues la sonificación permite analizar la información mientras se observa alguna otra cosa. La sonificación de EEG permite por ejemplo analizar de manera más eficiente los eventos de epilepsia grabados de un paciente con dicho trastorno [8–11], esto se logra comprimiendo la señal de la lectura de EEG mediante el uso de alguna técnica la sonificación (permitiendo así analizar los episodios epilépticos al escuchar un audio que dura por ejemplo 10 veces menos que la lectura de EEG original) y después aumentar la duración del audio en los eventos encontrados para apreciar los detalles. La sonificación de EEG en general permite distribuir la información requerida en un amplio rango de frecuencias audibles. Los

audios de un EEG sonificado han sido utilizados para ayudar en la toma de decisiones en los diagnósticos médicos, por ejemplo la enfermedad neurológica de Alzheimer, donde algunos trabajos reportan que se ha logrado que personas que no son especialistas puedan discriminar entre personas sanas y personas con una etapa temprana de Alzheimer con solo escuchar el audio generado al aplicar la sonificación [12, 13].

1.2. Solución propuesta

En la clasificación del habla imaginada algunos trabajos manejan un enfoque de reconocer sílabas bajo el concepto de que internamente estamos construyendo palabras con fonemas antes de pronunciarlas, y en algunos de estos trabajos la señal cerebral del habla imaginada ha sido tratada con sistemas para reconocimiento de voz [14–16], pero la señal cerebral tiene diferentes características que la señal de voz, por ejemplo está en distintas frecuencias y puede llegar a ser bastante compleja en cuanto a dimensionalidad y ruido. Por ello es interesante aplicar dichos métodos de extracción de características usados en reconocimiento de voz en audios que representen una palabra pronunciada internamente en el cerebro, es decir caracterizar la señal cerebral como señal de audio usando la sonificación de EEG.

En la presente investigación se utilizará la sonificación dentro de la extracción de características de palabras imaginadas, motivados por el hecho de que existen diversas técnicas de sonificación de EEG que permiten al investigador personalizar las características que son de nuestro interés dentro de la señal cerebral (eliminando o resaltado patrones en la señal de audio). Además de que se ha utilizado en la clasificación de pacientes con enfermedades neurológicas obteniendo buenos resultados. Se conjetura que si el oído humano puede diferenciar patrones de los pacientes al escuchar su sonificación de EEG, entonces un método computacional de aprendizaje automático podría lograr hacerlo de igual forma. Además de que en los casos en los que se ha usado un método de clasificación automática se ha tenido buenos resultados por lo que se conjetura que la sonificación de EEG ayudará a mejorar la clasificación de palabras no pronunciadas.

En esta investigación se aplicará un método de aprendizaje automático sobre la sonificación de registros de EEG que seleccione frecuencias dominantes, para luego evaluar si la sonificación de la señal de EEG puede discriminar o resaltar patrones que mejoren los resultados de clasificación de palabras durante el habla imaginada con respecto al trabajo del estado del arte con el que se compara. Se usarán técnicas de extracción de características comúnmente utilizada para reconocimiento de voz, pero a

diferencia de trabajos anteriores, la extracción se aplicará sobre la señal sonificada de EEG. Hasta donde sabemos no existe trabajo previo alguno que use algún método de sonificación de EEG para la clasificación de habla imaginada.

Todo lo anterior con la finalidad de desarrollar un método eficiente de comunicación que pudiera ser utilizado en una BCI basada en el habla imaginada. La interfaz proporcionaría una alternativa de comunicación a cualquier persona (sana o con alguna condición de las mencionadas anteriormente) que pueda generar conscientemente las señales cerebrales que representan la pronunciación de palabras. En esta tesis no realizará el proceso de desarrollo de una BCI, pues este es un objetivo que va más allá de la presente investigación.

1.2.1. Preguntas de Investigación

- ¿Puede la sonificación aportar una caracterización de la señal de EEG tal que haga mejorar los porcentajes de clasificación en el habla imaginada?.
- ¿Cómo se comportan los parámetros de la técnica de sonificación de EEG en la clasificación automática?
- ¿Se mantendrá la diferencia de exactitud en la clasificación entre usar 14 canales y usar 4 canales correspondientes al modelo Geschwind-Wernicke al usar sonificación de EEG?.
- ¿Pueden los métodos de extracción de características usados en reconocimiento de voz ayudar a extraer características en audios provenientes de la sonificación de palabras imaginadas?.

1.2.2. Hipótesis de investigación

La sonificación de EEG antes de la extracción de características puede resaltar los patrones que ayudan a mejorar la clasificación de cinco palabras imaginadas del idioma español, grabadas desde un dispositivo de EEG de 14 canales.

1.2.3. Objetivo general

Implementar un método para procesar y clasificar palabras no pronunciadas de lecturas de EEG, utilizando una transformación del dominio original al dominio del audio, con la finalidad de obtener resultados en la exactitud de clasificación mayores a los existentes en el estado del arte.

1.2.4. Objetivos específicos

Partiendo del objetivo general se definen los siguientes objetivos específicos:

- Seleccionar e implementar un método de sonificación existente de las señales EEG.
- Probar 2 conjuntos de 14 y 4 canales, para evaluar si un conjunto reducido de canales del modelos del habla permite obtener buenos resultados.
- Establecer y aplicar los métodos de extracción de características del audio.
- Definir y aplicar métodos de clasificación para las características previamente extraídas.

1.3. Contribuciones

- La presente investigación aportará resultados sobre clasificación de palabras no pronunciadas sobre EEG sonificados, lo cual no se ha explorado en trabajos previos.
- Se aportarán resultados que apoyen a saber si las técnicas de extracción de características regularmente usadas para reconocimiento de voz pueden ser usadas en señales sonificadas de palabras imaginadas.
- Se obtendrá el conocimiento del impacto de los parámetros de la sonificación de EEG sobre la clasificación de palabras no pronunciadas.

1.4. Alcances y limitaciones

Se realizará la experimentación de habla imaginada usando una base de datos existente obtenida de sujetos sanos y cuya lengua materna es el español. Por lo que se podrá comparar directamente los resultados obtenidos con los trabajos que la utilizaron. Solo se realizará un esquema de clasificación de palabras imaginadas usando sonificación de EEG pero no se realizará el bucle completo que conlleva la construcción de una interfaz cerebro-computadora. En esta investigación se aplican técnicas de extracción de características sobre el audio de la señal de EEG tales como MFCC o DWT, por lo que no se exploran técnicas adicionales que otros trabajos contemplan como viables. Las etapas de esta investigación se realizarán sobre la

información disponible de 4 y 14 canales, por lo que no se busca obtener información de canales personalizada para cada sujeto.

1.5. Estructura del documento

El resto del documento está estructurado de la siguiente manera. En el capítulo 2 se muestra el marco teórico, el cual incluye la explicación de los conceptos necesarios sobre los temas a tratar en esta tesis. En el capítulo 3 se muestran los trabajos más relevantes relacionados con la clasificación de habla no pronunciada y la clasificación usando sonificación de EEG. El capítulo 4 contiene la descripción del método propuesto para clasificar palabras del habla no pronunciada usando sonificación de EEG. En el capítulo 5 se muestran los resultados de los experimentos realizados, así como de la comparación con trabajos similares. Por último el capítulo 6 incluye las conclusiones de la investigación y las ideas o posibles líneas de investigación a seguir como trabajo futuro. Al final del documento se incluyen anexos que contienen información adicional a lo explicado en los capítulos del documento, como: las publicaciones realizadas y los resultados de clasificación detallados.

MARCO TEÓRICO

2.1. El cerebro

El cerebro humano es una red de más de 100 mil millones de células nerviosas individuales interconectadas formando circuitos neuronales, que construyen nuestras percepciones del mundo externo, nos permiten fijar nuestra atención, y controlar la maquinaria de nuestras acciones [17]. Cada neurona del cerebro está compuesta de un cuerpo celular y prolongaciones tubulares llamadas neuritas. La neuritas son de 2 tipos: axones y dendritas (ver figura 2.3). Las dendritas son ramificaciones cortas que actúan como receptoras de información. Los axones son terminaciones más largas que las dendritas de diámetro uniforme y se encargan de enviar la señal de salida de la neurona [18].

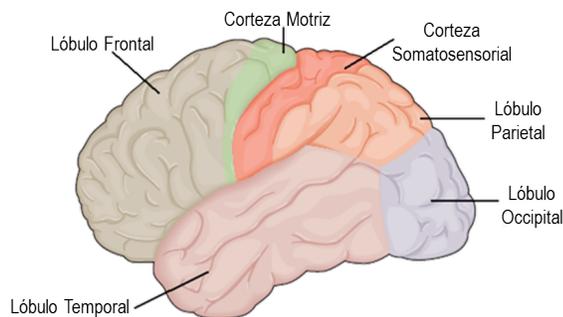


Figura 2.1: División cerebral por lóbulos (figura modificada de [17]).

El cerebro se compone de dos hemisferios cerebrales (izquierdo y derecho), cada uno con estructura simétricamente idéntica pero con distinta funcionalidad. El cerebro tiene una capa exterior muy arrugada (la corteza cerebral) y tres estructuras profundas (ganglios basales, el hipocampo, y la amígdala cerebral). Tal como se muestra en la figura 2.1 cada uno de los hemisferios de la corteza cerebral se puede dividir en 4

lóbulos: frontal, parietal, occipital y los lóbulos temporales, nombrados así de acuerdo a los huesos del cráneo que los recubren. Cada lóbulo tiene varios repliegues profundos característicos, lo cual es un proceso evolutivo estratégico para el embalaje de más células nerviosas en un espacio limitado. Cada lóbulo tiene un conjunto especializado de funciones. El lóbulo frontal (frontal lobe) se relaciona en gran parte con la memoria a corto plazo y la planificación de acciones futuras y con control de movimiento; el lóbulo parietal (parietal lobe) con la sensación somática, con la formación de una imagen corporal y lo relacionado con el espacio extra-personal; el lóbulo occipital (occipital lobe) con la visión; y el lóbulo temporal (temporal lobe) con la audición y a través de sus estructuras profundas, el hipocampo y los núcleos de la amígdala con el aprendizaje, la memoria y las emociones [17].

2.1.1. Procesamiento del lenguaje

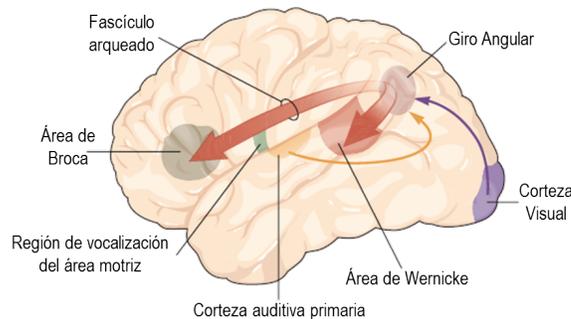


Figura 2.2: Modelo Geschwind-Wernicke (figura reproducida de [17]).

En la figura 2.2 se muestra el modelo Geschwind-Wernicke, el cual muestra las áreas involucradas en el procesamiento del lenguaje en nuestro cerebro [19]. En la mayoría de las personas el procesamiento del lenguaje lo realiza el hemisferio izquierdo de forma dominante. Según el modelo Geschwind-Wernicke, los primeros pasos en el procesamiento neuronal de las palabras habladas o escritas ocurren en las áreas sensoriales de la corteza especializadas en la información auditiva (Primary auditory cortex) o visual (Visual cortex). Esta información es entonces transportada a un área de asociación cortical, el giro angular (Angular gyrus), especializado para procesar tanto la información auditiva y visual. Aquí, según Wernicke, las palabras habladas o escritas se transforman en un código sensorial neuronal compartido para el habla y la escritura. Esta representación se transporta hasta el área de Wernicke (Wernicke's area), donde es reconocida como lenguaje y es asociada con un significado. También se transmite al área de Broca (Broca's area), que contiene las reglas, o gramática, para

la transformación de la representación sensorial. La representación se transforma en una representación motora que puede usarse para efectuar lenguaje hablado o escrito, usando la región de vocalización del área motora (Vocalization region of motor area) [17].

2.2. Electroencefalograma (EEG)

Un dispositivo de EEG mide la actividad eléctrica generada por las corrientes que fluyen durante la excitación sináptica de las dendritas de las neuronas piramidales de la corteza cerebral [20]. Esta medición se hace mediante electrodos ubicados en el cuero cabelludo. Un solo electrodo estima el promedio de las actividades sinápticas de millones de neuronas [21].

En la figura 2.3 se muestra la generación de un pequeño campo eléctrico creado durante la sinapsis de las neuronas piramidales. Cuando se dispara el axón aferente, la terminal presináptica libera glutamato, que abre los canales de cationes. La corriente positiva fluye en la dendrita, dejando una ligera carga negativa en el fluido extracelular. La corriente se propaga hacia abajo de la dendrita y escapa a sus partes más profundas, dejando el líquido ligeramente positivo en esos sitios. El electrodo registra la actividad solo si miles de neuronas contribuyen con su pequeño voltaje y pasa a través de las capas de tejido grueso [18].

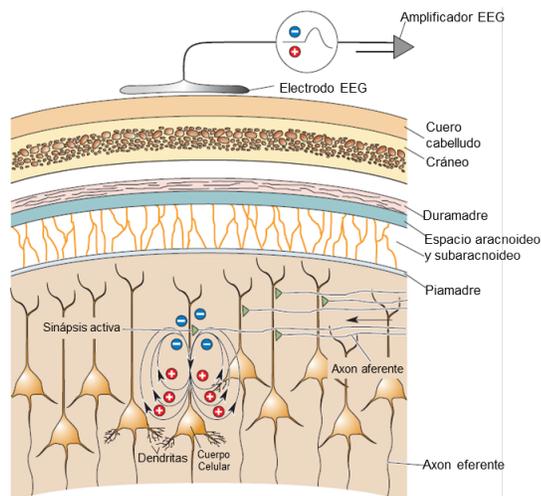


Figura 2.3: Electrodo de EEG registrando actividades sinápticas de las neuronas (figura modificada de [18]).

Los electrodos en un dispositivo de EEG también se conocen como canales, pues cada uno capta una parte de la señal total emitida por el cerebro. Para realizar la

medición de la señal cada electrodo debe colocarse en una posición específica en el cuero cabelludo, el método más utilizado para colocarlos es el sistema internacional 10-20 [22]. En el sistema 10-20, los electrodos están colocados a una distancia de separación relativa entre ellos de 10 ó 20 por ciento. Cada posición dentro del sistema 10-20 tiene una letra para identificar el lóbulo y un número el hemisferio. Las letras F, P, T, C y O corresponden a los lóbulos Frontal, Parietal, Temporal, Central y Occipital respectivamente. Los números impares corresponden al hemisferio izquierdo y los pares al hemisferio derecho (ver figura 4.2).

Para adquirir la señal mediante electroencefalograma existen diversos dispositivos los cuales tienen características diferentes (ver figura 2.4). Los dispositivos como EEG Cap y BrainVision tienen mejores prestaciones en cuanto a calidad de la señal y pueden adquirir la información de más canales de forma simultánea, pero dado su diseño, no es fácil colocarlos en una persona y requieren más tiempo de calibración, por lo que están enfocados a investigaciones de laboratorio. Por otra parte los dispositivos como MindWave de Neurosky o Eporc de Emotiv, poseen 1 y 14 electrodos respectivamente, son más económicos, son más sencillos de colocar y requieren muy poco tiempo de calibración, aunque tienen la desventaja de que la calidad de la señal es menor, ya sea por tener menos canales (MindWave) o por tener menor frecuencia de muestreo (Eporc).



Figura 2.4: Dispositivos para hacer lectura de EEG, de izquierda a derecha, EEG de BrainVision, EEG Cap, MindWave de Neurosky, Ephoc de Emotiv (figuras reproducidas de [23], [24], [25], [26]) .

La actividad de EEG puede medir oscilaciones continuas o cambios de potencial relacionado a la aparición de algún evento (ritmos cerebrales). Se han establecido ritmos cerebrales de acuerdo con: bandas de frecuencia dentro del espectro del EEG, la amplitud de la señal y aspectos inherentes de cada uno de ellos [27], aunque los parámetros pueden variar en cada sujeto. Los 6 ritmos cerebrales más importantes que se presentan en un sujeto son:

- Ritmo Delta (0.5 - 3.5 Hz): Es de gran amplitud (100-200 μ V). En adultos, este ritmo aparece solo en un estado de sueño muy profundo. Si se presenta en un

sujeto despierto puede deberse a un síntoma patológico [27,28].

- Ritmo Theta (4 - 7 Hz): Posee una amplitud menor a 30 μV . Es un ritmo poco común, se presenta más a menudo en niños durante el sueño, mientras que en adultos se percibe en estados de mucho estrés [27,28].
- Ritmo Alfa (8 - 13 Hz): Tiene una amplitud de 30 a 50 μV . Se produce en cualquier persona con los ojos cerrados o en estado de relajación y poca actividad mental. Este ritmo se atenúa con los ojos abiertos. Es más prominente en la región posterior del cerebro [27,28].
- Ritmo Beta (14 - 30 Hz): Tiene una amplitud de 5 a 20 μV . Se presenta en estados de concentración mental, también se encuentra estrechamente relacionado con el movimiento de las extremidades y se concentra principalmente en la región central y frontal del cerebro [27,28].
- Ritmo Gamma (31 ~ 50Hz): Se presenta con una amplitud entre 5 y 10 μV . La actividad en esta banda está relacionada con la atención selectiva. Esta actividad también está relacionada con la cognición, la actividad perceptual y sensorial [28,29].
- Ritmo Mu (7 - 13Hz): Está presente en la mayoría de los adultos, y aunque tenga características similares a las del ritmo alfa, sus propiedades son distintas. Se relaciona con las funciones motoras del cerebro y no se atenúa con los ojos abiertos sino cuando se realiza algún tipo de movimiento o se intenta realizar dicho movimiento. Este ritmo se registra en la región central del cerebro [30,31].

2.2.1. Artefactos EEG

Al realizar la lectura de la señal cerebral a través de un dispositivo de electroencefalografía se puede obtener ruido provocado por diversas fuentes ajenas al cerebro, a estas alteraciones se les conoce como artefactos [32]. Formalmente se define artefacto como un registro EEG de actividad no cerebral y se pueden dividir en fisiológicos y extrafisiológicos. Los primeros surgen del usuario directamente, como el movimiento ocular y los segundos se generan por fuera del paciente como los equipos o el medio ambiente. Los artefactos pueden ser causados por la corriente alterna o por otras fuentes fisiológicas del cuerpo como el movimiento de los ojos, el latido del corazón, (conocidos como artefactos electrocardiográficos (ECG)), la respiración, el sudor, las contracciones musculares (conocidos como Artefactos electromiográficos (EMG)), entre

otros [32]. En tareas de clasificación, remover los artefactos en la señal de EEG puede ayudar a limpiar la señal de ruido.

2.3. Brain Computer Interfaces (BCI)

Una interfaz cerebro computadora (Brain Computer Interface) hace referencia a un mecanismo de comunicación entre una persona y un sistema computacional usando solamente la actividad cerebral generada por dicho individuo. Una BCI es un sistema de comunicación que no depende de las vías de salida normal del cerebro, como los nervios periféricos y los músculos [33]. Uno de los principales objetivos de BCI es manipular las señales eléctricas generadas por las neuronas del cerebro y generar las señales necesarias para controlar algunos sistemas externos [2]. Al usar EEG, las señales eléctricas necesarias para controlar una BCI se conocen como fuentes electrofisiológicas. Las principales fuentes electrofisiológicas para controlar BCIs son:

- Potenciales relacionados a eventos (Event related potentials): Estos potenciales se generan cuando existen estímulos visuales, auditivos o somatosensoriales infrecuentes o particularmente significativos. Cuando estos estímulos son intercalados con estímulos rutinarios o frecuentes, regularmente disparan un pico positivo en el EEG sobre la corteza parietal después del paso de algunos milisegundos. P300 o N100 son ejemplos de estos potenciales [34].
- Potenciales corticales lentos (Slow cortical potentials): Son cambios lentos en el voltaje de la señal en las frecuencias más bajas. Estos cambios de potencial ocurren durante 0.5 a 10 segundos. Los SCPs negativos están típicamente asociados con movimientos y otras funciones que impliquen activación cortical, mientras que los SCPs positivos están generalmente asociados con la activación cortical reducida [34].
- Potenciales visuales evocados (Visual evoked potentials): Es un tipo de potencial relacionado a eventos. Son pequeños cambios de potencial registrados sobre la corteza visual, derivados de una respuesta cerebral a un estímulo visual. Estos potenciales son más prominentes en la zona occipital. Si un estímulo visual se presenta repetidamente a un ritmo de 5 a 6 Hz o mayor, se presenta una respuesta eléctrica oscilatoria continua en las vías visuales. Esta respuesta se conoce como potenciales evocados visuales de estado estable (SSVEP) [34, 35].
- Ritmos sensoriales motrices: Es el uso de las características de los ritmos de las señales cerebrales. De especial interés es el ritmo mu, pues la amplitud

de las oscilaciones en este ritmo disminuye cuando se realiza o se imagina un movimiento. Los cambios en el ritmo mu son localizados sobre la corteza sensorial-motora correspondiente a la parte del cuerpo que se desea mover. Para generar cambios en el ritmo mu, generalmente se requiere una etapa de entrenamiento con retroalimentación [34,36].

En la actualidad existen diversas BCI que utilizan las fuentes electrofisiológicas antes mencionadas. De especial interés son las BCI que apoyen a la tarea de generar mensajes [37]. Por ejemplo en la BCI mostrada en la figura 2.5 se construyen palabras letra por letra. En esta BCI se utiliza el potencial P300, registrando la información que se dispara cuando el estímulo visual se relaciona con la letra deseada.



Figura 2.5: Interfaz Cerebro Computadora 'Intendix', basada en la fuente electrofisiológica P300 (figura reproducida de [38]).

2.4. Sonificación

El concepto de Sonificación (en inglés 'Sonification' o 'Auditory display'), se refiere al uso de sonido no hablado (non-speech audio) para transmitir información [39]. La sonificación es la contraparte de la visualización, es un método que en lugar de asignar posiciones específicas a los valores abstractos de acuerdo a las reglas uniformes en un diagrama, asigna sonidos a ellos, también siguiendo normas uniformes. La sonificación es la disciplina de exploración de datos mediante procesos de transformación de los datos en sonido con el propósito de facilitar la comunicación y la interpretación [40,41]. La sonificación usa las capacidades del oído humano para interpretar datos escuchando frecuencias en el rango de 20 Hz a 20 KHz aproximadamente.

Existen diversas técnicas de sonificación y elegir la correcta depende de los objetivos planteados en la investigación. En el trabajo [42] se describen algunas de ellas. La sonificación se puede utilizar para aumentar la percepción cuando los ojos están ocupadas por otra tarea, como en cirugías médicas, o se puede aplicar como medio de

interacción con el ambiente para personas con alguna patología ocular [43].

Con la sonificación se puede aumentar el uso de la información acústica en aplicaciones de realidad virtual o mejorar los sistemas de alarma [44]. En la sonificación se usa el sentido del oído el cual está adaptado a escuchar varias señales a la vez, y a diferencia de los ojos siempre está activo y es omnidireccional. Además nuestro sentido del oído es bueno en condiciones de ruido, pudiendo filtrar alguna señal específica dentro de un conjunto de señales. Por ejemplo la sonificación se ha utilizado para escuchar los atributos de la base de datos 'Iris', la cual comúnmente es utilizada para probar algoritmos de agrupamiento [45].

2.4.1. Sonificación de EEG

La sonificación de EEG es el proceso de transformar la señal de EEG a patrones dentro de una señal audible. La sonificación de EEG, es el uso de cualquier método de sonificación para transformar la lectura de ondas cerebrales a sonidos con el objetivo de transmitir información que facilite la comprensión y la comunicación. Cuando se muestra al usuario el resultado audible de la sonificación se presenta una forma de retroalimentación neurológica (Neurofeedback), la cual se ha usado entre otras cosas, para entrenamiento [46] o como terapia para pacientes [47].

En los primeros usos de la sonificación se conectaba la salida de electrodos de EEG a un conjunto de instrumentos musicales con la finalidad de analizar el comportamiento que las oscilaciones naturales del cerebro tienen sobre los cambios sonoros [48,49].

En años recientes diversos investigadores han desarrollado trabajos relacionados con las técnicas y aplicaciones de la sonificación de EEG. Thomas Hermann y sus colaboradores han investigado las ventajas y desventajas de hacer sonificación de EEG [50]. Han presentado métodos de sonificación de EEG para mostrar la correspondencia entre las actividades neurales y cognitivas, como por ejemplo el método de sonificación por matriz de distancias entre filas de electrodos. En este método cada fila de electrodos tiene un tono distinto, su volumen es representado de acuerdo a las similitudes obtenidas entre cada par de electrodos y dependiendo de la ubicación de los electrodos se representan en el canal del audio correspondiente al hemisferio en el que se ubican [51]. También se han propuesto los métodos de sonificación basada en eventos [10,52], y el método de sonificación de EEG Vocal [53,54], los cuales han servido para analizar a pacientes con eventos de epilepsia (estos trabajos se describen en la sección 3.2).

Francois Vialatte ha desarrollado la técnica de modelado de elipsoides (bump modeling) para hacer sonificación de EEG y aplicarla a la detección de etapas tempranas de la enfermedad neurológica de Alzheimer [55,56] (este trabajo se describe en la

sección 3.2).

Por otra parte la sonificación de EEG también ha sido utilizada para buscar entender mejor el comportamiento que ocasionan los estímulos visuales periódicos, haciendo sonificación de EEG multicanal para luego analizando la coherencia entre audios de pares de electrodos [57]. La sonificación de EEG ha sido utilizada para hacer composiciones musicales en tiempo real estimulando a los sujetos por medio de SSVEP [58] o representando a distintos sujetos como instrumentos individuales, en el cual los sujetos accionaban su instrumento manipulando su ritmo mu por medio de movimientos imaginados [59]. También se ha usado para entretenimiento o educación dentro de aplicaciones interactivas [60].

2.5. Transformada de Fourier

La serie de Fourier permite que una señal periódica se descomponga en una suma de las ondas seno y coseno con valores reales o más generalmente una suma de exponenciales complejos.

La transformada rápida de Fourier (FFT) es la implementación optimizada de la transformada discreta de Fourier (DFT); dicha optimización aprovecha la redundancia de operaciones, reduciendo la cantidad de ellas al mínimo posible [61]. La DFT, en su forma directa, se define como:

$$X(f) = F_f[x(n)] = \sum_{n=0}^{M-1} x(n) e^{j2\pi n/M} \quad (2.1)$$

donde $x(n)$ es la muestra actual de la señal a transformar y M es la longitud (cantidad de muestras) de la DFT.

Al aplicar la transformada de Fourier se puede realizar el análisis espectral de una señal, el cual consiste en obtener su distribución en el dominio de la frecuencia, es decir, se trata de descomponerla en las frecuencias que la forman, dentro de un rango dado, obteniendo, además, que peso tiene cada una de ellas en el conjunto.

El espectrograma es la representación visual del espectro de frecuencias de una señal, mostrando las variaciones con respecto a una variable (generalmente el tiempo).

La transformada de Fourier de tiempo corto (STFT) es simplemente una secuencia de FFT de segmentos de datos con ventanas, donde las ventanas son generalmente solapadas en el tiempo. En la figura 2.6 se muestra un ejemplo de espectrograma de una señal de audio.

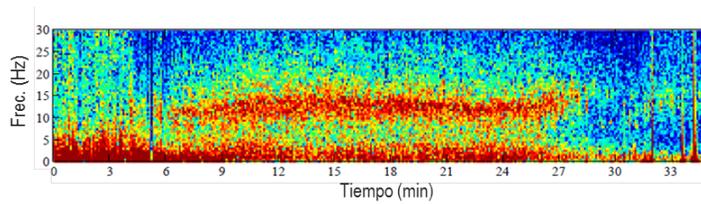


Figura 2.6: Ejemplo de espectrograma de una señal de EEG (imagen reproducida de [62]).

2.6. Transformada Wavelet

La transformada Wavelet (transformada de onduleta) combina las dimensiones tiempo y escala permitiendo la extracción de los detalles y de las aproximaciones de la señal. La transformada Wavelet es una técnica de estimación espectral en donde una función puede ser expresada como una serie infinita de onduletas. El objetivo de esto es representar la función como una combinación lineal de un conjunto particular de funciones, obtenidas mediante la traslación y escalamiento de una función básica llamada Wavelet Madre. La transformada Wavelet permite encontrar discontinuidades o picos de corta duración en la señales que de otra manera sería complicado detectar y tratar. El principio de la transformada se basa en una traslación telescópica de la Wavelet básica Ψ , denotada por:

$$\Psi_{a,b}(t) = \frac{1}{\sqrt{a}} \Psi\left(\frac{t-b}{a}\right) \quad (2.2)$$

Donde a y b son llamados el parámetro de escala y de posición respectivamente. La transformada Wavelet de cualquier función $f(t)$ (frecuencia-tiempo) se obtiene mediante el producto interno de la función $f(t)$ con la función wavelet.

$$W_f(a,b) = \{f(t), \Psi_{a,b}(t)\} \quad (2.3)$$

Existe un largo número de familias Wavelet conocidas y funciones que proveen un rico espacio en el que se puede buscar una onda que represente eficientemente la señal de interés. Entre las familias se encuentran: Bi-ortogonal [63], Coiflet [64], Haar [65], Daubechies [66]. La elección de la Wavelet depende de la aplicación en la que se esté trabajando, buscando seleccionar la que más coincida con la señal que se desea procesar.

2.6.1. Energía Relativa Wavelet (RWE)

Una vez aplicada la transformada discreta Wavelet (DWT) sobre la señal obtendremos coeficientes de aproximación y de detalle, desde los cuales es posible calcular

la energía relativa wavelet [67]. La energía relativa wavelet representa la energía que algún nivel de descomposición aporta al total de la energía wavelet de la señal, con lo que se puede caracterizar la distribución de energía de la señal en diferentes bandas de frecuencia. Al aplicar la DWT sobre una señal usando una wavelet madre Ψ y un número de niveles de descomposición N , se obtiene un conjunto de coeficientes de detalle ($d_{j,k}$; para $j = 1, \dots, N$) y un conjunto de coeficientes de aproximación $a_{N,k}$ a los que se denotará como coeficientes a_k en el nivel $N + 1$. A partir dichos coeficientes es posible calcular la energía relativa wavelet en cada uno de los niveles de descomposición.

La energía relativa wavelet para el j -ésimo nivel de descomposición se define como:

$$RWE_j = \frac{E_j}{E_{total}}; \quad \text{Para } j = 1, \dots, N + 1; \quad (2.4)$$

donde E_j representa la energía en el j -ésimo nivel de descomposición y E_{total} representa la energía total de los coeficientes wavelet de una señal dada. La energía en el j -ésimo nivel de descomposición E_j se define como:

Para $j = 1, \dots, N + 1$.

$$E_j = \begin{cases} \sum_k |d_{j,k}|^2; & \text{Si } j \leq N; \\ \sum_k |a_k|^2; & \text{en caso contrario.} \end{cases} \quad (2.5)$$

Mientras que, la energía total E_{total} se calcula como sigue:

$$E_{total} = \sum_{j=1}^{N+1} E_j. \quad (2.6)$$

2.7. Coeficientes cepstrales en la frecuencia Mel (MFCC)

Los coeficientes cepstrales en la frecuencia Mel (Mel-frequency Cepstral Coefficients) son características espectrales que se calculan a partir del análisis espectral en tiempo corto derivado de la transformada rápida de Fourier (cepstrums). Estos coeficientes aproximan el comportamiento de la señal en escalas de frecuencia lineal a la escala de frecuencias Mel, la cual es no lineal.

Esta técnica comúnmente se utiliza en señales de audio, dado que la escala de frecuencias Mel está diseñada para captar características fonéticamente importantes en el habla. Los MFCCs se calculan mediante el uso de grupos de filtros (filterbank). Los filtros suelen ser triangulares, como se muestra en la figura 2.7. Tales filtros calculan el

espectro en torno a cada frecuencia central con el aumento de ancho de banda [68]. La frecuencia central de cada filtro es obtenida usando:

$$\text{Mel}(x) = 2595 \log_{10} \left(1 + \frac{x}{700} \right) \quad (2.7)$$

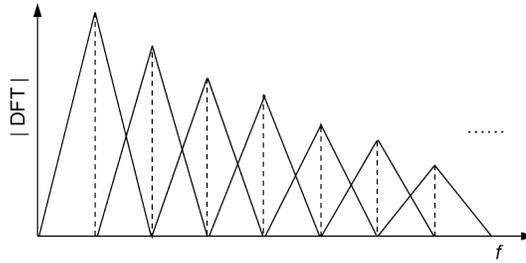


Figura 2.7: Ejemplo de grupo de filtros triangulares usado para calcular los MFCCs (figura reproducida de [69]).

Después de definir las frecuencias más bajas y más altas del banco de filtros y el número de filtros, las frecuencias límite del banco de filtros estarán uniformemente espaciados en la escala Mel. Después se calcula la energía en la salida de cada filtro. El cepstrum en la frecuencia-Mel es la Transformada Discreta del Coseno (DCT) de las salidas de los filtros de energía:

$$C[n] = \sum_{m=0}^{M-1} S[m] \cos(\pi n(m-1/2)/M); \quad 0 \leq n < M \quad (2.8)$$

Donde $C[n]$ es la energía $S[m]$ es la salida de cada filtro, y M es el número de filtros, lo cual varía para diferentes implementaciones. La ventaja del cálculo de los MFCCs usando las energías del banco de filtros, es que son más robustos al ruido y a los errores de estimación espectral. Se recomienda calcular la energía-log por separado para cada segmento de la señal [69]:

$$E = \log \sum_{n=1}^N x^2[n] \quad (2.9)$$

Las características señaladas anteriormente no disponen de información temporal. Con el fin de incorporar el cambio en curso sobre varios coeficientes, se añaden derivadas temporales al vector de características básicas. Se aplica la primera y segunda derivada de la función para obtener los coeficientes denominados como Delta y Doble Delta, respectivamente [69].

2.8. Clasificación automática

De acuerdo con Michie *et al.* [70], la clasificación cubre cualquier contexto en el que alguna decisión o pronóstico es hecho sobre la base de información histórica disponible. Esta base de información disponible D se define de la siguiente forma:

$$D = \{\langle x_1, y_1 \rangle, \langle x_2, y_2 \rangle, \dots, \langle x_m, y_m \rangle\} = \langle X, Y \rangle \quad (2.10)$$

donde los valores $x_i \in X$ son típicamente vectores multi-dimensionales de la forma: $x_i = \{z_1, z_2, \dots, z_n\}$ cuyos elementos pueden tomar valores reales o discretos. Estos componentes se denominan atributos (o características). El objetivo es inferir una función (o relación) f .

$$f : X \rightarrow Y. \quad (2.11)$$

donde los valores de Y están contenidos en un conjunto finito de clases $C = \{C_1, \dots, C_k\}$ que caracterizan los datos dados. Los modelos aprendidos de los datos de entrenamiento son evaluados con un conjunto de prueba distinto para determinar si los modelos pueden ser generalizados a nuevos casos.

Actualmente existen diversas formas de obtener la función que clasifica a los datos, cada uno con propiedades distintas, por lo que elegir el adecuado dependerá de las características conjunto de datos a clasificar y de los objetivos de la investigación. A continuación se describe el funcionamiento del clasificador Random Forest, el cual fue utilizado en este trabajo de tesis.

2.8.1. Random Forest

Random Forest es un algoritmo de clasificación basado en árboles de decisión. RF es una combinación de árboles predictores, donde cada uno de los árboles depende de los valores de un vector aleatorio muestreado independientemente y con la misma distribución para todos los árboles en el bosque. En Random Forest cada árbol arroja un único voto para la clase más popular para una entrada dada, y al final la elección de la clase se realiza usando voto mayoritario [71]. Random Forest se caracteriza por poder clasificar en corto tiempo bases de datos que poseen ruido y una gran cantidad de atributos [72].

En el algoritmo 1 se muestra el proceso de clasificación de Random Forest.

2.9. Prueba estadística de resultados

En este trabajo de investigación para la prueba de hipótesis se usará la prueba paramétrica *ANOVA de un factor* (análisis de varianza de un factor) [73], la cual se

Algorithm 1 Etapa de clasificación del clasificador RF

Entradas: $x \in X$ (una instancia x del conjunto de instancias de prueba X), $|\text{dom}(y)|$ (el número de clases del problema en cuestión), M_i (el i -ésimo árbol de decisión generado con el algoritmo de generación de árboles), y T (el número de árboles miembros de RF).

Salida: C (predicción de la clase)

Contador₁, ..., Contador _{$|\text{dom}(y)|$} $\leftarrow 0$ {inicializar los contadores de votos para las clases}

for $i \leftarrow 1$ hasta T **do**

voto _{i} $\leftarrow M_i(x)$ {obtener la clase predicha por el clasificador M_i para la instancia x }

Contador_{voto i} \leftarrow Contador_{voto i} + 1 { incrementar uno al contador de la clase correspondiente }

end for

$C \leftarrow$ la clase con el mayor número de votos.

Retornar C .

eligió para comparar los promedios de clasificación de 4 grupos: EEG sin sonificar y Sonificación de EEG en los dos métodos de extracción de características con DWT y MFCC. Esta prueba es una generalización de la prueba de comparación de medias de 2 grupos independientes, la cual se puede aplicar a 3 o más grupos independientes. La prueba trabaja bajo los supuestos de normalidad y de homogeneidad de varianzas. Además esta prueba requiere de un factor el cual agrupa las muestras de acuerdo a los distintos tratamientos. Si al aplicar la prueba ANOVA de un factor el resultado es positivo se concluye que al menos un par de grupos tiene una diferencia significativa, por lo que para conocer el resultado de los grupos de manera individual se debe de realizar adicionalmente una prueba estadística que permita la comparación por pares, como la prueba de Tukey.

TRABAJOS RELEVANTES RELACIONADOS

En este capítulo se incluyen los trabajos más relevantes relacionados con este proyecto de tesis. Los trabajos del estado del arte se muestran en dos secciones. La primer sección se describirán trabajos sobre clasificación de habla imaginada. Se describirán por bloques de acuerdo al enfoque seguido, ya sea clasificación de sílabas, fonemas, vocales o palabras. En la segunda sección se describirán los trabajos que han usado técnicas de sonificación de EEG para alguna tarea de clasificación.

3.1. Clasificación de habla imaginada

La idea de tener un mecanismo para transmitir mensajes usando la actividad cerebral ha avanzado. Los primeros estudios se remontan a 1967, cuando el laboratorio de investigación de la fuerza aérea de los Estados Unidos desarrolló un sistema para transmitir mensajes mediante la actividad cerebral [74] usando el código Morse. Con los resultados obtenidos ya se mostraban indicios de que el cerebro podría proveer de un mecanismo alternativo de comunicación, pero era un área que debía ser investigada más a fondo. Recientemente, el Departamento de Defensa de EE.UU. ha invertido 4 millones de dólares a su proyecto *MURI* en el cuál el objetivo es investigar y construir un modelo que sienta las bases de la telepatía sintética.

Como se describió anteriormente en la problemática, los problemas actuales en las BCIs derivados del uso de las fuentes electrofisiológicas, han motivado trabajos que tratan de utilizar los potenciales más relacionados con la producción del habla. Los trabajos relacionados con la clasificación de habla no pronunciada han seguido dos enfoques: clasificación de sílabas, vocales o fonemas y clasificación de palabras. Se sigue el enfoque de sílabas o fonemas bajo la hipótesis de que se podría reconocer los elementos básicos del lenguaje dentro del habla imaginada al igual que el habla convencional y con ello diseñar un reconocedor de habla imaginada basado en los

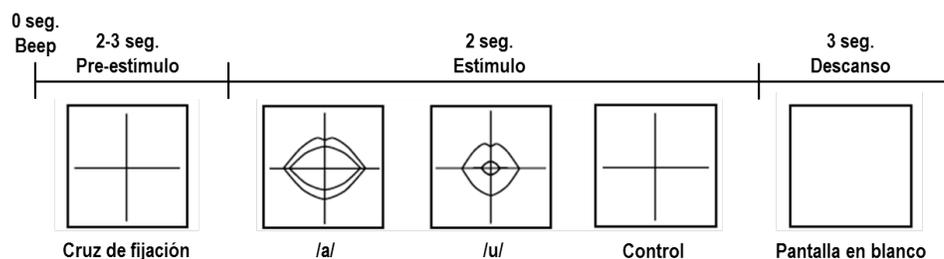


Figura 3.1: Esquema de grabación de una repetición para la investigación realizada por [75].

métodos utilizados en los reconocedores de voz, los cuales actualmente tienen muy buenos resultados reconociendo palabras del lenguaje humano.

El enfoque que busca clasificar palabras del habla imaginada tiene como base la evidencia que existe sobre el almacenamiento de palabras como conceptos en el cerebro. De acuerdo a esto cada palabra que un sujeto pronuncia de forma convencional es parte de un concepto único almacenado en el cerebro, por lo tanto se podría clasificar cada palabra al ser pronunciada internamente.

3.1.1. Clasificación por vocales, sílabas o fonemas

El estudio realizado por [75] clasifica la pronunciación imaginada de vocales. En dicho estudio se grabó la señal de EEG de 3 sujetos sanos, mientras pronunciaban de forma imaginada las vocales del lenguaje Inglés /a/ y /u/, además se graba un estado de no acción para usarlo como control. Se adquirió la señal usando un dispositivo de EEG con 72 electrodos con una frecuencia de muestreo de 256 Hz. En el esquema de grabación cada sujeto escucha un sonido que indica que debe fijar su mirada en una cruz en la pantalla que dura de 2 a 3 segundos. Después durante 2 segundos aparece en la pantalla una imagen que simula el final de la pronunciación de las vocales, lo que indica que se debe de pronunciar de forma imaginada las vocales /a/ y /u/ respectivamente. En el estado de control la cruz de fijación permanece constante. Cada sujeto repitió esto 50 veces para cada categoría. Luego se removió ruido que pudiera provenir de otras fuentes filtrando la señal a 1-45 Hz. Después se usó el método de patrones espaciales comunes (CSP), donde se promedian las repeticiones de acuerdo a la posición del electrodo. Después se divide lo que resultó en conjuntos de entrenamiento y de prueba para clasificar usando de Máquina de Vectores de Soporte (SVM). En dicho estudio se obtuvieron porcentajes exactitud entre $60 \pm 12\%$ y $78 \pm 5\%$.

El trabajo de [76], se basa en estudios que reportaban que hay suficiente información

en la señal de EEG para clasificar habla imaginada a través de estímulos audibles [77,78]. Esta investigación busca clasificar 2 sílabas /ba/ /ku/ en 3 diferentes ritmos de pronunciación. En el estudio se grabaron las señales de 4 sujetos usando un dispositivo de EEG de 118 canales, con una frecuencia de muestreo de 1024 Hz. Los sujetos grabaron 20 repeticiones de la pronunciación imaginada de las sílabas en los 3 diferentes ritmos. En cada repetición el sujeto escucha mediante audífonos la sílaba que deberá imaginar y mediante clics el ritmo al cual debe repetir la pronunciación tal como se muestra en el ejemplo de la figura 3.2. En este trabajo se filtro la señal de

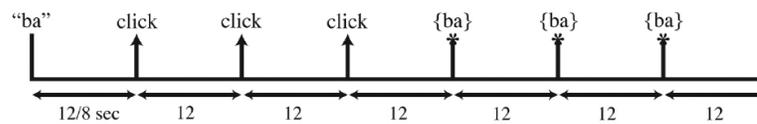


Figura 3.2: Esquema de grabación de los trabajos [76,79]. El tiempo está dado en octavas de segundo. Los clics que indican el ritmo están distribuidos cada 12 octavas (1.5 segundos) (figura reproducida de [76]).

acuerdo a 3 bandas de frecuencia (*theta*, *alpha* y *beta*) y se removió la tendencia lineal, la media de cada segmento de la señal y el artefacto electromiográfico. Después se usó la clasificación por emparejamiento de filtros (*matched-filter classification*) donde las envolturas obtenidas eran los filtros emparejados de entrada del clasificador. Este trabajo obtuvo 52.25% 74.25% y 56.75% de exactitud promedio para las 3 bandas respectivamente.

La investigación de [80] está basada en el trabajo de [76]. En este trabajo se tiene el mismo objetivo de clasificar las sílabas /ba/ y /ku/, pero imaginadas a un único ritmo y se extendió el estudio a 7 sujetos. Las etapas de procesamiento usadas en el trabajo base se mantienen y se añade un análisis de componentes independientes para remover artefactos. Se utilizaron coeficientes auto-regresivos (AR) para extraer características. Después se usó un clasificador de k-vecinos más cercanos con distancia euclidiana. Con el enfoque descrito se logró tener una exactitud de clasificación promedio de 61% para los 7 sujetos con 85 repeticiones en total.

En el trabajo realizado por [81] se busca clasificar fonemas dentro del habla imaginada. El esquema se compone de diez fonemas del idioma inglés correspondientes a cinco categorías de acuerdo a su forma de articularlos (ver tabla 3.1).

En este estudio participaron 5 sujetos. Se grabó las señales con un dispositivo EEG de 52 electrodos con una frecuencia de muestreo de 1000 Hz. Cada sujeto realizó 12 repeticiones de cada fonema y 12 repeticiones de dos clases de periodos de relajación o pausa. Las señales fueron procesadas para remover una línea base de ruido con la

Tabla 3.1: Conjunto de fonemas utilizados en el estudio [81].

Articulación	Mandíbula	Lengua	Nasal	Labial	Fricativa
Fonema	-aa, -ae	-l, -r	-m, -n	-uu, -ow	-s, -z
Ejemplos	(saw,jaw), (hat,cat)	(light, led), (right, red)	(mat, mice), (net, nice)	(who, drew), (boat, over)	(same, hiss), (zoo, his)

referencia promedio común. Después se les aplicó funciones Gabor truncadas para obtener espectrogramas en el espacio logarítmico entre 4 y 28 Hz. Cada segmento del espectrograma fue usado como entrada para dos clasificadores: Análisis discriminante lineal y Naive Bayes. Ambos clasificadores evaluaron un electrodo a la vez. Después se usó un esquema de votación, donde los mejores nueve electrodos clasificados entre los clasificadores representaban el porcentaje de exactitud del sujeto. Los resultados obtenidos van de 66.4 % hasta 76.0 % de exactitud promedio para los 5 sujetos.

3.1.2. Clasificación de palabras

Los primeros trabajos bajo el enfoque de clasificar palabras del habla imaginada fueron realizados por Suppes [77, 78], en el primero buscaba reconocer palabras y en el segundo buscaba reconocer sentencias. En el estudio [77] se grabó la actividad cerebral de 7 sujetos mientras pronunciaban internamente 7 palabras del idioma inglés: (*first, second, third, left, right, yes, no*). Para grabar la señal se usó un dispositivo EEG de 16 electrodos y un dispositivo de magnetoencefalografía (MEG) de 148 sensores. Durante la grabación de la señal el sujeto observaba en un monitor la siguiente palabra que debía pronunciar internamente. Para la parte de habla interna solo participaron 5 sujetos y cada uno completó 100 repeticiones para cada palabra. Después de obtener la señal cerebral se aplicó un promedio de línea base para eliminar ruido. De las 100 repeticiones de cada palabra se tomaron 50 para formar un prototipo promediando las ondas de la señal cerebral. Con las otras 50 repeticiones se formaron 70 prototipos de prueba. A cada prototipo de señal se le aplicó la FFT y el resultado fue filtrado usando un filtro de Butterworth. Se utilizó mínimos cuadrados (Least squares), como una medida de ajuste óptimo entre el prototipo y la muestra de prueba. Usando el enfoque anteriormente descrito este trabajo obtuvo un 52.57 % de exactitud promedio para 5 sujetos.

En el trabajo [16] se investigó clasificar la producción de distintas formas del habla (normal, susurrada, murmurada, imaginada). En este trabajo se obtuvo la señal cerebral mediante un dispositivo EEG de 16 electrodos con una frecuencia de muestreo de 300 Hz. En el estudio se grabaron las señales de 6 sujetos. Los sujetos pronunciaban

la palabra mientras una pantalla negra aparecía en el monitor por 2 segundos. A las señales se les aplicó la STFT para obtener características de la señal incluyendo coeficientes delta, doble delta y la media de ambos. De las características obtenidas se aplicó un LDA para reducirlas. Para clasificar los vectores reducidos se utilizaron modelos ocultos de Markov (HMM), los cuales forman parte del software JANUS diseñado para hacer reconocimiento de habla. En este trabajo se muestra que las áreas del cerebro que más se activaron durante la pronunciación imaginada de las palabras fueron: Las áreas de la corteza motriz, de Broca y Wernicke. Este trabajo reporta que para 5 palabras un individuo obtiene 47.27% de exactitud en la clasificación promedio.

En el trabajo [82] se destaca el uso de la DTCWT (Dual-Tree Complex Wavelet Transform) y la RDWT (Redundant Discrete Wavelet Transform (RDWT) como métodos para la extracción de características. Después en el trabajo [14,15] se aplicaron métodos de extracción de características que se usan en reconocimiento de habla. En este trabajo se grabó la señal cerebral de 18 sujetos sanos usando 16 canales de un dispositivo EEG. Se realizó la grabación de 5 palabras del idioma inglés utilizadas en el ámbito militar (*alpha, bravo, charlie, delta, echo*), donde cada una se repitió 20 veces. Luego de obtener las señal cerebral se utilizó la DTCWT para hacer la extracción de 96 características el cual es reducido mediante un análisis discriminante lineal a 35 características. Después el vector es pasado por el software JANUS para reconocimiento de habla donde se utilizó un clasificador de modelos ocultos de Markov, obteniendo un 45.5% de exactitud promedio en la clasificación para los 18 sujetos.

En los trabajos [7,67,83] se hace la clasificación de 5 palabras del lenguaje español; 4 tienen que ver con ubicación física (*arriba, abajo, derecha, izquierda*) y una corresponde a una acción (*seleccionar*). En el trabajo [67] se realizó la grabación de la señal cerebral de 27 sujetos sanos los cuales pronunciaron de forma imaginada cada palabra 33 veces. Una vez obtenidas las señales se les aplicó un procesamiento usando la referencia promedio común de los 14 electrodos utilizados. Después se aplicó la transformada discreta wavelet para extraer coeficientes wavelets de la señal usando 5 niveles de descomposición. Luego los coeficientes wavelets de la señal son normalizados usando la energía relativa wavelet, con lo cual se obtuvo un conjunto de 5 valores de energía wavelet, 4 de los niveles de descomposición y uno de aproximación. Estos valores obtenidos representan las características extraídas que pasaron como entrada a clasificadores como máquina de vectores de soporte, Naive Baye o Random Forest. En este trabajo se clasificó con dos conjuntos distintos de canales 4 y 14, obteniendo 48.10% y 60.11% de exactitud promedio para los 27 sujetos.

Recientemente en [84] se clasifican dos caracteres monosilábicos del idioma chino,

donde cada carácter representa un significado (*izquierda* y *uno*). Se elijen dichos caracteres ya que tienen distintas formas y pronunciaciones, además de que son usados de manera frecuente diariamente. En esta investigación se grabó la señal cerebral a 8 sujetos. El estudio se dividió haciendo la grabación de 30 electrodos sobre toda la corteza cerebral en 2 sujetos y usando 15 electrodos sobre las áreas del modelo Geschwind-Wernicke para 6 sujetos. Cada sujeto hizo una única grabación donde durante 4 segundos se imaginaba repetidamente la pronunciación del carácter y por último se grabó un periodo de descanso para continuar con la grabación del otro carácter.

Una vez obtenidas las señales se procedió a filtrar las señales al rango de 6 a 30 Hz con un filtro paso-bajo. Después se aplicaron los patrones espaciales comunes (CSP) para extraer características las cuales fueron proyectadas a un sub-espacio usando filtros espaciales, con lo que se obtenían mapas topográficos de cada pronunciación. Para clasificar los caracteres se utilizó SVM, con lo que se obtuvo un porcentaje de clasificación promedio para los dos caracteres de 70.64 % usando los 30 canales con 2 sujetos y de 65.6 % usando los 15 canales con 6 sujetos.

En la tabla 3.2 se muestra un resumen de los trabajos relacionados con la clasificación automática de habla imaginada.

3.2. Sonificación de EEG para clasificación

La sonificación de EEG ha sido utilizada para fines diversos; para estudiar la distribución espacial de la actividad cerebral [85], para averiguar la coherencia del sonido con la actividad cerebral de la banda alfa [86] o para asistencia en rehabilitación [87]. En esta sección solo se muestran los trabajos recientes que han utilizado la sonificación de EEG para tareas de clasificación, incluyendo ambos casos, métodos automáticos y métodos no automáticos. En la tabla 3.2 se muestra un resumen de las características de dichos trabajos.

El trabajo [88] se tomó como base para aplicar el método de sonificación de EEG. En este trabajo se usó la sonificación de EEG para mostrar que se puede caracterizar una actividad mental en un audio. Este método genera audios basados en tonos (ondas sinusoidales con frecuencia audible) mediante la elección de amplitudes dominantes en las frecuencias de la señal cerebral. Al aplicar el algoritmo de este trabajo se aprecia que el oído de un ser humano podría catalogar de forma sencilla los ejemplos de sonificaciones de ruido, parpadeos, estado mental de cálculo o rotación de manos. En la figura 3.4 se muestran 4 ejemplos de los espectrogramas obtenidos, al utilizar

Tabla 3.2: Trabajos relacionados con la clasificación de habla imaginada.

Trabajo (año)	Canales EEG	Enfoque (Idioma)	Clases	Num. Suj.	Proc.	Extracción de caract.	Clasificación	Exactitud Promedio
[75], (2009)	64	Vocales (Inglés)	2	3	CAR	CSP	SVM no lineal	68.0 - 78.0 %
[76], (2009)	118	Sílabas (Inglés)	3	4	LP Filter	No aplica	Matching filters	52.25 - 74.25 %
[80], (2010)	118	Sílabas (Inglés)	2	7	CAR, ICA	Coef. auto regresivos	KNN	61.0 %
[81], (2011)	52	Fonemas (Inglés)	10	5	CAR	No aplica	LDA, NB	66.4 - 76 %
[77], (1997)	16	Palabras (Inglés)	7	5	CAR	No aplica	Least squares	52.57 %
[15], (2009)	16	Palabras (Inglés)	5	18	No	DTCWT	HMM	45.50 %
[16], (2006)	16	Palabras (Inglés)	5	6	No	STFT	HMM	47.27 %
[67], (2013)	4,14	Palabras (Español)	5	27	CAR	DWT	SVM, NB, RF	48.1 - 60.1 %
[84], (2013)	15,30	Palabras (Chino)	2	2-6	LP Filter	CSP	SVM	65.6 - 70.64 %

el algoritmo de sonificación de EEG para 4 tareas distintas, correspondientes a 10 segundos de grabación de señal EEG.

En [13], el neurofeedback se genera a partir de electroencefalogramas (EEG) de sujetos sanos y sujetos con la enfermedad de Alzheimer. La retroalimentación audible del EEG de los pacientes con Alzheimer difieren sustancialmente de la retroalimentación audible de pacientes saludables, con esto se generan nuevas herramientas potenciales de diagnóstico, pues se aplica una técnica de sonificación de EEG por audificación para hacer clasificación entre pacientes sanos y pacientes con una etapa temprana de la enfermedad. Dicha clasificación fue realizada por personas no especializadas, las cuales lograron reconocer la diferencia escuchando la sonificación entre sujetos sanos y sujetos no sanos con un 95 % de exactitud. Adicionalmente en este trabajo también se aplicó un método de automático mediante un análisis discriminante lineal para clasificar los mismos datos y compararlos con los obtenidos por los voluntarios, usando LDA se obtuvo una exactitud promedio del 90 %.

En [55] se propone una retroalimentación auditiva para aplicarla en una interfaz cerebro computadora (BCI). Se basa en la sonificación de los estados mentales de

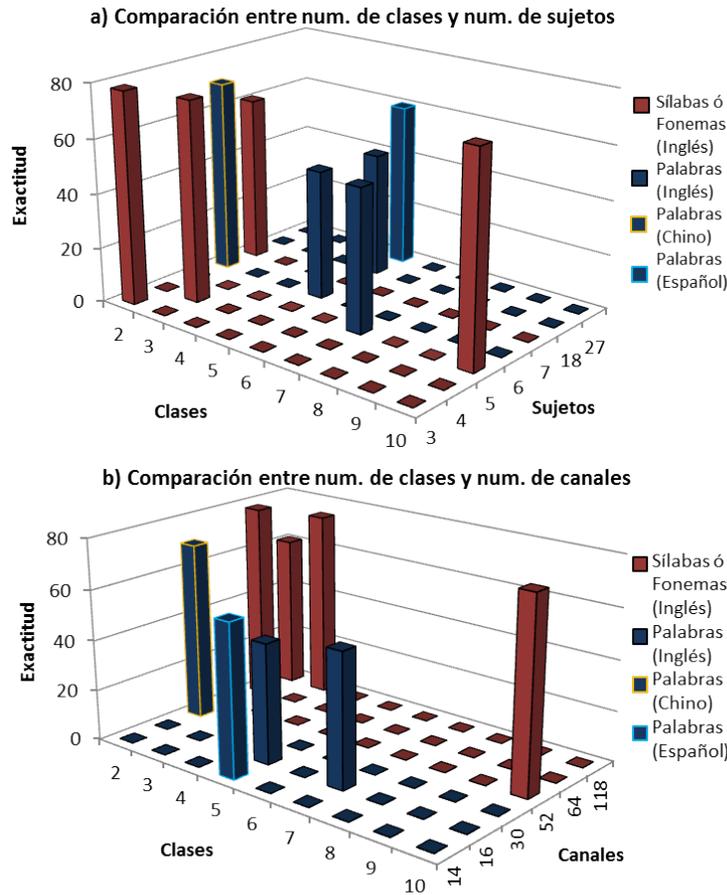


Figura 3.3: Comparación de exactitudes entre num. de clases y num. de sujetos (Arriba). Comparación de exactitudes entre num. de clases y num. de canales (Abajo).

los seres humanos capturados por electroencefalograma. Se utilizan dos técnicas de descomposición de la señal en tiempo y frecuencia, además de la técnica Empiric Descomposición Modeling (EMD) para asignar los registros de EEG en las partituras musicales. Esta retroalimentación auditiva puede ser usada en alguna BCI al poder representar de manera simple los potenciales corticales lentos en el dominio del audio. En dicho trabajo también se muestra la sonificación basada en modelado de elipsoides (Bump modeling sonficiation), la cual se obtiene aplicando la Wavelet Morlet compleja (Complex Morlet Wavelet) a la señal cerebral para aproximar los periodos de activación cerebral a elipsoides. Posteriormente siguiendo un mapeo de parámetros se crearon archivos de audio basados en notas musicales *MIDI*. En este trabajo no se usaron los archivos de audio para clasificar simplemente se obtuvieron atributos de los modelos empírico y del basado en elipsoides para clasificar. Mediante una Red Neuronal

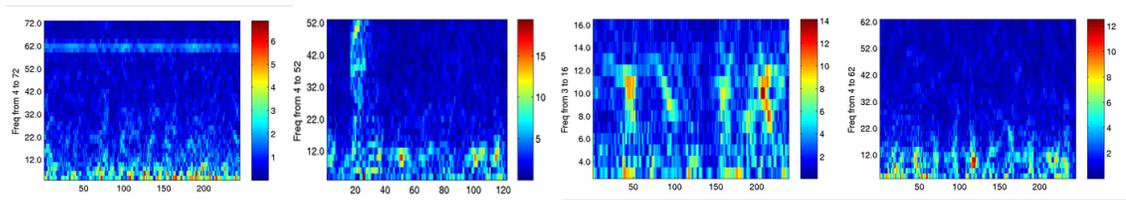


Figura 3.4: Espectrogramas generados por el algoritmo de sonificación *EEG a tonos*. De izquierda a derecha espectrograma para: Ruido, Parpadeo, Estado mental de cálculo, Rotación de manos (figuras reproducidas de [88]).

Artificial se clasificó automáticamente a los pacientes obteniendo porcentajes entre el 80 y 93 % de exactitud.

Se ha investigado usar la sonificación de EEG para apoyar en la identificación de episodios de pacientes con epilepsia [10]. Para ello se generaron ejemplos de sonido con ritmos de trastornos epilépticos conocidos y se encontraron objetos auditivos rítmicos estereotipados en un solo canal. Los ritmos fueron llevados a audio en tiempo real usando la técnica de sonificación basada en eventos (Event-based sonification), con lo que se fue capaz de separar los ritmos de actividad de fondo para permitir que el oyente pueda percibir su frecuencia, duración e intensidad mientras se monitorea. Aunque no se aplica algún método de extracción de características o de clasificación automática, en este trabajo se muestra que una persona puede escuchar en tiempo real las variaciones características de cada sujeto al estar percibiendo un episodio de epilepsia.

Como continuación a la investigación previa, el mismo autor propone obtener características de la señal de EEG para caracterizar audios de patologías humanas como irregularidades en el sueño o la ocurrencia de episodios epilépticos [11]. En este trabajo se introduce una técnica de sonificación que busca construir audios por medio de vocalizaciones (Vocal Sonification). Esta técnica de sonificación de EEG utiliza un modelo articulatorio de voz y la asignación de parámetros especialmente seleccionada para obtener objetos auditivos característicos, los cuales corresponden a las características de las señales de EEG multivariantes. En este trabajo se usan técnicas como análisis de componentes principales (PCA) para identificar fuentes anormales en la señal. En este trabajo no se realiza clasificación automática, pero se demuestra que usando la sonificación de EEG vocal, se puede tener una alternativa para identificar rasgos anómalos en la señal cerebral. A continuación en la tabla 3.3 se muestra el resumen del estado del arte en cuanto la sonificación de EEG para propósitos de clasificación.

Tabla 3.3: Trabajos relacionados con la sonificación de EEG usada para clasificación.

Trabajo	Utilizado en	Método de sonificación	Extracción de características	Método de Clasificación
C. Anderson, 2002 [88]	Estado de cálculo, Mov. Manos	EEG a tonos	(no aplica)	(no aplica)
T. Hermann, 2007 [10]	Episodios de epilepsia	Sonificación basada en eventos	(no aplica)	(no aplica)
M. Elgendi et. Al, 2013 [12]	Diagnóstico de Alzheimer	Audificación	(no aplica)	LDA 90 %
T. Rutkowski, F. Vialatte, 2006 [55]	Detección de Alzheimer	Sonificación por modelado de elipsoides	Complex Morlet Wavelet	ANN (80-93 %)
T. Hermann, G. Baier, 2006 [11]	Estado de sueño, pacientes con epilepsia	Sonificación de EEG Vocal	PCA	(No automático)

3.3. Análisis de los trabajos relacionados

Los trabajos que buscan clasificar ya sean sílabas, fonemas o palabras bajan sus porcentajes de exactitud conforme el vocabulario aumenta, en este sentido los trabajos donde solo existe la clasificación de dos categorías de sílabas, aunque tengan buenos porcentajes de exactitud, se tiene que demostrar que el método continuará teniendo un desempeño similar al aumentar el número de sílabas a clasificar (ver figura 3.3). En la tabla 3.2 se muestra que el número de electrodos empleados para la lectura influye en la clasificación; con más electrodos se obtiene información de más áreas del cerebro pero con demasiados electrodos podría haber redundancia de información. Se conjetura que no cada sujeto podría tener un número máximo de electrodos con los que logré la mejor clasificación (ver figura 3.3b). También se sugiere que aumentar el número de sujetos regularmente hará disminuir los porcentajes de exactitud promedio, al existir más variabilidades en las señales (ver figura 3.3a). En los trabajos [76,84] se usa un filtro para remover señales arriba de los 45 Hz, pero se desconoce que frecuencias involucran a la pronunciación imaginada de vocales y esta remoción podría implicar pérdida de señal útil. Los trabajos [15, 16] podrían ser comparables con el enfoque propuesto, pues tienen 5 palabras a clasificar y el número de canales es 16, solo dos más a los empleados en esta investigación.

En la mayoría de los trabajos descritos previamente se experimenta con pocos suje-

tos, por lo que se tendría que demostrar que esos métodos mantendrán los porcentajes de clasificación cuando se busca expandir y generalizar el esquema implementado. La base de datos utilizada en este trabajo se creó considerando tener más sujetos (27 sujetos) que los reportados en el estado de arte (máximo 18 sujetos).

En los trabajos de [14, 16, 89] se aplican métodos de extracción de características que funcionan en la tarea de reconocimiento de voz, pero se procesó la señal original de EEG, la cual es de distinta naturaleza a la de la voz humana. Por ejemplo la voz se concentra en el rango de 500 Hz a 4 KHz, mientras que la señal de EEG tiene un rango de frecuencias más reducido (1~70 Hz). Es por ello que en este trabajo de tesis se utilizarán los métodos de extracción de características sobre los audios resultantes de la sonificación de EEG.

En cuanto a el uso de la sonificación de EEG usada para tareas de clasificación existen muy pocos trabajos. Como se mostró en la descripción de los trabajos relacionados esta técnica se ha usado para diversas finalidades donde en el caso de los trastornos de epilepsia permitió acelerar la inspección de los datos, identificar características de los sucesos independientes de los sujetos y la extracción de características para identificar anomalías. También se muestra que en donde se ha aplicado algún método de clasificación automática los resultados son alentadores y que si el oído humano fue capaz de discriminar entre distintas clases al escuchar los audios, también los podría hacer un método automático. Por lo que se sugiere que aplicar la sonificación de EEG en permitirá utilizar métodos de extracción de características comúnmente utilizados en procesamiento de audio y podría ayudar a mejorar la clasificación de palabras no pronunciadas.

De acuerdo con lo mencionado anteriormente, el trabajo de esta tesis se encuentra ubicado dentro de los trabajos que buscan clasificar palabras del habla imaginada, usando la técnica de sonificación de EEG como paso previo a la clasificación, pues se usará la señal de audio proveniente de la sonificación de EEG para la extracción de características.

METODOLOGÍA

4.1. Descripción General

En este capítulo se mostrarán los pasos seguidos para realizar la clasificación de palabras no pronunciadas utilizando la sonificación de EEG. La investigación siguió la metodología mostrada en la figura 4.1. Estas etapas, a excepción de la etapa de sonificación, son las comúnmente utilizadas en procesos de clasificación. En la presente investigación se comienza con una base de datos de palabras no pronunciadas representadas en vectores de datos, los cuales serán procesados y luego sonificados para ayudar a resaltar detalles relevantes.

La base de datos fue adquirida en el trabajo [4] y contiene la señal cerebral de cada sujeto correspondiente a la pronunciación imaginada de las 5 palabras. Más adelante en esta sección se describe con mas detalle dicha base de datos y el proceso de adquisición. Después se hace el montaje de referencia de la señal cerebral, en la cual se aplica le referencia promedio común. Luego las señales se sonifican usando el método de sonificación *EEG to Tones*. Después se aplican los métodos para extraer las características de los audios resultantes de la sonificación, en este caso DWT y MFCC. Posteriormente se procede a clasificar los vectores de características de las 5 palabras usando Random Forest.

4.2. Dispositivo EEG utilizado

Las señales cerebrales utilizadas en este trabajo fueron registradas utilizando el dispositivo de EEG Epoc de la compañía Emotiv, el cual se muestra en la figura 2.4. El dispositivo es inalámbrico cuenta con 16 electrodos; 14 de lectura de alta resolución AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4 y 2 electrodos de referencia CMS (*Common Mode Signal*) y DRL (*Driven Right Leg*) en las posiciones P3 y P4,

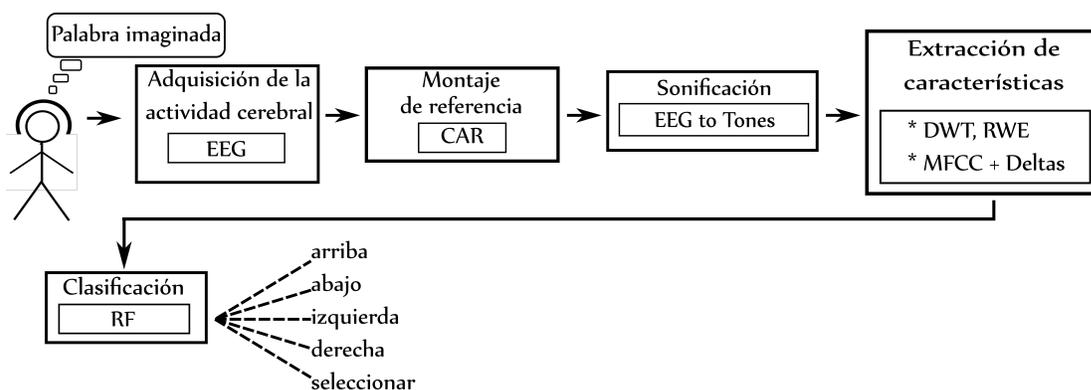


Figura 4.1: Metodología seguida para la clasificación de palabras no pronunciadas usando sonificación de EEG.

respectivamente), los electrodos tienen una frecuencia de muestreo de 128 Hz. En la figura 4.2 se muestra la distribución de los electrodos en el dispositivo de EEG. Este dispositivo registra voltajes con una resolución de $0.51 \mu\text{V}$ [90].

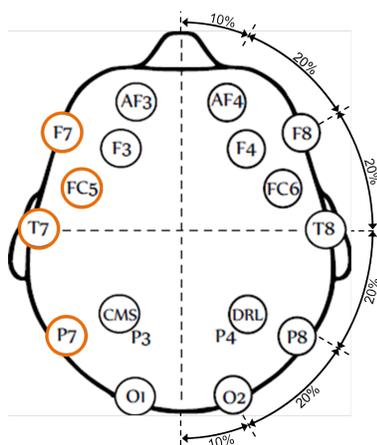


Figura 4.2: Localización de los electrodos en el kit EPOC de Emotiv. Se resaltan en naranja los 4 electrodos seleccionados para representar las áreas que se activan durante el procesamiento de lenguaje según el modelo Geschwind-Wernicke.

4.3. Adquisición de la señal cerebral

La señal cerebral empleada es la misma que se adquirió en el trabajo [4]. En esta base de datos se grabó la señal cerebral de 14 canales de 27 sujetos sanos con el dispositivo de EEG de la compañía Emotiv, *EpoC*. Cada sujeto imaginó la pronunciación de cada una de las 5 palabras 33 veces, se pensó en ese número de repeticiones para

obtener un mejor promedio. El proceso para delimitar el espacio de tiempo que incluye una palabra fue realizado mediante el uso de clics por parte del usuario que funcionan como marcadores dentro de la grabación de la señal, tal como se aprecia en la figura 4.3. Las sesiones de grabación fueron realizadas en el Instituto nacional de astrofísica, óptica y electrónica (INAOE), en un sala libre de distracciones y ruido externo, esto para evitar que la señal cerebral tuviera información no deseada que pudiera afectar el proceso de clasificación.

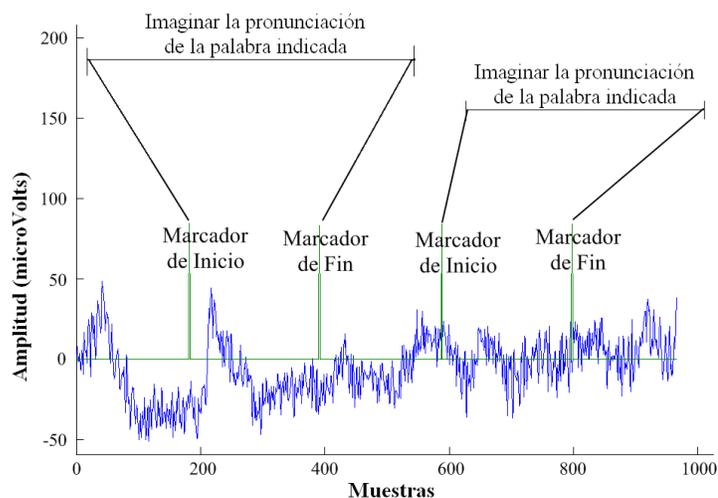


Figura 4.3: Señal de EEG del canal F7 mientras imagina la dicción de la palabra *Abajo* (figura reproducida de [4])

4.4. Montaje de referencia

Al realizar la grabación de EEG, las variaciones de voltaje registradas de cada electrodo incluyen ruido. Para cada instante de tiempo el conjunto de electrodos tiene ligeras variaciones en el valor capturado con respecto a un valor base capturado por los todos electrodos. Las señales de EEG obtenidas son procesadas con el método de referencia promedio común (CAR) (tal como se aplicó en los trabajos relacionados anteriormente descritos [67, 81, 91]). Este montaje mejora la relación señal a ruido de la señal de EEG. La referencia promedio común busca quitar la información que es común en todas las lecturas simultáneas de los electrodos [92]. La CAR puede ser calculada mediante la resta del potencial entre cada electrodo y la referencia (el potencial promedio de todos los canales), se repite esto para cada instante de tiempo

en la frecuencia de muestreo, tal como se muestra en la siguiente fórmula:

$$V_i^{CAR} = V_i^{ER} - \frac{1}{n} \sum_{j=1}^n V_j^{ER} \quad (4.1)$$

donde V_i^{ER} es el potencial entre el i -ésimo electrodo y la referencia, y n es el número de electrodos en el montaje [67]. Existen otros montajes como la referencia respecto a orejas o el laplaciano, comúnmente usado para tareas de movimientos imaginados [93]. Se decidió utilizar CAR dado que no se busca encontrar diferencias entre electrodos de una zona específica usando por ejemplo el montaje laplaciano, sino resaltar las diferencias de cada electrodo con respecto a lo que es común [94].

4.5. Sonificación de la señal EEG

En el presente trabajo se utilizó la técnica de Sonificación *EEG to tones* (*EEG a tonos*) [88]. Esta técnica es un proceso de sonificación, en la cual se escalan las frecuencias de EEG a frecuencias audibles seleccionando las frecuencias dominantes.

El algoritmo de sonificación *EEG to tones*, primero toma de entrada una señal de EEG de 1 canal. Luego se crea un espectrograma de la señal EEG, según los parámetros establecidos: Rango de frecuencias en la señal EEG, tamaño de la ventana y el tamaño del traslape entre ventanas. El rango de frecuencias es establecido de acuerdo al valor mínimo y valor máximo de frecuencias de la señal de EEG a tomar en cuenta. El tamaño de la ventana indica cuantas muestras de la señal de EEG se tomarán para formar sub-conjuntos de la señal. Finalmente el tamaño del traslape entre ventanas indica cuantas muestras incluirá el siguiente sub-conjunto del sub-conjunto anterior, por lo que el tamaño de traslape es menor al tamaño de la ventana. Ya con los subconjuntos formados, se procede a aplicar la transformada rápida de Fourier (*FFT*) a cada uno con una definición de al menos el doble de la frecuencia de muestreo (teorema de Nyquist [95]), en nuestro caso si la frecuencia de muestreo es 128 usaremos 256 para la FFT. Al aplicar la FFT obtenemos valores que representan la energía contenida en cada frecuencia de la señal de EEG para cada segmento de tiempo a los que también se conocen como amplitudes. Las amplitudes según el detalle de la FFT quedan agrupadas por bloques de frecuencias (en nuestro caso por bloques de 4).

Una vez obtenido el espectrograma de la señal de EEG, las amplitudes se escalan de acuerdo a la mayor amplitud, es decir se dividen los valores de las amplitudes entre la máxima amplitud y así poder representar cada amplitud en valores que indican el porcentaje de amplitud respecto a la máxima amplitud (con valores entre 0 y 1). Las amplitudes escaladas de todo el espectrograma se ordenan de manera descendente y

Algorithm 2 Sonificación de EEG a tonos

```

Entradas: eeg //Señal EEG de un canal
fs //Frec. de muestreo EEG ntn //Número de tonos
fl //Frec. mínima en EEG fh //Frec. máxima en EEG
w //Longitud de ventana sf //Traslape de ventana
dur //Duración de los tonos fla //Frec. mínima del audio
fha //Frec. máxima del audio fsa //Frec. de muestreo del audio
Salida: audio //Sonificación de la señal EEG
//Obtener matriz de espectrograma del EEG
col=1;
for i = 1 hasta longitud(eeg)-w con paso sf do
    spec_col = sfft(eegi:i+w, fl, fh, fs)
    col = col + 1
end for
maxamp = máximo(spec)
//Obtener las frecuencias de mayor amplitud
for c = 1 hasta columnas(spec) do
    //Escalar amplitudes entre la máxima amplitud
    for r = 1 hasta renglones(spec) do
        specr,c = specr,c/maxamp
    end for
    //Ordenar amplitudes descendientemente
    ordc = ordendesc(specc)
    //Tomar las frec. dominantes por columna
    fdac = ord1:ntn,c
    //Escalar frecuencias únicas de fda a un tono audible
    for r = 1 hasta renglones(fda) do
        if unica(fdar,c) then
            ftn = [fdar,c - fl]/(fh - fl) * (fha - fla) + fla
            tnfdar,c = sin(2π/fsa * ftni * (fsa * dur))
        end if
    end for
end for
//Crear audio con la suma de las señales en tonos
for c = 1 hasta columnas(fda) do
    for r = 1 hasta renglones(fda) do
        sumtn = sumtn + tnfdar,c
    end for
    audio = audio + sumtn
end for
return audio

```

se toman las de mayor valor. El número de amplitudes a tomar es el establecido por el parámetro de número de tonos a representar en el audio. Con lo anterior se crea un conjunto de bloques de frecuencia dominantes por cada segmento de tiempo.

A continuación se establece que tono le corresponde a cada bloque de frecuencias dominantes del espectrograma. Para ello se establece la mínima y máxima frecuencia posible a representar en el archivo de audio, luego a partir de ese rango se distribuyen proporcionalmente los tonos respecto al rango de frecuencias de la señal de EEG. Para hacer la distribución de los tonos se escala en el audio cada frecuencia dominante del EEG, con lo que se obtiene la separación que existirá entre cada tono (4.2).

$$FTono = \left(\frac{frecEEG - fminEEG}{fmaxEEG - fminEEG} \right) * (fmaxAudio - fminAudio) + fminAudio \quad (4.2)$$

Ya que se conoce la distribución de los tonos, se crea una onda sinusoidal que representará cada tono en el archivo de audio.

$$Tono = \sin \left(\frac{2 * \pi}{Fs * FTono * (Fs * segundos)} \right) \quad (4.3)$$

Por último, se forma el audio de salida. Para ello se recorre cada ventana del espectrograma y se verifica si existen frecuencias dominantes. De ser así, se asignan los tonos correspondientes con la duración previamente establecida en segundos. En otras palabras, se suman las señales sinusoidales de los tonos para cada segmento de tiempo.

El pseudocódigo que explica el método de sonificación de EEG utilizado es mostrado en el algoritmo 2.

En la figura 4.4 se muestran un ejemplo del resultado de la ejecución del algoritmo con dos configuraciones distintas, en el cual se muestra que al aplicar la sonificación de EEG podemos expandir o contraer la señal a partir de la modificación del parámetro de entrada de duración de tonos.

Si de la señal de audio generada por el algoritmo de sonificación obtenemos la densidad espectral de su energía, podremos apreciar la distribución de los tonos y la intensidad de la energía de cada uno de ellos dentro de la señal de audio, la cual se puede apreciar en la figura 4.5.

La señal de audio generada ahora posee otras características en la distribución de las frecuencias como se puede apreciar en el espectrograma calculado a partir de las señales de audio de ejemplo de la figura 4.4. En los espectrogramas de la figura 4.6, se aprecia el impacto que tiene la cantidad de tonos en el audio de salida, pues para una sonificación con 4 tonos las frecuencias dominantes estarán con poca separación, mientras que con 14 tonos las frecuencias están mejor distribuidas.

El proceso de sonificación de EEG previamente descrito se efectúa para las 33 repeticiones en los 14 canales de las 5 palabras del vocabulario, esto para cada uno

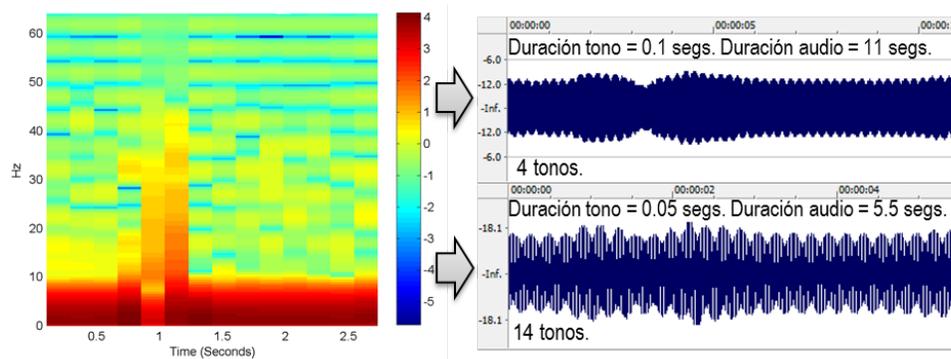


Figura 4.4: A la izquierda se muestra el espectrograma de una repetición de la palabra arriba. A la derecha el audio resultante de la sonificación de EEG con distinta duración de tono y distinto número de tonos.

de los 27 sujetos del estudio. Se pueden escuchar ejemplos de la sonificación de EEG en [96]

4.6. Extracción de características

Después de aplicar el proceso de sonificación de EEG a las palabras no pronunciadas, la siguiente etapa es buscar la manera de extraer características de la señal de audio con el objetivo de que un clasificador pueda discriminar entre clases.

Las características utilizadas en las BCI son no estacionarias [97]. Además, estas características deben contener información del tiempo debido a que los patrones de actividad cerebral están generalmente relacionados a variaciones específicas del EEG en el tiempo. Por lo que es necesaria una representación que considere dichos aspectos.

La transformada wavelet discreta (Discret Wavelet Transform) es una técnica que permite modelar dichas variaciones en el dominio tiempo-escala. La DWT provee una representación wavelet altamente eficiente mediante la restricción de la variación en la traslación y la escala, usualmente a potencias de dos. Cuando es así, la DWT es llamada transformada wavelet diádica.

4.6.1. Usando DWT y Energía Relativa

Una vez aplicada la DWT sobre la señal se obtienen coeficientes de aproximación y de detalle, desde los cuales es posible calcular la energía relativa wavelet [67]. La energía relativa provee información para caracterizar la distribución de energía de la señal en diferentes bandas de frecuencia, con lo que se obtiene una independencia del tamaño de la ventana de señal de EEG o de audio, según sea el caso.

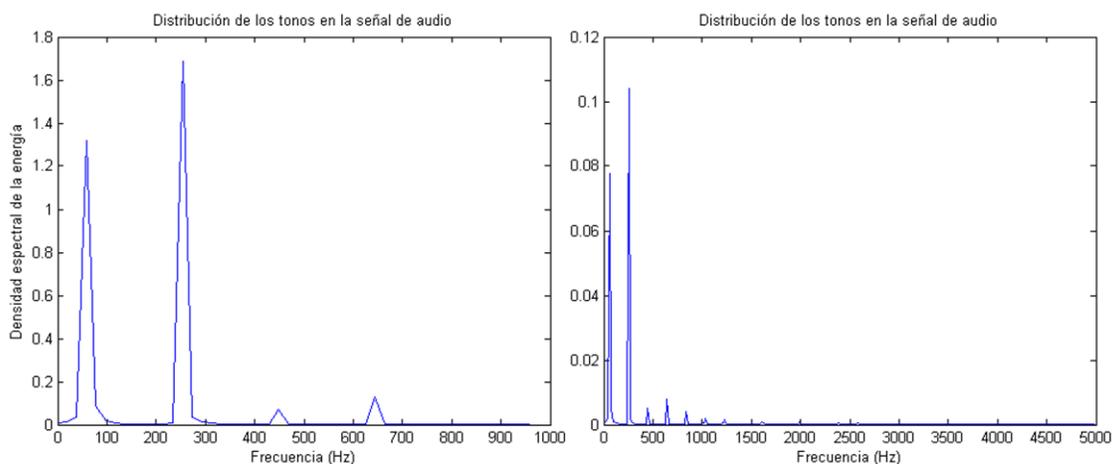


Figura 4.5: Distribución e intensidad de la energía de los tonos de las señales de audio de la figura 4.4. Se muestra la distribución de los 4 (izquierda) y 14 tonos (derecha) a lo largo del rango audible predefinido.

A partir de la descripción anterior, en las señales EEG sonificadas se determinó usar 7 valores que representan la energía wavelet en todos los niveles de descomposición y el último de aproximación (D1-D6 y A6).

Por otra parte, en las señales de EEG no sonificadas, cada ventana de habla imaginada se representa mediante un conjunto de 5 valores de energía wavelet, 4 de los niveles de descomposición y uno de aproximación (D2-D5 y A5) con respecto a la energía wavelet total. Tal como se realizó en [67], el valor asociado con D1 es descartado.

En la figura 4.7 se muestran los niveles de descomposición y detalle obtenidos de la transformada Wavelet, donde cada nivel representa un rango de frecuencias en la señal de EEG o de audio. Para la señal de EEG se calcularon los niveles a partir del rango de frecuencias de 0 a 60 Hz, mientras que en la señal de EEG sonificado los niveles van de 50 a 5000 Hz.

En el proceso de extracción de características se siguió el mismo enfoque que en la selección de parámetros para la sonificación, se realizaron simulaciones con distintos wavelets y niveles de descomposición de diferentes Wavelets Daubechies (db2, db4, db6 y db20) con el objetivo de obtener los mejores resultados de clasificación. En el proceso de clasificación de las palabras imaginadas usando sonificación de EEG, los niveles de descomposición se calculan en base a los archivos de audio generados, por lo que el rango de frecuencias es diferente al de las señal de EEG sin sonificar. Entonces los niveles de descomposición, a diferencia de [67], evaluarán todo el espectro

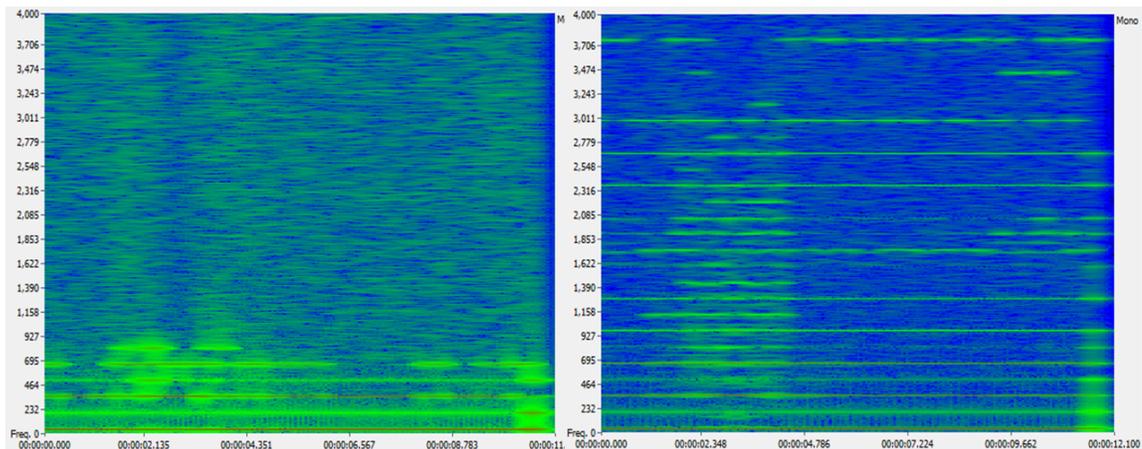


Figura 4.6: Espectrogramas del resultado de la sonificación de EEG para los ejemplos del la figura 4.4 correspondientes a 4 tonos (izquierda) y 14 tonos (derecha). El audio con 14 tonos, tiene una mejor distribución que la de 4 tonos.

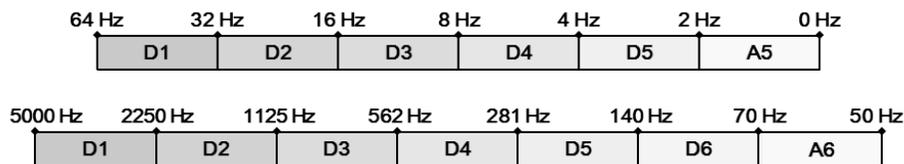


Figura 4.7: Descomposición en niveles de la transformada Wavelet. Cada nivel representa un rango de frecuencias en la señal de EEG (arriba) o en la señal de EEG sonificado (abajo).

de frecuencias sin descartar ningún nivel de descomposición, pues en nuestro caso se desconoce si la señal de la palabra imaginada dentro del audio se encuentra también solo en un cierto rango de frecuencias específicas.

4.6.2. Usando MFCC

En esta técnica se creó un espectrograma basados en los parámetros de tamaño de ventana y traslape entre ventanas. Luego a partir del espectrograma se establecen filtros en las frecuencias Mel por medio de funciones triangulares. A partir de las frecuencias filtradas el algoritmo obtiene coeficientes cepstrales que indican la energía en las distintas bandas del espectro. Con lo anterior se obtiene una matriz de coeficientes cepstrales por cada segmento de tiempo.

Los coeficientes cepstrales totales obtenidos son de grandes dimensiones por lo que es conveniente hacer reducciones o representar los valores con sus atributos estadísticos. En nuestro caso se optó por representar los coeficientes con los atributos

estadísticos:

- Valor máximo
- Valor mínimo
- Media
- Desviación estándar

También se implementó un enfoque para permitir extraer los cambios diferenciales entre segmentos de tiempo los cuales son conocidos como valores delta y los cambios entre valores *delta*, son conocidos como valores *doble delta*. Para obtener los valores delta se tiene que establecer una ventana la cual indicará cuantas muestras vecinas se tomaran para calcular los cambios en el tiempo [98]. La ventana es representada por coeficientes los cuales representan un tiempo anterior y un tiempo posterior de la muestra central. Estos coeficientes se definen como: $C = \{k, k - 1, k - 2, \dots, -k\}$ Donde k es el tamaño de la ventana. Se repiten las muestra inicial y final k veces como ajuste. Después para calcular la derivada d_i en el tiempo de cada muestra, se recorre el vector aplicando una operación filtro usando los coeficientes calculados:

$$\Delta_i = C_1 X_n + C_2 X_{n-1} + \dots + C_{|C|} X_{n-|C|} \quad (4.4)$$

Donde X_n es la muestra en la posición final de la ventana. Una vez calculadas las derivadas para cada muestra del vector se divide cada valor entre la suma de los cuadrados de los coeficientes C .

$$\Delta F_i = \frac{\Delta_i}{\sum_{n=1}^{|C|} C_n^2} \quad (4.5)$$

Al final se quitan las muestras de ajuste agregadas. Los valores delta representan a un segmento de MFCC por lo que tienen que ser representados igualmente en forma de atributos estadísticos. Por otra parte, los valores doble delta se calculan con el mismo proceso que los coeficientes delta, pero con los coeficientes delta como vector de entrada y también son representados con sus respectivos atributos estadísticos.

4.7. Clasificación

Como etapa final se aplica un proceso de clasificación automática para los vectores de características provenientes de los 4 enfoques (EEG usando DWT, EEG con sonificación usando DWT, EEG usando MFCC y EEG con sonificación usando MFCC). La

clasificación se realizó bajo dos enfoques de selección de canales. Un enfoque clasifica solo 4 canales correspondientes al modelo Geschwind-Wernicke correspondientes al hemisferio izquierdo en las áreas que involucran el procesamiento del lenguaje. Los electrodos seleccionados se muestran en la figura 4.2. El otro enfoque buscará clasificar usando todos los 14 canales disponibles, esto con la finalidad de evaluar si un conjunto reducido de canales permite mantener o mejorar los porcentajes de exactitud en la clasificación, con lo que se podría reducir los datos a procesar.

Se utilizó el clasificador Random Forest. Se eligió RF porque obtuvo los mejores porcentajes de clasificación en palabras imaginadas en los estudios previos [7,67] comparado con otros clasificadores como Naive Bayes o SVM. Además que este clasificador tiene un buen desempeño en la clasificación de bases de datos con muchos atributos y bases de datos con ruido [72], tal como es nuestro caso, pues se puede llegar a tener un gran número de características extraídas de los audios al usar DWT o MFCC.

En el presente trabajo se utilizaron los siguientes hiper-parámetros para la implementación del clasificador Random Forest: el número de árboles es 50, el número de atributos considerados en cada nodo es $\log_2(\text{numeroCaracteristicas}) + 1$ y el árbol base es un random tree. El esquema de clasificación está basado en un enfoque de validación cruzada [99] el cual hace particiones sobre el conjunto de datos de entrada. Este enfoque divide los datos en 10 partes iguales de forma aleatoria para después usar 9 de ellos para entrenar el modelo y se usan los datos de los pliegues restante como prueba, se repite este enfoque 10 veces para varias particiones aleatorias. La exactitud de clasificación de cada prueba por cada una de las 10 repeticiones se promedian para generar el valor de exactitud promedio que se reporta como el porcentaje de clasificación que obtuvo ese conjunto de datos. La partición del conjunto de datos para formar los 10 pliegues es aleatoria y esta cambia cuando se cambia la semilla del algoritmo de generación aleatoria de particiones. En nuestro caso para mantener las comparaciones de manera justa, se ejecutó la partición por pliegues bajo la misma semilla para el generador aleatorio de particiones. Así entonces se establece una semilla igual a 1 para la etapa de partición, con lo que el resultado sabemos no cambiarán.

EXPERIMENTACIÓN Y EVALUACIÓN

En este capítulo se describen los diversos experimentos y simulaciones realizados en la investigación, así como también la descripción y evaluación de los resultados.

Los procesos de sonificación de EEG, de extracción de características usando la DWT y de extracción de características usando MFCC, tienen diversos parámetros que afectan al desempeño en la clasificación, por lo que se realizó una búsqueda empírica para elegir configuraciones que se aproximen a las mejores. Las simulaciones fueron realizadas para buscar una tendencia en el comportamiento de los parámetros. Para la búsqueda se tomó un sub-conjunto de 3 sujetos de prueba (S8,S11,S20), los cuales son los que obtuvieron los mejores porcentajes de clasificación en el trabajo [67].

5.1. Descripción de los datos

La base de datos a utilizar contiene por cada uno de los 27 individuos 5 matrices de datos correspondientes a cada palabra (en el orden *arriba*, *abajo*, *derecha*, *izquierda*, *seleccionar*). Cada columna de la matriz representa la lectura de 1 canal de EEG y los renglones son los valores de voltaje de cada muestra. Una vez aplicada la referencia promedio común los valores oscilan entre -1000 y 1000 microvolts. Mientras que el rango de duración de cada palabra imaginada varía entre 0.8 a 3.2 segundos.

5.2. Configuración de parámetros del método de Sonificación de EEG

El algoritmo de sonificación de *EEG a tonos* posee distintos parámetros los cuales afectan las características del audio que se genera. En consecuencia, es conveniente evaluar su comportamiento para seleccionar los mejores parámetros. En este experimento se mostró el comportamiento de los parámetros: Número de tonos, duración de los

tonos, frecuencia de muestreo en el audio, rango de frecuencia en el audio, muestras por ventana y muestras para traslape entre ventanas. Los resultados mostrados en cada gráfica son los porcentajes de exactitud promedio en la clasificación de las palabras al hacer validación cruzada con 10 pliegues.

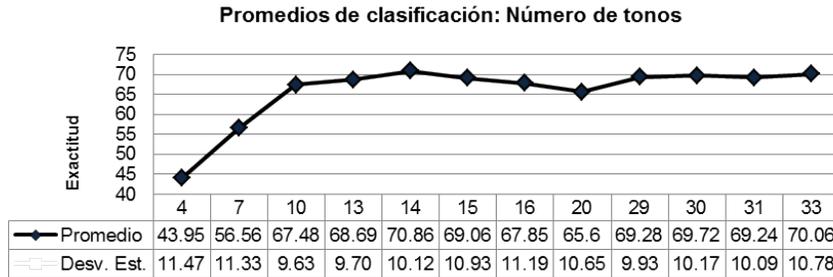


Figura 5.1: Promedios de exactitud en la clasificación con distinto número de tonos.

El número de tonos en el proceso de sonificación indica el número de frecuencias dominantes del espectrograma de EEG que estarán representadas en la señal de audio EEG. El número de tonos afecta también a como se percibe por el oído. Por ejemplo, un audio con pocos tonos tendrá pocas variaciones en tipos de sonido al escucharla y será más difícil encontrar detalles en una frecuencia en específico. En la gráfica 5.1 se muestran los resultados de clasificación de los 3 sujetos usando distinta cantidad de tonos en la señal de audio. Cuando hay menos de 10 tonos aproximadamente en la señal de audio los resultados de clasificación son bajos respecto a cuando se eligen 10 o más. Al usar 14 tonos se obtuvo el mejor promedio de clasificación, esto nos indica que no necesariamente una sonificación de EEG con gran cantidad de tonos obtiene los mejores resultados de clasificación.

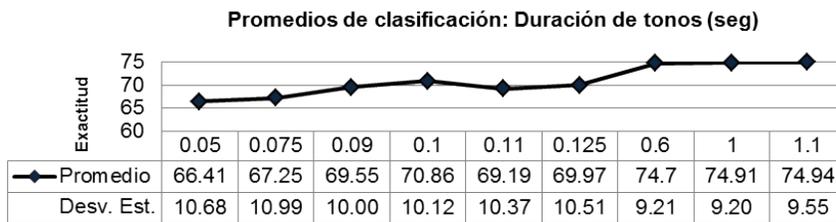


Figura 5.2: Promedios de exactitud en la clasificación con distinta duración (en segundos) de los tonos.

Como ya se mencionó anteriormente la duración de los tonos ayuda a modificar la duración del audio de salida del algoritmo de sonificación. En la figura 5.2 se muestra que una duración de tonos baja no ayuda al proceso de clasificación, mientras que

conforme se va aumentando la duración dicha clasificación también aumenta. Cabe señalar que la duración de los tonos afecta directamente al peso del archivo de audio almacenado, por lo que el incremento en este parámetro está limitado a los recursos disponibles.

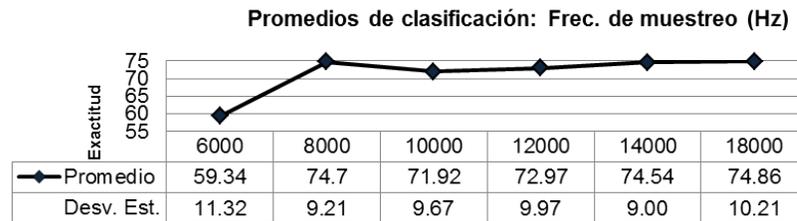


Figura 5.3: Promedios de exactitud en la clasificación con distinta frecuencia de muestreo para el audio de salida.

Otro aspecto evaluado fue la frecuencia de muestreo del archivo de audio generado por la sonificación de EEG. La frecuencia de muestreo afecta el detalle en que es discretizada la señal. Es de esperarse que a mayor frecuencia de muestreo los porcentajes de clasificación sean mejores. En la gráfica 5.3 se muestra que para frecuencias de 8 KHz en adelante los promedios de clasificación se mantienen aumentando en pequeñas cantidades. La frecuencia de muestreo al igual que la duración de los tonos, afecta al peso que tendrá el archivo de audio generado pues se están almacenando más muestras por segundo de la señal.

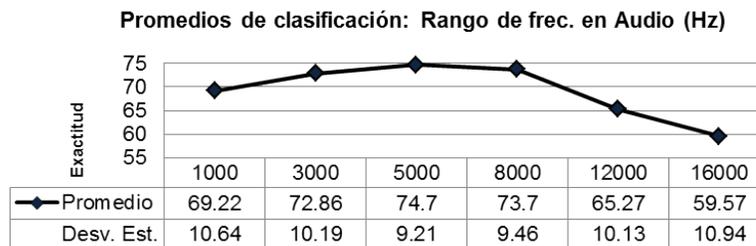


Figura 5.4: Promedios de exactitud en la clasificación con distinto rango de frecuencias en los audios. El límite inferior es 50 Hz y el límite superior es la frecuencia señalada.

En la gráfica 5.4 se muestran los resultados de clasificación de los 3 sujetos variando el tamaño del rango de frecuencias, estableciendo un mínimo de 50 Hz y máximos de 1 KHz hasta 16 KHz. Se muestra que un rango de frecuencias mayor a 8 KHz no tiene buenos resultados en la clasificación, mientras que establecer un rango entre 3 KHz y 8 KHz obtiene mejores resultados.

El tamaño de las ventanas para crear el espectrograma influye en el detalle de los

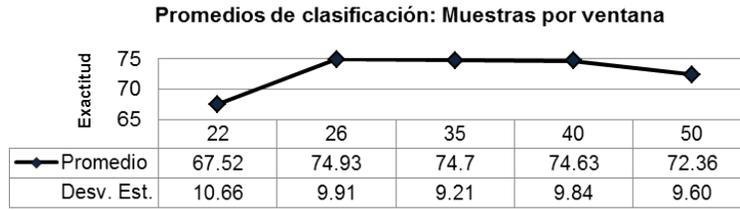


Figura 5.5: Promedios de exactitud en la clasificación con distinto número de muestras por ventana para formar el espectrograma.

cambios capturados en el tiempo de la señal. En la gráfica 5.5 se muestra que una ventana formada por 22 muestras no genera un espectrograma que capture los detalles necesarios para obtener el mejor porcentaje de clasificación, mientras que una ventana de 26 muestras o mayor obtiene mejores porcentajes.

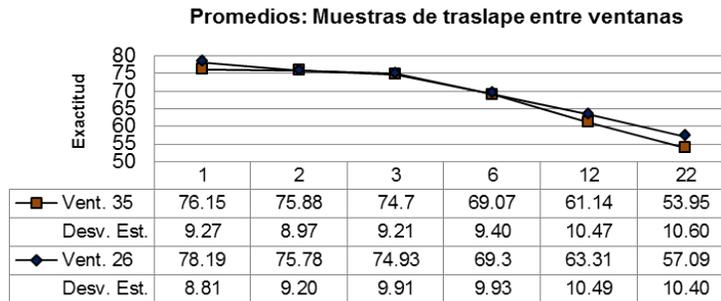


Figura 5.6: Promedios de exactitud en la clasificación con las dos mejores muestras por ventana para el espectrograma.

El traslape influye en la longitud y detalle de cambios capturados por el espectrograma. En la gráfica 5.6 se muestran los resultados de variar el traslape entre ventanas de dos configuraciones de tamaño de ventana. Se muestra que al generar el espectrograma de la señal de EEG con poco traslape entre ventanas se obtienen los mejores resultados de clasificación para las palabras imaginadas.

Los valores seleccionados de los parámetros fueron: número de tonos (14), frecuencia mínima de la señal EEG (1 Hz), frecuencia máxima de la señal EEG (60 Hz), tamaño de la ventana del espectrograma(26 muestras), traslape entre ventanas del espectrograma (1 muestra), duración de los tonos (0.6 seg.), frecuencia mínima del audio (50 Hz), frecuencia máxima del audio (5000 Hz), frecuencia de muestreo del audio (8000 Hz).

5.3. Selección de Wavelet para EEG Sonificado

En esta sección se explican las simulaciones realizadas usando la DWT como método de extracción de características. Se hizo la búsqueda sobre la elección de la Wavelet y el número de niveles de descomposición a utilizar.

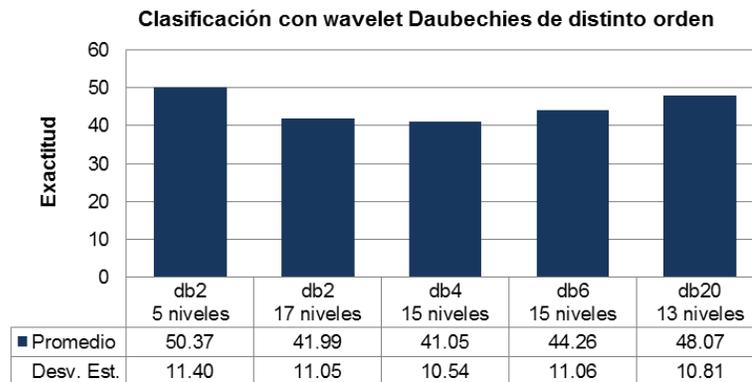


Figura 5.7: Comparativa para la elección de wavelets.

Se evaluó el comportamiento de distintas Wavelets de la familia Daubechies (db), en este caso se evaluaron las Wavelets Daubechies 2, 4, 6 y 20. Cada Daubechies puede soportar distinto rango de niveles de descomposición y entre más grande es el orden de la Wavelet menos niveles de descomposición puede aceptar. Para ello se calculó el máximo número de niveles de descomposición que puede soportar en el rango del audio. El rango de audio usado fue el elegido previamente de 4950 Khz (50 Hz a 5 KHz). En la gráfica 5.7 se incluyen las Wavelets Daubechies ya mencionadas (db2, db4, db6 y db20) con su máximo número de niveles de descomposición, además de la configuración bajo la cual en [67] obtuvo mejores porcentajes de clasificación (db2 con 5 niveles de descomposición). Se muestra que el comportamiento es similar para los 3 sujetos independientemente de la Wavelet que se elija. La Wavelet db2 con 5 niveles de descomposición obtuvo los mejores resultados de clasificación y con esto se muestra que no necesariamente el máximo número de niveles de descomposición es el que más ayuda a clasificar las palabras.

Después se procedió a calcular los porcentajes de clasificación que se obtienen al variar los niveles de descomposición de las Wavelets, para ello se seleccionó de la simulación anterior la configuración que mejor resultado obtuvo en general (Daubechies 2 con 5 niveles) y la que mejor resultado obtuvo de entre las que se les calculó su máximo nivel (Daubechies 20). En la gráfica 5.8 se muestra el comportamiento que tiene el variar los niveles de las Wavelets db2 y db20, se muestra que con pocos niveles

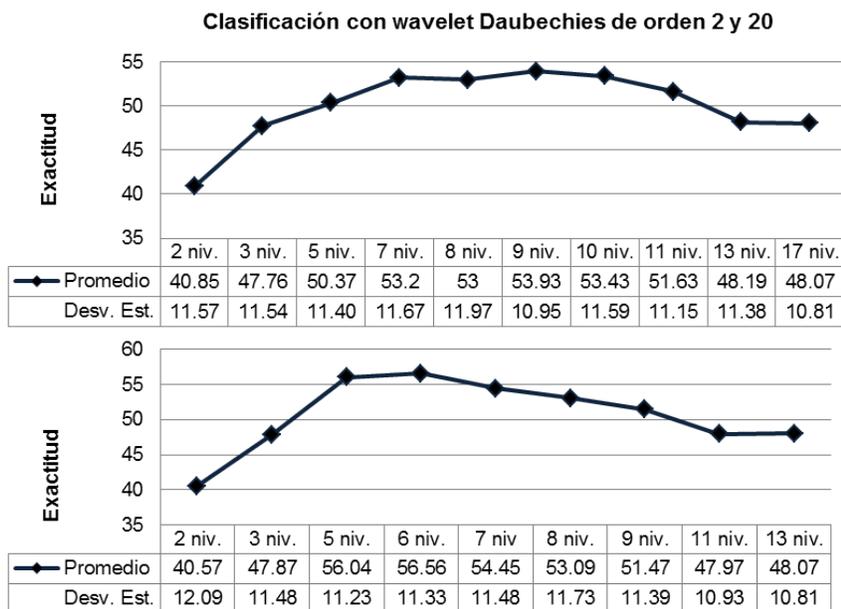


Figura 5.8: Resultados de exactitud promedio en la clasificación usando Daubechies de orden 2 y de orden 20, con distintos niveles de descomposición.

de descomposición (menos de 5) el porcentaje de clasificación es bajo. También se muestra que con el máximo número de niveles en ninguno de los dos casos obtiene el mejor porcentaje de clasificación. Al final del análisis se decidió elegir la Wavelet Daubechies de orden 20 con 6 niveles de descomposición.

5.4. Configuración de parámetros de MFCC para sonificación de EEG

Se realizaron diversas simulaciones para evaluar el comportamiento de los parámetros del método de MFCC para EEG sonificado.

En la gráfica 5.9 se muestra el comportamiento de variar el número de coeficientes cepstrales calculados de la señal de audio que contiene la palabra imaginada. Con los resultados de clasificación obtenidos se muestra que es mejor calcular alrededor de entre 17 coeficientes cepstrales de la señal.

Al igual que en el algoritmo de sonificación de EEG, el tamaño de la ventana para formar el espectrograma influye en los resultados de clasificación. En la gráfica se muestran los resultados de clasificación a partir de los 3 sujetos seleccionados al evaluar el tamaño de la ventana (dado en milisegundos) utilizada para formar el espectrograma de la señal de audio. El comportamiento de la gráfica nos muestra que

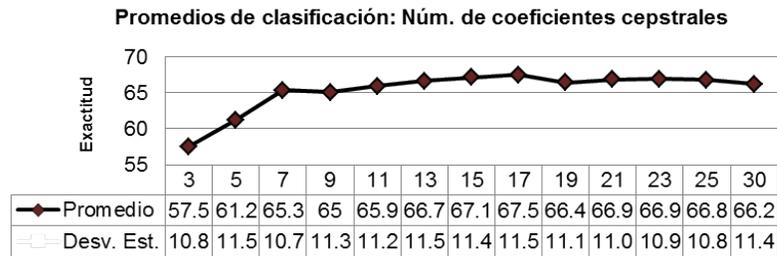


Figura 5.9: Promedios de exactitud en la clasificación con distinto número de coeficientes cepstrales, usando MFCC para sonificación de EEG.

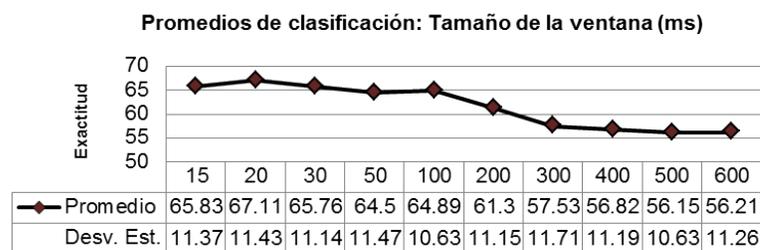


Figura 5.10: Promedios de exactitud en la clasificación con distinto tamaño de ventana, usando MFCC para sonificación de EEG.

para tamaños de aproximadamente 100 ms en adelante los promedios de clasificación bajan con respecto al rango de entre 15 y 100 ms donde los resultados son mejores.

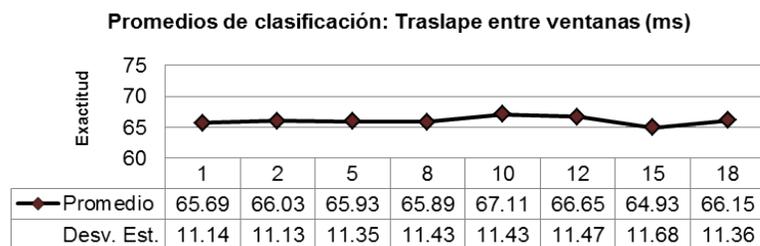


Figura 5.11: Promedios de exactitud en la clasificación con distinto traslape entre ventanas, usando MFCC para sonificación de EEG.

Se evaluó también el comportamiento de varios tamaños del traslape entre ventanas, manteniendo un tamaño de ventana de 20 milisegundos. En la gráfica 5.11 se muestra que la variación en el traslape entre 1 y 18 milisegundos solo tiene cambios mínimos en sus porcentajes de clasificación.

Se evaluó el comportamiento de calcular los cambios entre segmentos de coeficientes calculados (valores delta) y los cambios entre los valores delta (valores doble delta).

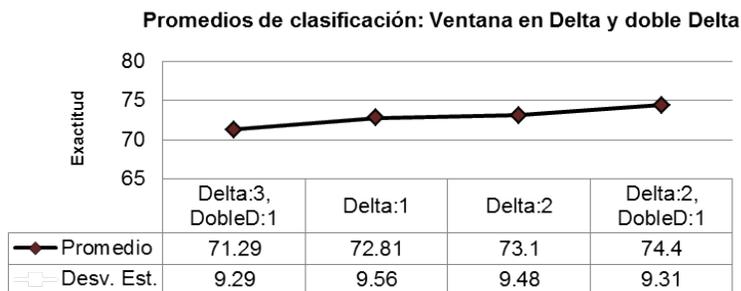


Figura 5.12: Promedios de exactitud en la clasificación con distinto tamaño de ventanas en delta y doble delta, usando MFCC para sonificación de EEG.

En la gráfica 5.12 se muestra que las ventanas amplias ya sea en delta o doble delta no mejoran la clasificación, se muestra que es mejor elegir la ventana mínima.

La configuración elegida para la extracción de características de los audios usando MFCC son: duración de la ventana (20 ms), duración del traslape entre ventanas (10 ms), frecuencia inferior del audio (50 Hz), frecuencia superior (5000 Hz), numero de coeficientes cepstrales (23). Y los valores estadísticos elegidos son: (máximo, mínimo, media, Desv. Estándar, valores delta con ventana (2) y valores doble delta con ventana de (1)).

5.5. Configuración de parámetros de MFCC para EEG

En esta sección se muestra el comportamiento de evaluar distintas configuraciones usando MFCC para la señal de EEG. Aunque se conoce que la de EEG no tiene las características que si tiene una señal de audio, fue necesario el calculo para poder comparar entre el uso de MFCC para una señal cerebral normal y una señal sonificada.

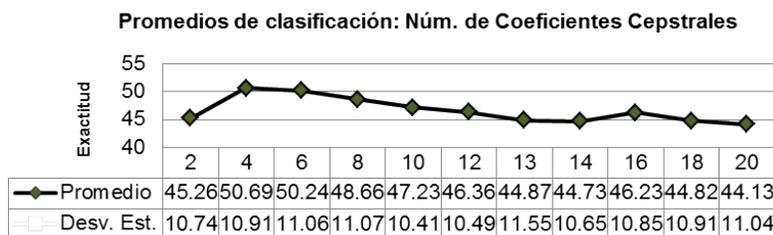


Figura 5.13: Promedios de exactitud en la clasificación con distinto número de coeficientes cepstrales, usando MFCC para EEG sin sonificar.

En la gráfica 5.13 se muestran los resultados de clasificación al evaluar la extracción

de diferentes números de coeficientes cepstrales. Se muestra que los mejores resultados se obtienen cuando se calculan entre 4 y 6 coeficientes cepstrales.

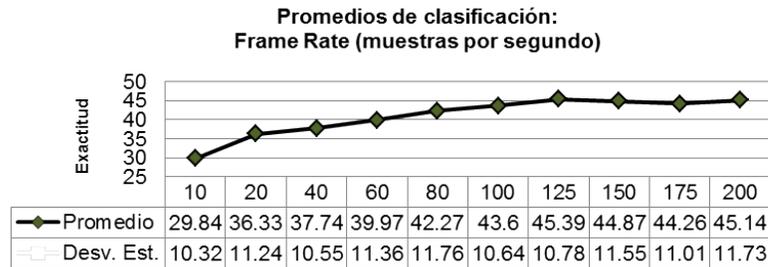


Figura 5.14: Promedios de exactitud en la clasificación con distinto tamaño de ventana y traslape que en combinación dan el número de muestras por segundo, usando MFCC para EEG sin sonificar.

Se evaluó también el comportamiento que tiene la tasa de muestreo (*frame rate*) en la clasificación usando MFCC para la señal de EEG. El frame rate se calcula en base a las muestras por ventana y el traslape por ventana en referencia con la frecuencia de muestreo de la señal que en este caso es de 128 Hz. Se muestra en la gráfica 5.14 que es mejor elegir un *frame rate* de entre 125 y 200 muestras por segundo.

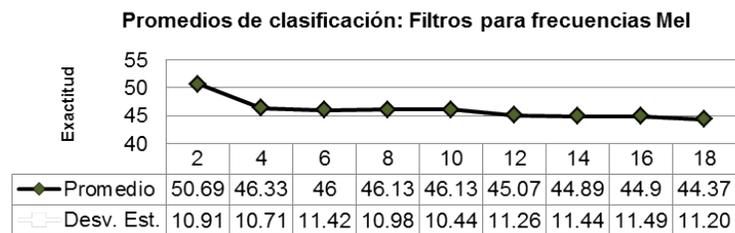


Figura 5.15: Promedios de exactitud en la clasificación con distinto número de filtros en las frecuencias Mel, usando MFCC para EEG sin sonificar.

El número de filtros utilizados para la extracción de características de EEG usando MFCC afecta el rendimiento en la clasificación, tal como se muestra en la gráfica 5.15 donde a partir de más de 2 filtros los resultados de clasificación bajan. Esto sugiere que pase debido a que en la señal de EEG el rango de frecuencias es reducido comparadas con las del audio y una cantidad baja de filtros es suficiente para caracterizarla mejor.

Al igual que con la señal de audio se evaluaron los porcentajes de clasificación calculando los valores delta y doble delta de los coeficientes cepstrales. En la gráfica 5.16 se muestra que los mejores resultados se obtienen al no utilizar los valores doble delta. Esto sugiere que en la señal de EEG sin sonificar no es conveniente ampliar el

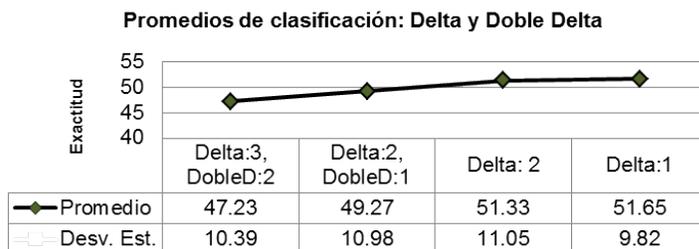


Figura 5.16: Promedios de exactitud en la clasificación con distinto tamaño de ventanas en delta y doble delta, usando MFCC para EEG sin sonificar.

calculo de variaciones en el tiempo al menos para nuestro caso donde las palabras imaginadas tienen pocas muestras por segundo.

En el algoritmo para la extracción de características usando MFCC para la señal de EEG se usaron los siguientes parámetros: coeficientes cesptrales (4), frame rate (125), frecuencia mínima (1 Hz), frecuencia máxima (60 Hz), filtros para las frecuencias (2). Y se utilizaron los siguientes valores estadísticos: máximo, mínimo, media, desviación estándar y valores delta con ventana (1).

5.6. Resumen de configuraciones utilizadas

Las configuraciones de parámetros utilizadas en los experimentos finales son descritas a continuación.

- Sonificación de EEG a tonos: Número de tonos = 14, frecuencia mínima de la señal EEG = 1 Hz, frecuencia máxima de la señal EEG = 60 Hz, tamaño de la ventana del espectrograma = 26 muestras, traslape entre ventanas del espectrograma = 1, duración de los tonos = 0.6 seg., frecuencia mínima del audio = 50 Hz, frecuencia máxima del audio = 5000 Hz, frecuencia de muestreo del audio = 8000 Hz.
- Extracción de características usando DWT en EEG: Wavelet daubechies orden 2, niveles de descomposición = 5.
- Extracción de características usando DWT en EEG sonificado: Wavelet daubechies orden 20, niveles de descomposición = 6.
- Extracción de características usando MFCC en EEG sonificado: Duración de la ventana = 20 ms, duración del traslape entre ventanas = 10 ms, frecuencia inferior = 50 Hz, frecuencia superior = 5000 Hz, coeficientes cesptrales = 23. Atributos

estadísticos: máximo, mínimo, media, Desv. Estándar. Valores delta con ventana = 2 y valores doble delta con ventana = 1.

- Extracción de características usando MFCC en EEG: Coeficientes cesptrales = 4, frame rate = 125, frecuencia mínima = 1 Hz, frecuencia máxima = 60 Hz, filtros para las frecuencias = 2. Atributos estadísticos: máximo, mínimo, media, desviación estándar. Valores delta con ventana = 1.

5.7. Experimento de clasificación para los 27 sujetos

Después de elegir las mejores configuraciones para las etapas propuestas se procedió a usarlos en la clasificación de las palabras imaginadas de los 27 sujetos. Los resultados obtenidos incluyen los 3 enfoques propuestos: Extracción de características con DWT para EEG sonificado, extracción de características con MFCC para EEG y extracción de características con MFCC para EEG sonificado. Y se comparan con el enfoque propuesto en el trabajo [67] el cual no utiliza la sonificación de EEG y extrae características por medio de DWT para comprobar si existe mejora en la exactitud de clasificación. Los resultados se separan en dos criterios de elección de canales 4 y 14 canales con el objetivo de evaluar si un conjunto reducido de 4 canales del modelo del procesamiento del habla permite obtener mantener o mejorar los resultados de exactitud.

5.7.1. Resultados de clasificación en los 4 enfoques

En las tablas 5.1 y 5.2, se muestran los resultados de clasificación de los 27 sujetos con los 4 enfoques evaluados, usando 4 y 14 respectivamente. En base a las tablas se muestra en las gráficas 5.17 y 5.18 los resultados de exactitud promedio para cada sujeto y para cada método utilizado. En dichas gráficas se muestra que no hay método dominante, pero que especialmente los enfoques que usan sonificación logran mejorar los porcentajes de forma notoria en algunos sujetos. Por ejemplo para 4 canales el sujeto 26 que utiliza sonificación de EEG con MFCC, aumenta 11.08 % su promedio con respecto a no usar sonificación y el sujeto 24 aumenta su promedio de clasificación 26.67 % usando sonificación de EEG con DWT, mientras que para 14 canales es notoria la mejora que tiene el sujeto 10 al usar la sonificación de EEG con MFCC es de 15.15 % o la del sujeto S13 usando la sonificación de EEG con DWT cuya diferencia es de 18.79 %.

Tabla 5.1: Promedios de exactitud en la clasificación de los 4 enfoques usando 4 canales

Suj.	EEG DWT	Desv.	EEG Son. DWT	Desv.	EEG MFCC	Desv.	EEG Son. MFCC	Desv.
S1	66.06	9.54	70.30	8.78	50.91	12.72	59.39	10.86
S2	38.18	14.44	40.00	14.73	29.70	15.87	43.64	13.79
S3	39.39	14.44	55.76	11.05	51.20	13.09	47.88	12.99
S4	46.67	13.28	58.18	10.74	44.24	13.21	55.76	11.76
S5	61.82	9.95	75.15	8.22	48.48	12.96	71.52	9.46
S6	31.52	15.24	30.91	15.57	32.73	15.85	31.52	14.69
S7	48.78	13.15	52.44	12.15	38.41	14.34	46.95	12.95
S8	66.06	9.46	73.33	8.01	40.61	14.37	67.88	9.39
S9	58.79	11.35	60.61	10.90	51.52	12.14	62.42	11.20
S10	41.21	14.36	48.48	13.28	29.09	15.89	45.45	13.82
S11	73.33	7.79	83.03	5.37	62.42	9.96	79.39	6.23
S12	49.70	12.89	59.39	10.61	50.00	13.30	55.76	11.58
S13	49.09	11.93	60.61	10.30	32.73	14.75	58.79	11.56
S14	32.73	15.40	41.82	14.30	24.24	16.55	31.52	15.21
S15	59.39	10.99	64.85	9.97	40.00	14.64	62.42	11.15
S16	37.58	14.29	46.06	13.18	34.34	15.01	43.64	14.11
S17	49.70	12.45	55.76	11.48	30.91	15.99	54.55	11.40
S18	58.79	11.60	54.55	11.53	46.06	13.49	49.09	12.24
S19	27.88	15.99	32.12	15.36	19.39	17.24	29.70	15.44
S20	75.76	7.40	73.33	7.83	49.70	12.76	75.15	8.12
S21	31.52	15.67	49.70	12.60	23.64	15.83	39.39	14.75
S22	55.15	10.40	68.48	8.97	44.24	13.47	58.79	9.90
S23	47.27	13.73	57.58	11.71	35.15	14.33	56.97	11.98
S24	33.33	14.70	60.00	10.63	31.93	15.88	44.85	13.58
S25	28.48	17.32	33.94	16.65	23.03	19.00	32.12	17.15
S26	41.21	14.23	44.85	13.39	42.86	14.08	53.94	12.10
S27	49.39	12.71	56.10	11.90	32.93	15.69	55.49	12.35
Prom.	48.10	12.77	55.83	11.45	38.54	14.53	52.37	12.21

En la figura 5.19 se muestran los promedios de exactitud de clasificación para los 4 enfoques. En los resultados se muestra que para 4 canales se obtiene mejor porcentaje de acierto en la clasificación con el enfoque de sonificación de EEG usando DWT con un 55.83 %, teniendo una diferencia de 7.73 % con respecto a clasificar sin usar sonificación de EEG. También se muestra que el enfoque de sonificación de EEG usando MFCC obtuvo 52.37 % de promedio de clasificación y es mejor que el reportado en el trabajo [67] con 48.1 %, existiendo una diferencia de 4.27 %, por otra parte el enfoque de EEG usando MFCC es el que resultó con los porcentajes más bajos.

El método que mejores porcentajes de clasificación obtiene es el que hace sonificación de EEG y usa DWT para la extracción de características, obteniendo un 63.82 %, dicho porcentaje tiene una diferencia de 5.41 % con respecto al método que no usa

Tabla 5.2: Promedios de exactitud en la clasificación de los 4 enfoques usando 14 canales

Suj.	EEG DWT	Desv.	EEG Son. DWT	Desv.	EEG MFCC	Desv.	EEG Son. MFCC	Desv.
S1	73.94	8.28	76.36	8.13	50.30	12.33	67.88	9.99
S2	42.42	13.98	44.85	13.90	30.30	15.79	45.45	13.45
S3	63.03	10.50	61.82	9.85	37.95	14.46	62.42	10.77
S4	60.00	11.24	70.30	8.71	50.91	12.18	67.27	10.04
S5	72.12	9.80	77.58	8.94	48.48	12.68	74.55	9.81
S6	36.36	14.54	40.00	14.17	33.33	15.76	38.18	14.47
S7	56.71	11.41	60.98	11.30	39.63	14.76	63.41	11.48
S8	82.42	7.74	87.27	6.20	40.00	12.79	84.85	7.33
S9	64.24	11.05	65.45	10.59	35.15	15.66	63.64	11.15
S10	57.58	11.89	67.27	10.61	34.55	15.52	72.73	10.66
S11	83.64	5.82	88.48	4.71	53.33	11.44	83.64	5.66
S12	60.00	11.57	66.06	10.24	39.90	14.05	61.82	11.76
S13	55.76	11.40	74.55	8.95	29.09	15.97	64.24	10.95
S14	44.24	14.26	46.67	13.74	27.88	16.46	32.73	14.78
S15	63.64	10.41	72.73	8.84	33.94	15.61	66.06	11.26
S16	45.45	12.93	50.30	13.13	33.13	15.10	43.64	13.81
S17	60.61	10.61	69.09	9.68	40.61	14.34	70.30	10.20
S18	68.48	10.22	69.70	9.77	40.00	14.35	64.24	10.73
S19	52.73	13.11	47.27	12.80	21.82	17.08	47.27	13.16
S20	75.15	8.04	76.36	8.64	38.18	14.79	74.55	8.81
S21	43.03	14.08	50.30	11.98	29.70	15.93	45.45	13.73
S22	65.45	9.33	72.12	8.35	50.91	11.70	67.88	9.74
S23	51.52	13.24	62.42	11.35	35.15	15.07	55.15	12.28
S24	46.67	13.36	68.48	9.74	27.11	16.31	52.73	12.97
S25	43.64	16.06	36.97	16.67	31.52	17.87	37.58	17.00
S26	56.36	12.41	60.61	11.93	30.95	16.07	64.85	11.16
S27	51.83	11.80	59.15	10.81	29.27	15.68	65.24	11.10
Prom.	58.41	11.45	63.82	10.51	36.78	14.81	60.66	11.42

sonificación de EEG. Al igual que lo mostrado en los resultados de 4 canales, el método de sonificación de EEG que extrae características usando MFCC tiene mejores porcentajes de exactitud que el método que no utiliza la sonificación de EEG, obteniendo un 60.66 % de correcta clasificación promedio, esto da una diferencia de 2.25 % con respecto al trabajo con el que se compara que obtuvo un 58.41 %. El método de MFCC para la señal de EEG sin sonificación es el que obtiene el promedio de clasificación más bajo.

En la tabla 5.20 se muestran ordenados los enfoques de acuerdo a la cantidad de sujetos en las que es mejor cada enfoque. Se muestra que, ya sea usando 4 o 14 canales el orden en el acierto de clasificación se mantiene, es decir que si escogiéramos un método para aplicarlo a una BCI escogeríamos la sonificación de EEG con extracción

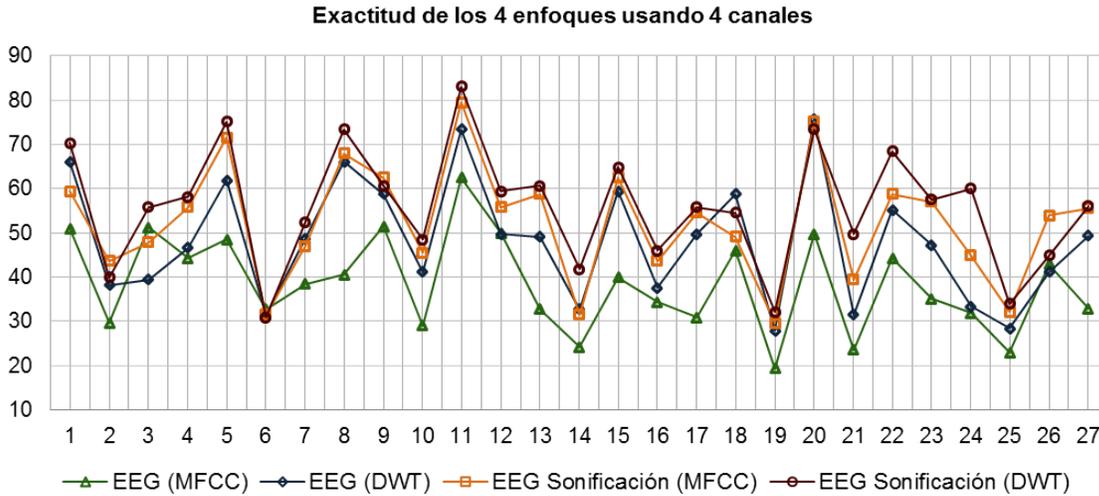


Figura 5.17: Exactitud de la clasificación para los 27 sujetos usando 4 canales.

usando DWT, pues en 21 de los 27 sujetos para 4 canales y 18 de los 27 sujetos para 14 canales obtiene el mejor porcentaje de clasificación.

5.7.2. Prueba estadística

Se realizaron pruebas estadísticas para comprobar si las diferencias existentes entre los resultados de clasificación de los métodos propuestos son significativas. Para ello se evaluarán los promedios de clasificación de los 27 sujetos en las 4 metodologías, es decir, se comparan los resultados de *EEG con extracción de características usando DWT*, *Sonificación de EEG con extracción de características usando DWT*, *EEG con extracción de características usando MFCC* y *Sonificación de EEG con extracción de características usando MFCC*, esto para los 2 enfoques de elección de canales 4 y 14 respectivamente, por lo que en total se realizaron 4 pruebas estadísticas.

Para conocer el tipo de prueba a utilizar se aplicó la prueba de normalidad *Shapiro-Wilk*. En la figura 5.21 se muestra el resultado de aplicar la prueba de *Shapiro-Wilk* con un nivel de significancia de 0.05. Como resultado tenemos todos los valores de la columna de significancia son mayores a 0.05, por lo que se concluye que los promedios utilizados siguen una distribución normal.

Se aplicó la prueba paramétrica *ANOVA de un factor*, dado que los promedios siguen una distribución normal y necesitamos comparar más de dos grupos de promedios. El nivel de significación fue establecido a 0.05 y el factor de la prueba son los métodos utilizados en esta investigación (*EEG con DWT*, *EEG sonificado con DWT*, *EEG con MFCC* y *EEG sonificado con MFCC*). Al aplicar la prueba para los dos enfoques de 4 y

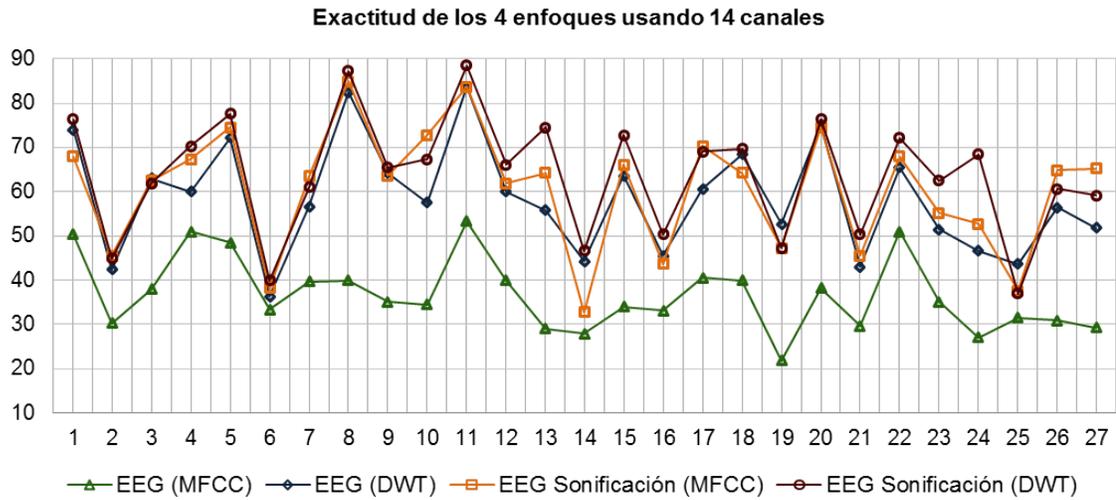


Figura 5.18: Exactitud de la clasificación para los 27 sujetos usando 14 canales.

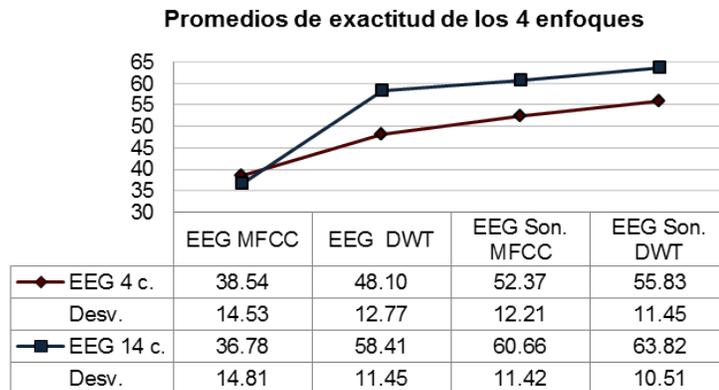


Figura 5.19: Promedios de exactitud de los 4 enfoques en 4 y 14 canales.

14 canales el resultado es menor que el nivel de significancia establecido, con lo que se concluye que existe al menos un par de grupos cuya diferencia es estadísticamente significativa.

Después se realizó una prueba estadística que permite la comparación por pares para conocer el resultado de los grupos de manera individual. En nuestro caso utilizaremos la prueba *Post-hoc* de *Tukey*. La figura 5.22 muestra el resultado de todas las comparaciones realizadas. Como podemos apreciar las diferencias existentes según la prueba estadística no son significativas entre usar EEG con DWT y usar EEG sonificado con DWT o MFCC, pues el valor de la significación es mayor al nivel establecido de 0.05. Por otra parte las comparaciones realizadas contra el método de EEG sin sonificar

Rank	Método	Cantidad de sujetos 4 canales	Cantidad de sujetos 14 canales
1	EEG Sonif. DWT	21	18
2	EEG Sonif. MFCC	3	6
3	EEG DWT	2	3
4	EEG MFCC	1	0
	Total	27	27

Figura 5.20: Cantidad de sujetos que obtienen sus mejores resultados para cada enfoque.

Método	Sig.	Dist. normal	Método	Sig.	Dist. normal
EEG DWT 4 c.	.333	Si	EEG DWT 14 c.	.686	Si
EEG Son. DWT 4 c.	.822	Si	EEG Son. DWT 14 c.	.478	Si
EEG MFCC 4 c.	.620	Si	EEG MFCC 14 c.	.104	Si
EEG Son. MFCC 4 c.	.599	Si	EEG Son. MFCC 14 c.	.160	Si

Figura 5.21: Resultados de la prueba de normalidad Shapiro-Wilk por pares.

con MFCC si presentan una diferencia significativa cuando se les compara con EEG con DWT, EEG sonificado con DWT y EEG sonificado con MFCC. Aunque los valores de significación varían el resultado de si existe o no diferencia significativa entre pares se mantiene en ambos enfoques de 4 y 14 canales. Con lo anterior solo se puede concluir que con el experimento planteado bajo las condiciones establecidas solo la técnica de EEG con MFCC tiene diferencias con las demás y que esos resultados podrían no ser causa del azar. Aunque no se asegura que bajo otros experimentos los promedios analizados mantengan las mismas diferencias.

Comparaciones (4 canales)	Sig.	Signific. estadística	Comparaciones (14 canales)	Sig.	Signific. estadística
EEG DWT - EEG Son. DWT	.125	No	EEG DWT - EEG Son. DWT	.361	No
EEG DWT - EEG MFCC	.035	Si	EEG DWT - EEG MFCC	.000	Si
EEG DWT - EEG Son. MFCC	.613	No	EEG DWT - EEG Son. MFCC	.904	No
EEG Son. DWT - EEG MFCC	.000	Si	EEG Son. DWT - EEG MFCC	.000	Si
EEG Son. DWT - EEG Son. MFCC	.754	No	EEG Son. DWT - EEG Son. MFCC	.774	No
EEG MFCC - EEG Son. MFCC	.001	Si	EEG MFCC - EEG Son. MFCC	.000	Si

Figura 5.22: Resultados de la prueba de significancia estadística por pares, usando comparación por pares.

5.7.3. Análisis de clasificación por palabra

Después de conocer los resultados promedio de clasificación de cada sujeto, se procedió a hacer un análisis por palabra para poder evaluar que método clasifica mejor cada palabra. También se realizó un análisis para conocer que palabras son más fáciles

clasificar y en cuales se tienen dificultades.

Para ello se evaluó el porcentaje de exactitud del clasificador en las 5 palabras por separado basados en las matrices de confusión obtenidas al clasificar (ver Apéndice B.1). Evaluando los resultados de una sola palabra por los 27 sujetos se obtuvieron promedios de clasificación por método y por palabra.

Los resultados por palabra para 4 y 14 canales son mostrados en la figuras 5.23 y 5.24 respectivamente. En dichas gráficas se muestra que el desempeño de los métodos investigados en este trabajo se mantiene con un comportamiento muy similar independientemente de la palabra que se trate.

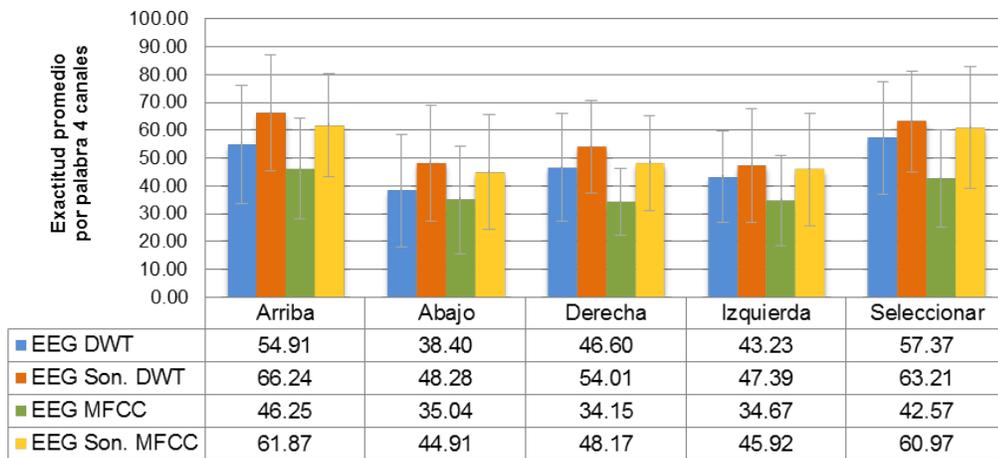


Figura 5.23: Promedios de exactitud en la clasificación por palabra usando 4 canales.

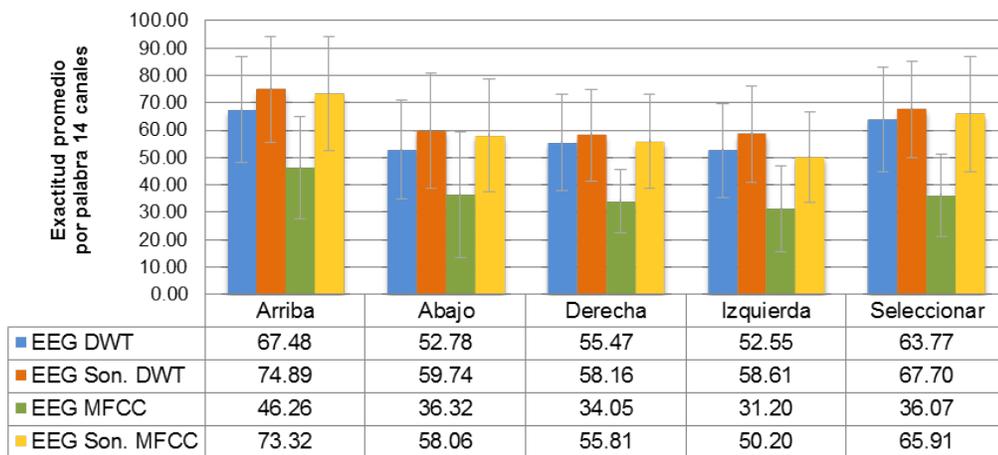


Figura 5.24: Promedios de exactitud en la clasificación por palabra usando 14 canales.

En la tabla 5.25 se muestra el ordenamiento de menor a mayor del resumen de los

resultados mostrados en las gráficas anteriores 5.23 y 5.24, con el objetivo de evaluar si las palabras mantienen sus porcentajes de clasificación con respecto a las demás al pasar de 4 a 14 canales. En dicha tabla se muestra que usando 4 y 14 canales, en general la palabra *arriba* es la que mejor se logra discriminar de entre las demás, mientras que las palabras que más se dificultan son las palabras *abajo* para 4 canales y la palabra *izquierda* para 14 canales, no se mantuvo la misma palabra como la peor clasificada ya que la palabra *izquierda* bajo en el orden de las mejores clasificadas, pues la palabra *abajo* aumenta 10.06 % en su exactitud de clasificación promedio al ser reconocida con 14 canales. Con esto último podemos considerar que existen palabras que aumentarán considerablemente su clasificación si se les lee desde los canales adecuados.

	4 canales	Promedio	Desv.	14 canales	Promedio	Desv.	Diferencia
1	Arriba	57.32	7.56	Arriba	65.49	11.44	8.17
2	Seleccionar	56.03	8.05	Seleccionar	58.36	12.94	2.33
3	Derecha	45.73	7.24	Abajo	51.72	9.26	10.06
4	Izquierda	42.80	4.93	Derecha	50.87	9.77	5.14
5	Abajo	41.66	5.22	Izquierda	48.14	10.25	5.34

Figura 5.25: Palabras ordenadas de acuerdo a la exactitud de clasificación promedio. Se muestra la diferencia que existió en la misma palabra al pasar de 4 a 14 canales

Se realizó un concentrado de los porcentajes de clasificación correcta por palabra para los 27 sujetos usando el método en el que obtuvieron el mejor promedio de exactitud. Los porcentajes se muestran en la tabla 5.26 para 4 canales y en la tabla 5.27 para 14 canales. Los resultados muestran, como se esperaba, que cada persona obtiene diferentes resultados de clasificación al pensar una palabra en específico. También muestran que hay sujetos cuyo proceso de generar alguna palabra logra ser reconocido de forma efectiva, con porcentajes arriba del 95 %, como ejemplo el sujeto 22 logra un 100 % de exactitud en la clasificación de la palabra *arriba* pero no generó los patrones que lograran discriminar igual de bien las demás palabras. Otro ejemplo es el sujeto 11 que logra mantener porcentajes arriba de 72 % (72.73 % en la palabra *arriba*, 81.82 % en la palabra *abajo*, 100 % en la palabra *derecha*, 78.79 % en la palabra *izquierda* y 81.82 % en la palabra *seleccionar*), logrando el mejor porcentaje para 4 canales de 83.03 %. En la tabla de 14 canales se muestra que el sujeto 11 también es el mejor clasificado 88.4 %, aunque ya no tiene alguna palabra clasificada al 100 %, mantiene porcentajes arriba de 78 % (84.85 % en la palabra *arriba*, 93.94 % en la palabra *abajo*, 96.97 % en la palabra *derecha*, 87.88 % en la palabra *izquierda* y 78.79 % en la palabra *seleccionar*).

Se calculó cual sería el promedio si se escogiera para cada sujeto el mejor enfoque de los mostrados en este trabajo, con esos porcentajes se puede tener una idea de lo

Suj.	Método	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
S1	EEG Sonif. DWT	93.94%	84.85%	57.58%	42.42%	72.73%	70.30%	8.78%
S2	EEG Sonif. MFCC	63.64%	9.09%	42.42%	42.42%	60.61%	43.64%	13.79%
S3	EEG Sonif. DWT	90.91%	33.33%	60.61%	33.33%	60.61%	55.76%	11.05%
S4	EEG Sonif. DWT	75.76%	57.58%	42.42%	69.70%	45.46%	58.18%	10.74%
S5	EEG Sonif. DWT	87.88%	81.82%	63.64%	57.58%	84.85%	75.15%	8.22%
S6	EEG MFCC	39.39%	51.52%	27.27%	30.30%	15.15%	32.73%	15.85%
S7	EEG Sonif. DWT	70.12%	48.78%	42.68%	51.83%	48.78%	52.44%	12.15%
S8	EEG Sonif. DWT	72.73%	60.61%	57.58%	81.82%	93.94%	73.33%	8.01%
S9	EEG Sonif. MFCC	69.70%	66.67%	69.70%	36.36%	69.70%	62.42%	11.20%
S10	EEG Sonif. DWT	36.36%	27.27%	48.49%	39.39%	90.91%	48.49%	13.28%
S11	EEG Sonif. DWT	72.73%	81.82%	100.00%	78.79%	81.82%	83.03%	5.37%
S12	EEG Sonif. DWT	66.67%	48.49%	51.52%	57.58%	72.73%	59.39%	10.61%
S13	EEG Sonif. DWT	75.76%	42.42%	66.67%	69.70%	48.49%	60.61%	10.30%
S14	EEG Sonif. DWT	63.64%	39.39%	63.64%	21.21%	21.21%	41.82%	14.30%
S15	EEG Sonif. DWT	42.42%	63.64%	72.73%	48.49%	96.97%	64.85%	9.97%
S16	EEG Sonif. DWT	78.79%	39.39%	24.24%	30.30%	57.58%	46.06%	13.18%
S17	EEG Sonif. DWT	90.91%	39.39%	51.52%	42.42%	54.55%	55.76%	12.45%
S18	EEG DWT	54.55%	39.39%	54.55%	72.73%	72.73%	58.79%	11.60%
S19	EEG Sonif. DWT	36.36%	21.21%	21.21%	33.33%	48.49%	32.12%	15.36%
S20	EEG DWT	72.73%	93.94%	72.73%	63.64%	75.76%	75.76%	7.40%
S21	EEG Sonif. DWT	54.55%	45.46%	45.46%	27.27%	75.76%	49.70%	12.60%
S22	EEG Sonif. DWT	100.00%	48.49%	36.36%	87.88%	69.70%	68.49%	8.97%
S23	EEG Sonif. DWT	60.61%	54.55%	54.55%	60.61%	57.58%	57.58%	11.71%
S24	EEG Sonif. DWT	87.88%	75.76%	45.46%	18.18%	72.73%	60.00%	10.63%
S25	EEG Sonif. DWT	6.06%	9.09%	45.46%	54.55%	54.55%	33.94%	16.65%
S26	EEG Sonif. MFCC	48.49%	48.49%	54.55%	39.39%	78.79%	53.94%	12.10%
S27	EEG Sonif. DWT	45.73%	33.54%	76.22%	54.88%	70.12%	56.10%	11.90%
	Promedio	65.12%	49.85%	53.67%	49.86%	64.90%	56.68%	11.41%

Figura 5.26: Mejores porcentajes de exactitud por método en las 5 palabras usando 4 canales.

que mejoraría el porcentaje global de clasificación para los 27 sujetos con un esquema de selección de enfoques por sujeto. En las figuras 5.26 y 5.27 se muestra que dicha mejora sería de un 0.85 % para 4 canales (de 55.83 % a 56.68 %) y de 1.23 % para 14 canales (de 63.83 % a 65.06 %).

También se reportan los peores porcentajes de exactitud que un sujeto tuvo en alguno de los 4 enfoques. En las tablas 5.28 y 5.29 se muestra que el método de extracción de características con MFCC en la señal de EEG sin sonificar es el que más ocasiones obtiene los menores porcentajes de exactitud. Esto muestra que éste método es poco eficiente especialmente en la clasificación con 14 canales.

5.7.4. Energía relativa por palabra

Como ya se mencionó anteriormente la energía relativa wavelet se obtiene a partir de la transformada wavelet y del cálculo de la energía que aporta cada nivel de descomposición con respecto a la energía total de todos los niveles. De manera similar

Suj.	Método	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
S1	EEG Sonif. DWT	96.97%	84.85%	60.61%	57.58%	81.82%	76.36%	8.13%
S2	EEG Sonif. MFCC	75.76%	33.33%	54.55%	27.27%	36.36%	45.45%	13.45%
S3	EEG DWT	81.82%	54.55%	45.46%	48.49%	84.85%	63.03%	10.50%
S4	EEG Sonif. DWT	93.94%	78.79%	48.49%	75.76%	54.55%	70.30%	8.71%
S5	EEG Sonif. DWT	84.85%	78.79%	60.61%	72.73%	90.91%	77.58%	8.94%
S6	EEG Sonif. DWT	54.55%	45.46%	24.24%	36.36%	39.39%	40.00%	14.17%
S7	EEG Sonif. MFCC	82.32%	67.07%	48.78%	57.93%	60.98%	63.41%	11.48%
S8	EEG Sonif. DWT	93.94%	81.82%	78.79%	84.85%	96.97%	87.27%	6.20%
S9	EEG Sonif. DWT	69.70%	81.82%	63.64%	57.58%	54.55%	65.45%	10.59%
S10	EEG Sonif. MFCC	81.82%	81.82%	54.55%	54.55%	90.91%	72.73%	10.66%
S11	EEG Sonif. DWT	84.85%	93.94%	96.97%	87.88%	78.79%	88.48%	4.71%
S12	EEG Sonif. DWT	84.85%	60.61%	54.55%	69.70%	60.61%	66.06%	10.24%
S13	EEG Sonif. DWT	81.82%	54.55%	84.85%	75.76%	75.76%	74.55%	8.95%
S14	EEG Sonif. DWT	66.67%	42.42%	57.58%	30.30%	36.36%	46.67%	13.74%
S15	EEG Sonif. DWT	63.64%	72.73%	60.61%	72.73%	93.94%	72.73%	8.84%
S16	EEG Sonif. DWT	69.70%	63.64%	24.24%	45.46%	48.49%	50.30%	13.13%
S17	EEG Sonif. MFCC	90.91%	57.58%	78.79%	51.52%	72.73%	70.30%	10.20%
S18	EEG Sonif. DWT	75.76%	39.39%	87.88%	66.67%	78.79%	69.70%	9.77%
S19	EEG DWT	93.94%	42.42%	63.64%	27.27%	36.36%	52.73%	13.11%
S20	EEG Sonif. DWT	87.88%	72.73%	75.76%	75.76%	69.70%	76.36%	8.64%
S21	EEG Sonif. DWT	72.73%	21.21%	39.39%	42.42%	75.76%	50.30%	11.98%
S22	EEG Sonif. DWT	100.00%	60.61%	54.55%	75.76%	69.70%	72.12%	8.35%
S23	EEG Sonif. DWT	48.49%	60.61%	54.55%	72.73%	75.76%	62.42%	11.35%
S24	EEG Sonif. DWT	96.97%	75.76%	48.49%	39.39%	81.82%	68.49%	9.74%
S25	EEG DWT	6.06%	9.09%	75.76%	63.64%	63.64%	43.64%	16.06%
S26	EEG Sonif. MFCC	63.64%	78.79%	57.58%	39.39%	84.85%	64.85%	11.16%
S27	EEG Sonif. MFCC	57.93%	70.12%	76.22%	64.02%	57.93%	65.24%	11.10%
	Promedio	76.35%	61.65%	60.41%	58.28%	68.60%	65.06%	10.51%

Figura 5.27: Mejores porcentajes de exactitud por método en las 5 palabras usando 14 canales.

al trabajo [100], en la figura 5.30 se muestra la energía relativa promedio por palabra mostrando por separando cada nivel obtenido al aplicar la transformada wavelet. Se muestran los 4 niveles de detalle y 1 de aproximación del método de EEG sin sonificar y los 6 niveles de detalle y 1 nivel de aproximación del método de sonificación de EEG. Cada nivel de descomposición representa a un rango de frecuencias (mostrados previamente en la figura 4.7). Los porcentajes mostrados corresponden al promedio de los 27 sujetos bajo el enfoque de 14 canales.

En la gráfica correspondiente a la descomposición del método de EEG se muestra que la distribución de la energía es prácticamente nula en los niveles D2 al D5, mientras que la energía está concentrada en el nivel de detalle A5. También se muestra que la variación entre palabras es muy baja. Por otra parte en la gráfica correspondiente al método de sonificación de EEG la energía relativa está ligeramente más distribuida entre los niveles de descomposición aunque la mayor parte se sigue concentrando en el nivel de aproximación del último nivel (A6) correspondiente a las frecuencias bajas de

Suj.	Método	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
S1	EEG MFCC	48.49%	78.79%	15.15%	36.36%	75.76%	50.91%	12.72%
S2	EEG MFCC	39.39%	24.24%	45.46%	18.18%	21.21%	29.70%	15.87%
S3	EEG DWT	57.58%	18.18%	51.52%	27.27%	42.42%	39.39%	14.44%
S4	EEG MFCC	57.58%	45.46%	39.39%	51.52%	27.27%	44.24%	13.21%
S5	EEG MFCC	66.67%	42.42%	42.42%	36.36%	54.55%	48.48%	12.96%
S6	EEG Sonif. DWT	42.42%	18.18%	36.36%	21.21%	36.36%	30.91%	15.57%
S7	EEG MFCC	73.17%	27.44%	27.44%	21.34%	42.68%	38.41%	14.34%
S8	EEG MFCC	18.18%	30.30%	39.39%	66.67%	48.49%	40.61%	14.37%
S9	EEG MFCC	48.49%	48.49%	51.52%	48.49%	60.61%	51.52%	12.14%
S10	EEG MFCC	42.42%	9.09%	15.15%	36.36%	42.42%	29.09%	15.89%
S11	EEG MFCC	69.70%	84.85%	51.52%	45.46%	60.61%	62.42%	9.96%
S12	EEG DWT	81.82%	21.21%	36.36%	30.30%	78.79%	49.70%	12.89%
S13	EEG MFCC	39.39%	27.27%	21.21%	21.21%	54.55%	32.73%	14.75%
S14	EEG MFCC	18.18%	42.42%	30.30%	6.06%	24.24%	24.24%	16.55%
S15	EEG MFCC	39.39%	45.46%	42.42%	24.24%	48.49%	40.00%	14.64%
S16	EEG MFCC	45.18%	30.12%	18.07%	36.15%	42.17%	34.34%	15.01%
S17	EEG MFCC	21.21%	33.33%	30.30%	33.33%	36.36%	30.91%	15.99%
S18	EEG MFCC	54.55%	36.36%	54.55%	24.24%	60.61%	46.06%	13.49%
S19	EEG MFCC	24.24%	30.30%	6.06%	6.06%	30.30%	19.39%	17.24%
S20	EEG MFCC	48.49%	66.67%	51.52%	36.36%	45.46%	49.70%	12.76%
S21	EEG MFCC	24.24%	21.21%	36.36%	12.12%	24.24%	23.64%	15.83%
S22	EEG MFCC	78.79%	45.46%	30.30%	24.24%	42.42%	44.24%	13.47%
S23	EEG MFCC	39.39%	18.18%	39.39%	39.39%	39.39%	35.15%	14.33%
S24	EEG MFCC	60.24%	15.06%	33.13%	27.11%	24.10%	31.93%	15.88%
S25	EEG MFCC	12.12%	0.00%	42.42%	39.39%	21.21%	23.03%	19.00%
S26	EEG DWT	30.30%	39.39%	33.33%	45.46%	57.58%	41.21%	14.23%
S27	EEG MFCC	42.68%	24.39%	30.49%	48.78%	18.29%	32.93%	15.69%
	Promedio	45.34%	34.23%	35.24%	31.99%	42.98%	37.96%	14.56%

Figura 5.28: Porcentajes de exactitud más bajos por método en las 5 palabras usando 4 canales.

las señales de audio. La energía relativa de los niveles D₁ y D₂ correspondientes a las frecuencias altas de la señal de audio un poco menor a la que se muestra en los niveles D₃ y D₄ mientras que la energía en el nivel de descomposición D₅ es prácticamente nula. Los niveles D₄ y A₆ son los que más variaciones presentan entre palabras, por lo que esto hace suponer que la información proporcionada al clasificador por estos niveles es de mayor utilidad.

Al comparar ambas gráficas apreciamos el efecto que tiene aplicar la técnica de sonificación de EEG en la distribución de la energía relativa por niveles. Con lo que se puede sugerir que la técnica de sonificación de EEG ayuda a representar la energía relativa por niveles al tener control de un rango de frecuencias mas amplio que incluye la selección previa de las frecuencias dominantes, tornándose en posibles aspectos que mejoran los porcentajes de clasificación.

Suj.	Método	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
S1	EEG MFCC	75.76%	81.82%	18.18%	27.27%	48.49%	50.30%	12.33%
S2	EEG MFCC	33.33%	15.15%	54.55%	24.24%	24.24%	30.30%	15.79%
S3	EEG MFCC	72.29%	24.10%	33.13%	27.11%	33.13%	37.95%	14.46%
S4	EEG MFCC	69.70%	69.70%	33.33%	36.36%	45.46%	50.91%	12.18%
S5	EEG MFCC	30.30%	87.88%	30.30%	36.36%	57.58%	48.49%	12.68%
S6	EEG MFCC	42.42%	42.42%	33.33%	21.21%	27.27%	33.33%	15.76%
S7	EEG MFCC	42.68%	36.59%	33.54%	36.59%	48.78%	39.63%	14.76%
S8	EEG MFCC	54.55%	27.27%	30.30%	51.52%	36.36%	40.00%	12.79%
S9	EEG MFCC	39.39%	45.46%	39.39%	3.03%	48.49%	35.15%	15.66%
S10	EEG MFCC	51.52%	54.55%	33.33%	18.18%	15.15%	34.55%	15.52%
S11	EEG MFCC	57.58%	84.85%	48.49%	24.24%	51.52%	53.33%	11.44%
S12	EEG MFCC	50.48%	26.44%	14.42%	33.65%	74.52%	39.90%	14.05%
S13	EEG MFCC	57.58%	9.09%	30.30%	27.27%	21.21%	29.09%	15.97%
S14	EEG MFCC	45.46%	18.18%	30.30%	21.21%	24.24%	27.88%	16.46%
S15	EEG MFCC	33.33%	24.24%	36.36%	45.46%	30.30%	33.94%	15.61%
S16	EEG MFCC	39.16%	24.10%	30.12%	60.24%	12.05%	33.13%	15.10%
S17	EEG MFCC	39.39%	48.49%	45.46%	21.21%	48.49%	40.61%	14.34%
S18	EEG MFCC	24.24%	18.18%	69.70%	45.46%	42.42%	40.00%	14.35%
S19	EEG MFCC	27.27%	30.30%	24.24%	12.12%	15.15%	21.82%	17.08%
S20	EEG MFCC	66.67%	57.58%	30.30%	30.30%	6.06%	38.18%	14.79%
S21	EEG MFCC	30.30%	24.24%	24.24%	36.36%	33.33%	29.70%	15.93%
S22	EEG MFCC	100.00%	39.39%	45.46%	30.30%	39.39%	50.91%	11.70%
S23	EEG MFCC	33.33%	36.36%	27.27%	45.46%	33.33%	35.15%	15.07%
S24	EEG MFCC	48.19%	24.10%	24.10%	6.02%	33.13%	27.11%	16.31%
S25	EEG MFCC	12.12%	3.03%	33.33%	72.73%	36.36%	31.52%	17.87%
S26	EEG MFCC	44.64%	11.91%	47.62%	11.91%	38.69%	30.95%	16.07%
S27	EEG MFCC	27.44%	15.24%	18.29%	36.59%	48.78%	29.27%	15.68%
	Promedio	46.26%	36.32%	34.05%	31.20%	36.07%	36.78%	14.81%

Figura 5.29: Porcentajes de exactitud más bajos por método en las 5 palabras usando 14 canales.

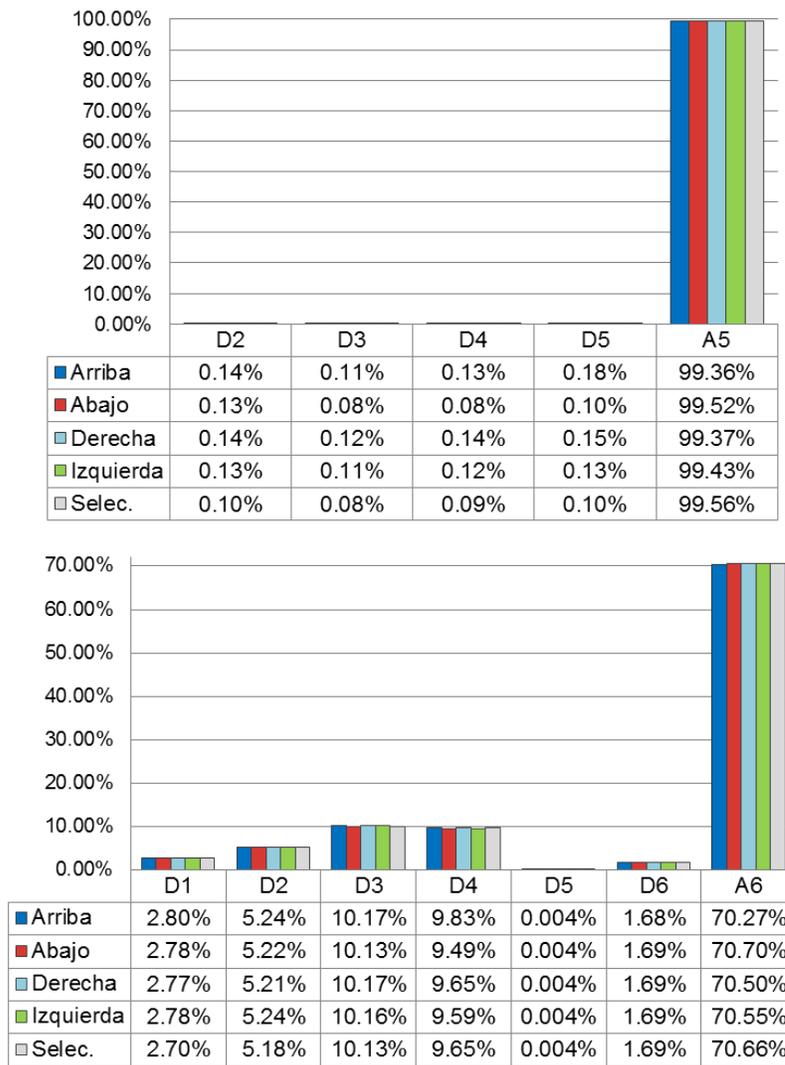


Figura 5.30: Energía relativa Wavelet promedio de los 27 sujeto usando 14 canales. Se muestra la distribución de la energía por niveles de descomposición para cada palabra para los métodos de EEG (arriba) y de sonificación de EEG (abajo).

CONCLUSIONES Y TRABAJO FUTURO

6.1. Conclusiones

Este trabajo de tesis aporta resultados novedosos sobre la clasificación de palabras imaginadas provenientes de audios obtenidos de aplicar sonificación de EEG. Se explora usar una representación distinta de la señal de EEG para clasificación automática.

En esta investigación se muestra que la clasificación de la señal de EEG aplicando la sonificación, mediante la elección de las frecuencias dominantes del espectrograma de la señal de EEG y el mapeo de las frecuencias de EEG a frecuencias del audio, logró resaltar patrones de la señal que ayudaron a tener mejores resultados de la correcta clasificación promedio para los 27 sujetos (55.83 % para 4 canales y 63.83 % para 14 canales), esto bajo la limitante de no usar aspectos de la metodología iguales a los del trabajo con el cual mayormente se comparó [83].

Basados en los resultados obtenidos para los 27 sujetos se concluye que, aplicar sonificación de EEG como paso previo a la extracción de características por ambos métodos DWT y MFCC se mejoran ligeramente los porcentajes de clasificación de palabras imaginadas respecto a no usar sonificación de EEG. Esto acotado con el resultado de la prueba estadística realizada, que nos indica que la diferencia es significativa solo cuando se compara con el método de EEG con MFCC.

Los porcentajes de exactitud en la clasificación promedio obtenidos de los 27 sujetos superan a los reportados en los trabajos [15,16] (47.24 % y 45.5 %). En dichos trabajos se clasificaban 5 palabras del habla imaginada usando 16 canales y con menor número de sujetos (6 y 18 respectivamente), comparando esos resultados con el resultado obtenido en esta investigación con 14 canales muestra que se supera de 47.24 % a 63.83 %.

La diferencia entre los porcentajes entre EEG y EEG sonificado son más notorias al usar 4 canales que al usar 14, pues para 4 canales la diferencia es de 7.73 %, mientras

que para 14 canales la diferencia es de 4.37 %, por lo que se puede decir que el método de sonificación de EEG sirve para clasificar mejor cuando hay poca información.

El mejor resultado de clasificación obtenido que un sujeto obtuvo fue de 83.03 % y 88.48 % para 4 y 14 canales respectivamente, lo cual sugiere que si existe información suficiente en la señal cerebral para que por medio de técnicas como las aquí propuestas se puedan tener buenos porcentajes de clasificación. Esto resulta alentador pues da un pequeño paso encaminado a lograr tener alguna BCI con la que cualquier persona que genere las señales cerebrales adecuadas, pueda usar una BCI para transmitir mensaje basados en el habla imaginada. Aunque se debe señalar que el método aún tiene mucha variabilidad entre sujetos.

Evaluando el enfoque por palabra, se concluye que, aunque varían los porcentajes de cada sujeto por palabra, existen palabras que en promedio son más difíciles de reconocer y otras que son más fáciles.

Aunque no fue el objetivo de la tesis, el obtener el comportamiento de los parámetros de la sonificación puede apoyar en la construcción de audios. Donde además de tener buena exactitud de clasificación también se aporte información útil para distinguir distintas señales con solo escucharlas. Esto muestra una ventaja con respecto a procesar las señales de EEG sin sonificación.

Cuando se extraen características por MFCC se muestra que los porcentajes de clasificación son relativamente mejores cuando se incluye información temporal usando los coeficientes delta, por lo que se conjetura que se podría experimentar con algún método que caracterice información sobre secuencia de eventos en la señal.

Se muestra también que los porcentajes de reconocimiento por palabra tienen diferentes porcentajes de clasificación al pasar de 4 a 14 canales, con ello se afirma que distintas palabras activan distintas regiones del cerebro, con lo que podría ser necesario no solo tener un método que permita escoger los mejores canales por sujeto si no también uno que permita escoger los mejores canales por palabra de cada sujeto.

6.2. Trabajo futuro

Como se describió previamente un sujeto obtiene un mejor porcentaje de clasificación usando alguna configuración que no es necesariamente la mejor para todos los sujetos. Entonces se podría usar un algoritmo de selección de modelo para optimizar los parámetros de la sonificación, extracción de características y clasificación, los cuales otorguen la mejor exactitud de clasificación de cada sujeto.

La investigación puede ser extendida aplicando los métodos aquí propuestos en

problemas que involucren el procesamiento y clasificación de la señal cerebral que no sea el habla imaginada, pues se espera que funcionen con la misma eficacia que lo aquí mostrado. Además que en problemas donde existen solo dos o tres clases (por ejemplo el diagnóstico de paciente sano, con etapa temprana de Alzheimer o con etapa avanzada de Alzheimer) se espera que incluso pueda mejorar al ser un problema teóricamente más sencillo para un clasificador automático. También se necesita probar si el desempeño de los métodos aquí mostrados se mantiene con un corpus de datos de un conjunto mas extenso de palabras, sílabas o fonemas del idioma español.

En la etapa de adquisición de datos es conveniente probar los algoritmos aquí ejecutados con datos de grabaciones usando otros dispositivos de EEG, los cuales tengan mayor cantidad de electrodos y una mejor frecuencia de muestreo mayor a 128 Hz, aunque para grabaciones de EEG provenientes de más canales se tendrían que aplicar métodos de reducción a la información y características extraídas de los canales.

En la etapa de montaje se puede probar otros montajes o buscar aplicar iterativamente la referencia promedio común para limpiar mejor la señal y así obtener mejores características, que ayuden a mejorar la clasificación de las palabras.

En la fase de sonificación de EEG, se puede experimentar con otros métodos como la sonificación basada en modelado de elipsoides, la sonificación vocal o la sonificación orquestal, los cuales podrían mejorar algunas características de los audios, por ejemplo en la técnica de modelado de bumps el cambio de tonos a notas musicales podría mejorar la caracterización de las frecuencias además que se podría tener una variación de intensidad que ayude a la clasificación. En la sonificación vocal los audios generados permitirían tener características muy similares a la voz humana, con lo que métodos de extracción de características como MFCC teóricamente funcionarían mejor. Por último en la sonificación orquestal, se podría diseñar un esquema por regiones como los basados en filtrado espacial [84, 101, 102], para que cada región de electrodos del EEG tenga la sonificación basada en un instrumento distinto, con ello se podría tener un audio mejor caracterizado para el oído humano y a la vez podría ayudar a mejorar la discriminación en la etapa de clasificación.

Los resultados obtenidos dan pie a realizar pruebas con otros métodos de extracción de características del audio, pues como se mostró, los dos métodos con los que se experimentó aquí los resultados son muy variantes entre sujetos, y podría existir alguno otro que logre obtener mejores resultados. En la fase de clasificación, se podría probar con un enfoque de selección de modelo el cual permita escoger los clasificadores que optimizan los porcentajes de exactitud alcanzados por cada sujeto individualmente.

PUBLICACIONES

A partir del presente trabajo de tesis se publicaron los siguientes trabajos en forma de presentación oral y en forma de artículo.

A.1. Presentación

- Analysis of sonification for EEG classification, *XX Asamblea general del ISTECS*, 2014.

A.2. Artículos

- Sonificación de EEG para la clasificación de palabras no pronunciadas, *Research in Computing Science*, no. 74, pp. 61–72, 2014.
- Reconocimiento de patrones para clasificación de palabras no pronunciadas a partir de EEG sonificados, *2° Seminario de Inteligencia Computacional (SNAIC)*, 2014.

A.3. Capítulo en libro

- A. A. Torres-García, E. F. González-Castañeda, C. A. Reyes-García, L. Villaseñor-Pineda. Hacia la decodificación de habla imaginada registrada mediante electroencefalogramas, *Mente-Cerebro: Diálogos entre Ciencia y Filosofía*, UPAEP, 2015 (en proceso de publicación).

DETALLES DE LOS EXPERIMENTOS

B.1. Matrices de confusión de todos sujetos por palabra

En esta sección se muestran los porcentajes de acierto que tuvo la clasificación con cada uno de los cuatro enfoques, usando 4 y 14 canales.

El porcentaje que aparece en cada palabra, nos indica cuantas veces de las 33 repeticiones el clasificador eligió correctamente la clase a la que pertenece la palabra, en la ultima columna de cada tabla se muestra el promedio de clasificación de las 5 palabras. Los datos fueron obtenidos a partir de las 216 matrices de confusión obtenidas al clasificar con los 4 métodos para los 27 sujetos en los 2 enfoques de elección de canales.

S1 - 4 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio
EEG Normal DWT	87.88%	75.76%	51.52%	48.49%	66.67%	66.06%
EEG Sonif. DWT	93.94%	84.85%	57.58%	42.42%	72.73%	70.30%
EEG Normal MFCC	48.49%	78.79%	15.15%	36.36%	75.76%	50.91%
EEG Sonif. MFCC	87.88%	72.73%	36.36%	36.36%	63.64%	59.39%

S1 - 14 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio
EEG Normal DWT	93.94%	66.67%	66.67%	54.55%	87.88%	73.94%
EEG Sonif. DWT	96.97%	84.85%	60.61%	57.58%	81.82%	76.36%
EEG Normal MFCC	75.76%	81.82%	18.18%	27.27%	48.49%	50.30%
EEG Sonif. MFCC	100.00%	81.82%	48.49%	36.36%	72.73%	67.88%

Figura B.1: Promedios de clasificación por palabra para el sujeto 1.

S2 - 4 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	66.67%	15.15%	36.36%	18.18%	54.55%	38.18%	14.44%
EEG Sonif. DWT	75.76%	39.39%	45.46%	6.06%	33.33%	40.00%	14.73%
EEG Normal MFCC	39.39%	24.24%	45.46%	18.18%	21.21%	29.70%	15.87%
EEG Sonif. MFCC	63.64%	9.09%	42.42%	42.42%	60.61%	43.64%	13.79%

S2 - 14 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	69.70%	30.30%	57.58%	18.18%	36.36%	42.42%	13.98%
EEG Sonif. DWT	72.73%	30.30%	54.55%	24.24%	42.42%	44.85%	13.90%
EEG Normal MFCC	33.33%	15.15%	54.55%	24.24%	24.24%	30.30%	15.79%
EEG Sonif. MFCC	75.76%	33.33%	54.55%	27.27%	36.36%	45.45%	13.45%

Figura B.2: Promedios de clasificación por palabra para el sujeto 2.

S3 - 4 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	57.58%	18.18%	51.52%	27.27%	42.42%	39.39%	14.44%
EEG Sonif. DWT	90.91%	33.33%	60.61%	33.33%	60.61%	55.76%	11.05%
EEG Normal MFCC	72.29%	30.12%	36.15%	75.30%	42.17%	51.20%	13.09%
EEG Sonif. MFCC	84.85%	33.33%	48.49%	33.33%	39.39%	47.88%	12.99%

S3 - 14 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	81.82%	54.55%	45.46%	48.49%	84.85%	63.03%	10.50%
EEG Sonif. DWT	87.88%	30.30%	63.64%	54.55%	72.73%	61.82%	9.85%
EEG Normal MFCC	72.29%	24.10%	33.13%	27.11%	33.13%	37.95%	14.46%
EEG Sonif. MFCC	93.94%	36.36%	51.52%	48.49%	81.82%	62.42%	10.77%

Figura B.3: Promedios de clasificación por palabra para el sujeto 3.

S4 - 4 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	57.58%	45.46%	33.33%	51.52%	45.46%	46.67%	13.28%
EEG Sonif. DWT	75.76%	57.58%	42.42%	69.70%	45.46%	58.18%	10.74%
EEG Normal MFCC	57.58%	45.46%	39.39%	51.52%	27.27%	44.24%	13.21%
EEG Sonif. MFCC	63.64%	60.61%	39.39%	60.61%	54.55%	55.76%	11.76%

S4 - 14 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	84.85%	72.73%	48.49%	45.46%	48.49%	60.00%	11.24%
EEG Sonif. DWT	93.94%	78.79%	48.49%	75.76%	54.55%	70.30%	8.71%
EEG Normal MFCC	69.70%	69.70%	33.33%	36.36%	45.46%	50.91%	12.18%
EEG Sonif. MFCC	93.94%	78.79%	39.39%	51.52%	72.73%	67.27%	10.04%

Figura B.4: Promedios de clasificación por palabra para el sujeto 4.

S5 - 4 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	78.79%	60.61%	48.49%	33.33%	87.88%	61.82%	9.95%
EEG Sonif. DWT	87.88%	81.82%	63.64%	57.58%	84.85%	75.15%	8.22%
EEG Normal MFCC	66.67%	42.42%	42.42%	36.36%	54.55%	48.48%	12.96%
EEG Sonif. MFCC	90.91%	69.70%	51.52%	60.61%	84.85%	71.52%	9.46%

S5 - 14 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	78.79%	78.79%	54.55%	54.55%	93.94%	72.12%	9.80%
EEG Sonif. DWT	84.85%	78.79%	60.61%	72.73%	90.91%	77.58%	8.94%
EEG Normal MFCC	30.30%	87.88%	30.30%	36.36%	57.58%	48.49%	12.68%
EEG Sonif. MFCC	81.82%	81.82%	51.52%	69.70%	87.88%	74.55%	9.81%

Figura B.5: Promedios de clasificación por palabra para el sujeto 5.

S6 - 4 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	33.33%	33.33%	36.36%	27.27%	27.27%	31.52%	15.24%
EEG Sonif. DWT	42.42%	18.18%	36.36%	21.21%	36.36%	30.91%	15.57%
EEG Normal MFCC	39.39%	51.52%	27.27%	30.30%	15.15%	32.73%	15.85%
EEG Sonif. MFCC	48.49%	36.36%	24.24%	21.21%	27.27%	31.52%	14.69%

S6 - 14 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	51.52%	36.36%	39.39%	24.24%	30.30%	36.36%	14.54%
EEG Sonif. DWT	54.55%	45.46%	24.24%	36.36%	39.39%	40.00%	14.17%
EEG Normal MFCC	42.42%	42.42%	33.33%	21.21%	27.27%	33.33%	15.76%
EEG Sonif. MFCC	45.46%	57.58%	45.46%	24.24%	18.18%	38.18%	14.47%

Figura B.6: Promedios de clasificación por palabra para el sujeto 6.

S7 - 4 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	70.12%	33.54%	42.68%	48.78%	48.78%	48.78%	13.15%
EEG Sonif. DWT	70.12%	48.78%	42.68%	51.83%	48.78%	52.44%	12.15%
EEG Normal MFCC	73.17%	27.44%	27.44%	21.34%	42.68%	38.41%	14.34%
EEG Sonif. MFCC	70.12%	45.73%	30.49%	36.59%	51.83%	46.95%	12.95%

S7 - 14 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	67.07%	60.98%	48.78%	48.78%	57.93%	56.71%	11.41%
EEG Sonif. DWT	64.02%	73.17%	54.88%	67.07%	45.73%	60.98%	11.30%
EEG Normal MFCC	42.68%	36.59%	33.54%	36.59%	48.78%	39.63%	14.76%
EEG Sonif. MFCC	82.32%	67.07%	48.78%	57.93%	60.98%	63.41%	11.48%

Figura B.7: Promedios de clasificación por palabra para el sujeto 7.

S8 - 4 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	69.70%	51.52%	45.46%	75.76%	87.88%	66.06%	9.46%
EEG Sonif. DWT	72.73%	60.61%	57.58%	81.82%	93.94%	73.33%	8.01%
EEG Normal MFCC	18.18%	30.30%	39.39%	66.67%	48.49%	40.61%	14.37%
EEG Sonif. MFCC	69.70%	57.58%	27.27%	90.91%	93.94%	67.88%	9.39%

S8 - 14 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	81.82%	69.70%	81.82%	84.85%	93.94%	82.42%	7.74%
EEG Sonif. DWT	93.94%	81.82%	78.79%	84.85%	96.97%	87.27%	6.20%
EEG Normal MFCC	54.55%	27.27%	30.30%	51.52%	36.36%	40.00%	12.79%
EEG Sonif. MFCC	87.88%	72.73%	81.82%	84.85%	96.97%	84.85%	7.33%

Figura B.8: Promedios de clasificación por palabra para el sujeto 8.

S9 - 4 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	60.61%	60.61%	63.64%	51.52%	57.58%	58.79%	11.35%
EEG Sonif. DWT	63.64%	66.67%	72.73%	42.42%	57.58%	60.61%	10.90%
EEG Normal MFCC	48.49%	48.49%	51.52%	48.49%	60.61%	51.52%	12.14%
EEG Sonif. MFCC	69.70%	66.67%	69.70%	36.36%	69.70%	62.42%	11.20%

S9 - 14 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	63.64%	66.67%	57.58%	63.64%	69.70%	64.24%	11.05%
EEG Sonif. DWT	69.70%	81.82%	63.64%	57.58%	54.55%	65.45%	10.59%
EEG Normal MFCC	39.39%	45.46%	39.39%	3.03%	48.49%	35.15%	15.66%
EEG Sonif. MFCC	69.70%	69.70%	66.67%	39.39%	72.73%	63.64%	11.15%

Figura B.9: Promedios de clasificación por palabra para el sujeto 9.

S10 - 4 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	60.61%	27.27%	27.27%	27.27%	63.64%	41.21%	14.36%
EEG Sonif. DWT	36.36%	27.27%	48.49%	39.39%	90.91%	48.49%	13.28%
EEG Normal MFCC	42.42%	9.09%	15.15%	36.36%	42.42%	29.09%	15.89%
EEG Sonif. MFCC	51.52%	27.27%	33.33%	27.27%	87.88%	45.45%	13.82%

S10 - 14 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	66.67%	75.76%	21.21%	42.42%	81.82%	57.58%	11.89%
EEG Sonif. DWT	60.61%	78.79%	48.49%	60.61%	87.88%	67.27%	10.61%
EEG Normal MFCC	51.52%	54.55%	33.33%	18.18%	15.15%	34.55%	15.52%
EEG Sonif. MFCC	81.82%	81.82%	54.55%	54.55%	90.91%	72.73%	10.66%

Figura B.10: Promedios de clasificación por palabra para el sujeto 10.

S11 - 4 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	51.52%	63.64%	100.00%	63.64%	87.88%	73.33%	7.79%
EEG Sonif. DWT	72.73%	81.82%	100.00%	78.79%	81.82%	83.03%	5.37%
EEG Normal MFCC	69.70%	84.85%	51.52%	45.46%	60.61%	62.42%	9.96%
EEG Sonif. MFCC	75.76%	93.94%	84.85%	60.61%	81.82%	79.39%	6.23%

S11 - 14 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	72.73%	75.76%	96.97%	84.85%	87.88%	83.64%	5.82%
EEG Sonif. DWT	84.85%	93.94%	96.97%	87.88%	78.79%	88.48%	4.71%
EEG Normal MFCC	57.58%	84.85%	48.49%	24.24%	51.52%	53.33%	11.44%
EEG Sonif. MFCC	81.82%	90.91%	90.91%	72.73%	81.82%	83.64%	5.66%

Figura B.11: Promedios de clasificación por palabra para el sujeto 11.

S12 - 4 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	81.82%	21.21%	36.36%	30.30%	78.79%	49.70%	12.89%
EEG Sonif. DWT	66.67%	48.49%	51.52%	57.58%	72.73%	59.39%	10.61%
EEG Normal MFCC	62.50%	19.23%	28.85%	55.29%	84.14%	50.00%	13.30%
EEG Sonif. MFCC	66.67%	51.52%	27.27%	48.49%	84.85%	55.76%	11.58%

S12 - 14 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	84.85%	45.46%	42.42%	60.61%	66.67%	60.00%	11.57%
EEG Sonif. DWT	84.85%	60.61%	54.55%	69.70%	60.61%	66.06%	10.24%
EEG Normal MFCC	50.48%	26.44%	14.42%	33.65%	74.52%	39.90%	14.05%
EEG Sonif. MFCC	72.73%	48.49%	45.46%	57.58%	84.85%	61.82%	11.76%

Figura B.12: Promedios de clasificación por palabra para el sujeto 12.

S13 - 4 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	60.61%	9.09%	78.79%	36.36%	60.61%	49.09%	11.93%
EEG Sonif. DWT	75.76%	42.42%	66.67%	69.70%	48.49%	60.61%	10.30%
EEG Normal MFCC	39.39%	27.27%	21.21%	21.21%	54.55%	32.73%	14.75%
EEG Sonif. MFCC	57.58%	33.33%	81.82%	54.55%	66.67%	58.79%	11.56%

S13 - 14 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	54.55%	33.33%	84.85%	45.46%	60.61%	55.76%	11.40%
EEG Sonif. DWT	81.82%	54.55%	84.85%	75.76%	75.76%	74.55%	8.95%
EEG Normal MFCC	57.58%	9.09%	30.30%	27.27%	21.21%	29.09%	15.97%
EEG Sonif. MFCC	84.85%	30.30%	81.82%	54.55%	69.70%	64.24%	10.95%

Figura B.13: Promedios de clasificación por palabra para el sujeto 13.

S14 - 4 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	48.49%	30.30%	33.33%	39.39%	12.12%	32.73%	15.40%
EEG Sonif. DWT	63.64%	39.39%	63.64%	21.21%	21.21%	41.82%	14.30%
EEG Normal MFCC	18.18%	42.42%	30.30%	6.06%	24.24%	24.24%	16.55%
EEG Sonif. MFCC	54.55%	33.33%	57.58%	9.09%	3.03%	31.52%	15.21%

S14 - 14 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	60.61%	48.49%	42.42%	42.42%	27.27%	44.24%	14.26%
EEG Sonif. DWT	66.67%	42.42%	57.58%	30.30%	36.36%	46.67%	13.74%
EEG Normal MFCC	45.46%	18.18%	30.30%	21.21%	24.24%	27.88%	16.46%
EEG Sonif. MFCC	48.49%	39.39%	36.36%	24.24%	15.15%	32.73%	14.78%

Figura B.14: Promedios de clasificación por palabra para el sujeto 14.

S15 - 4 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	24.24%	60.61%	63.64%	54.55%	93.94%	59.39%	10.99%
EEG Sonif. DWT	42.42%	63.64%	72.73%	48.49%	96.97%	64.85%	9.97%
EEG Normal MFCC	39.39%	45.46%	42.42%	24.24%	48.49%	40.00%	14.64%
EEG Sonif. MFCC	54.55%	60.61%	48.49%	54.55%	93.94%	62.42%	11.15%

S15 - 14 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	54.55%	51.52%	60.61%	63.64%	87.88%	63.64%	10.41%
EEG Sonif. DWT	63.64%	72.73%	60.61%	72.73%	93.94%	72.73%	8.84%
EEG Normal MFCC	33.33%	24.24%	36.36%	45.46%	30.30%	33.94%	15.61%
EEG Sonif. MFCC	60.61%	63.64%	57.58%	54.55%	93.94%	66.06%	11.26%

Figura B.15: Promedios de clasificación por palabra para el sujeto 15.

S16 - 4 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	60.61%	30.30%	15.15%	30.30%	51.52%	37.58%	14.29%
EEG Sonif. DWT	78.79%	39.39%	24.24%	30.30%	57.58%	46.06%	13.18%
EEG Normal MFCC	45.18%	30.12%	18.07%	36.15%	42.17%	34.34%	15.01%
EEG Sonif. MFCC	63.64%	42.42%	27.27%	33.33%	51.52%	43.64%	14.11%

S16 - 14 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	51.52%	63.64%	27.27%	36.36%	48.49%	45.45%	12.93%
EEG Sonif. DWT	69.70%	63.64%	24.24%	45.46%	48.49%	50.30%	13.13%
EEG Normal MFCC	39.16%	24.10%	30.12%	60.24%	12.05%	33.13%	15.10%
EEG Sonif. MFCC	51.52%	60.61%	27.27%	30.30%	48.49%	43.64%	13.81%

Figura B.16: Promedios de clasificación por palabra para el sujeto 16.

S17 - 4 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	75.76%	36.36%	54.55%	39.39%	42.42%	49.70%	12.45%
EEG Sonif. DWT	90.91%	39.39%	51.52%	42.42%	54.55%	55.76%	11.48%
EEG Normal MFCC	21.21%	33.33%	30.30%	33.33%	36.36%	30.91%	15.99%
EEG Sonif. MFCC	75.76%	36.36%	60.61%	48.49%	51.52%	54.55%	11.40%

S17 - 14 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	84.85%	42.42%	66.67%	39.39%	69.70%	60.61%	10.61%
EEG Sonif. DWT	87.88%	51.52%	63.64%	60.61%	81.82%	69.09%	9.68%
EEG Normal MFCC	39.39%	48.49%	45.46%	21.21%	48.49%	40.61%	14.34%
EEG Sonif. MFCC	90.91%	57.58%	78.79%	51.52%	72.73%	70.30%	10.20%

Figura B.17: Promedios de clasificación por palabra para el sujeto 17.

S18 - 4 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	54.55%	39.39%	54.55%	72.73%	72.73%	58.79%	11.60%
EEG Sonif. DWT	66.67%	24.24%	57.58%	51.52%	72.73%	54.55%	11.53%
EEG Normal MFCC	54.55%	36.36%	54.55%	24.24%	60.61%	46.06%	13.49%
EEG Sonif. MFCC	48.49%	21.21%	57.58%	42.42%	75.76%	49.09%	12.24%

S18 - 14 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	57.58%	57.58%	75.76%	78.79%	72.72%	68.48%	10.22%
EEG Sonif. DWT	75.76%	39.39%	87.88%	66.67%	78.79%	69.70%	9.77%
EEG Normal MFCC	24.24%	18.18%	69.70%	45.46%	42.42%	40.00%	14.35%
EEG Sonif. MFCC	60.61%	42.42%	81.82%	57.58%	78.79%	64.24%	10.73%

Figura B.18: Promedios de clasificación por palabra para el sujeto 18.

S19 - 4 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	33.33%	30.30%	21.21%	21.21%	33.33%	27.88%	15.99%
EEG Sonif. DWT	36.36%	21.21%	21.21%	33.33%	48.49%	32.12%	15.36%
EEG Normal MFCC	24.24%	30.30%	6.06%	6.06%	30.30%	19.39%	17.24%
EEG Sonif. MFCC	39.39%	24.24%	33.33%	24.24%	27.27%	29.70%	15.44%

S19 - 14 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	93.94%	42.42%	63.64%	27.27%	36.36%	52.73%	13.11%
EEG Sonif. DWT	87.88%	36.36%	42.42%	30.30%	39.39%	47.27%	12.80%
EEG Normal MFCC	27.27%	30.30%	24.24%	12.12%	15.15%	21.82%	17.08%
EEG Sonif. MFCC	96.97%	30.30%	42.42%	30.30%	36.36%	47.27%	13.16%

Figura B.19: Promedios de clasificación por palabra para el sujeto 19.

S20 - 4 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	72.73%	93.94%	72.73%	63.64%	75.76%	75.76%	7.40%
EEG Sonif. DWT	69.70%	84.85%	75.76%	69.70%	66.67%	73.33%	7.83%
EEG Normal MFCC	48.49%	66.67%	51.52%	36.36%	45.46%	49.70%	12.76%
EEG Sonif. MFCC	66.67%	84.85%	69.70%	84.85%	69.70%	75.15%	8.12%

S20 - 14 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	84.85%	87.88%	66.67%	72.73%	63.64%	75.15%	8.04%
EEG Sonif. DWT	87.88%	72.73%	75.76%	75.76%	69.70%	76.36%	8.64%
EEG Normal MFCC	66.67%	57.58%	30.30%	30.30%	6.06%	38.18%	14.79%
EEG Sonif. MFCC	87.88%	87.88%	72.73%	60.61%	63.64%	74.55%	8.81%

Figura B.20: Promedios de clasificación por palabra para el sujeto 20.

S21 - 4 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	27.27%	24.24%	39.39%	24.24%	42.42%	31.52%	15.67%
EEG Sonif. DWT	54.55%	45.46%	45.46%	27.27%	75.76%	49.70%	12.60%
EEG Normal MFCC	24.24%	21.21%	36.36%	12.12%	24.24%	23.64%	15.83%
EEG Sonif. MFCC	48.49%	33.33%	42.42%	27.27%	45.46%	39.39%	14.75%

S21 - 14 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	63.64%	33.33%	33.33%	33.33%	51.52%	43.03%	14.08%
EEG Sonif. DWT	72.73%	21.21%	39.39%	42.42%	75.76%	50.30%	11.98%
EEG Normal MFCC	30.30%	24.24%	24.24%	36.36%	33.33%	29.70%	15.93%
EEG Sonif. MFCC	60.61%	33.33%	39.39%	33.33%	60.61%	45.45%	13.73%

Figura B.21: Promedios de clasificación por palabra para el sujeto 21.

S22 - 4 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	96.97%	33.33%	21.21%	69.70%	54.55%	55.15%	10.40%
EEG Sonif. DWT	100.00%	48.49%	36.36%	87.88%	69.70%	68.49%	8.97%
EEG Normal MFCC	78.79%	45.46%	30.30%	24.24%	42.42%	44.24%	13.47%
EEG Sonif. MFCC	100.00%	30.30%	39.39%	87.88%	36.36%	58.79%	9.90%

S22 - 14 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	100.00%	54.55%	48.49%	75.76%	48.49%	65.45%	9.33%
EEG Sonif. DWT	100.00%	60.61%	54.55%	75.76%	69.70%	72.12%	8.35%
EEG Normal MFCC	100.00%	39.39%	45.46%	30.30%	39.39%	50.91%	11.70%
EEG Sonif. MFCC	100.00%	60.61%	48.49%	75.76%	54.55%	67.88%	9.74%

Figura B.22: Promedios de clasificación por palabra para el sujeto 22.

S23 - 4 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	45.46%	36.36%	57.58%	60.61%	36.36%	47.27%	13.73%
EEG Sonif. DWT	60.61%	54.55%	54.55%	60.61%	57.58%	57.58%	11.71%
EEG Normal MFCC	39.39%	18.18%	39.39%	39.39%	39.39%	35.15%	14.33%
EEG Sonif. MFCC	69.70%	39.39%	39.39%	72.73%	63.64%	56.97%	11.98%

S23 - 14 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	57.58%	39.39%	45.46%	54.55%	60.61%	51.52%	13.24%
EEG Sonif. DWT	48.49%	60.61%	54.55%	72.73%	75.76%	62.42%	11.35%
EEG Normal MFCC	33.33%	36.36%	27.27%	45.46%	33.33%	35.15%	15.07%
EEG Sonif. MFCC	69.70%	39.39%	39.39%	60.61%	66.67%	55.15%	12.28%

Figura B.23: Promedios de clasificación por palabra para el sujeto 23.

S24 - 4 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	21.21%	18.18%	18.18%	30.30%	78.79%	33.33%	14.70%
EEG Sonif. DWT	87.88%	75.76%	45.46%	18.18%	72.73%	60.00%	10.63%
EEG Normal MFCC	60.24%	15.06%	33.13%	27.11%	24.10%	31.93%	15.88%
EEG Sonif. MFCC	45.46%	42.42%	36.36%	27.27%	72.73%	44.85%	13.58%

S24 - 14 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	45.46%	30.30%	36.36%	48.49%	72.73%	46.67%	13.36%
EEG Sonif. DWT	96.97%	75.76%	48.49%	39.39%	81.82%	68.49%	9.74%
EEG Normal MFCC	48.19%	24.10%	24.10%	6.02%	33.13%	27.11%	16.31%
EEG Sonif. MFCC	75.76%	60.61%	27.27%	30.30%	69.70%	52.73%	12.97%

Figura B.24: Promedios de clasificación por palabra para el sujeto 24.

S25 - 4 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	9.09%	6.06%	60.61%	27.27%	39.39%	28.48%	17.32%
EEG Sonif. DWT	6.06%	9.09%	45.46%	54.55%	54.55%	33.94%	16.65%
EEG Normal MFCC	12.12%	0.00%	42.42%	39.39%	21.21%	23.03%	19.00%
EEG Sonif. MFCC	6.06%	9.09%	63.64%	36.36%	45.46%	32.12%	17.15%

S25 - 14 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	6.06%	9.09%	75.76%	63.64%	63.64%	43.64%	16.06%
EEG Sonif. DWT	6.06%	9.09%	48.49%	60.61%	60.61%	36.97%	16.67%
EEG Normal MFCC	12.12%	3.03%	33.33%	72.73%	36.36%	31.52%	17.87%
EEG Sonif. MFCC	3.03%	12.12%	60.61%	63.64%	48.49%	37.58%	17.00%

Figura B.25: Promedios de clasificación por palabra para el sujeto 25.

S26 - 4 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	30.30%	39.39%	33.33%	45.46%	57.58%	41.21%	14.23%
EEG Sonif. DWT	60.61%	33.33%	42.42%	27.27%	60.61%	44.85%	13.39%
EEG Normal MFCC	62.50%	17.86%	35.71%	35.71%	62.50%	42.86%	14.08%
EEG Sonif. MFCC	48.49%	48.49%	54.55%	39.39%	78.79%	53.94%	12.10%

S26 - 14 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	57.58%	42.42%	48.49%	60.61%	72.73%	56.36%	12.41%
EEG Sonif. DWT	72.73%	66.67%	48.49%	36.36%	78.79%	60.61%	11.93%
EEG Normal MFCC	44.64%	11.91%	47.62%	11.91%	38.69%	30.95%	16.07%
EEG Sonif. MFCC	63.64%	78.79%	57.58%	39.39%	84.85%	64.85%	11.16%

Figura B.26: Promedios de clasificación por palabra para el sujeto 26.

S27 - 4 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	45.73%	42.68%	60.98%	48.78%	48.78%	49.39%	12.71%
EEG Sonif. DWT	45.73%	33.54%	76.22%	54.88%	70.12%	56.10%	11.90%
EEG Normal MFCC	42.68%	24.39%	30.49%	48.78%	18.29%	32.93%	15.69%
EEG Sonif. MFCC	48.78%	48.78%	73.17%	42.68%	64.02%	55.49%	12.35%

S27 - 14 canales	Arriba	Abajo	Derecha	Izquierda	Seleccionar	Promedio	Desv.
EEG Normal DWT	51.83%	54.88%	60.98%	45.73%	45.73%	51.83%	11.80%
EEG Sonif. DWT	54.88%	67.07%	70.12%	48.78%	54.88%	59.15%	10.81%
EEG Normal MFCC	27.44%	15.24%	18.29%	36.59%	48.78%	29.27%	15.68%
EEG Sonif. MFCC	57.93%	70.12%	76.22%	64.02%	57.93%	65.24%	11.10%

Figura B.27: Promedios de clasificación por palabra para el sujeto 27.

RESULTADOS ADICIONALES

En este anexo se muestran los resultados de clasificación reportados en el artículo: *Reconocimiento de patrones para clasificación de palabras no pronunciadas a partir de EEG sonificados*. En el cual se probaron distintos clasificadores bajo los dos enfoques usando 4 canales (ver figura C.1) y 14 canales (ver figura C.2). De especial interés es el resultado obtenido al clasificar con 14 canales usando MFCC con el clasificador SVM, el cual es mayor al reportado usando DWT clasificando con Random Forest.

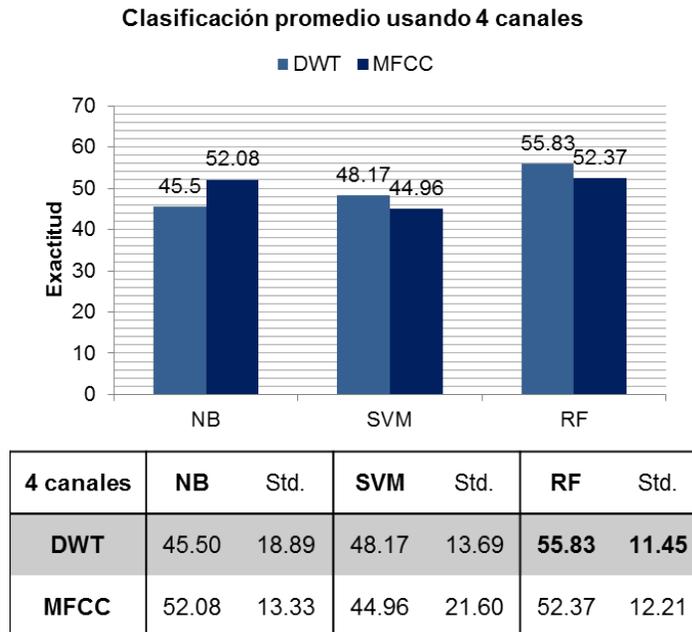
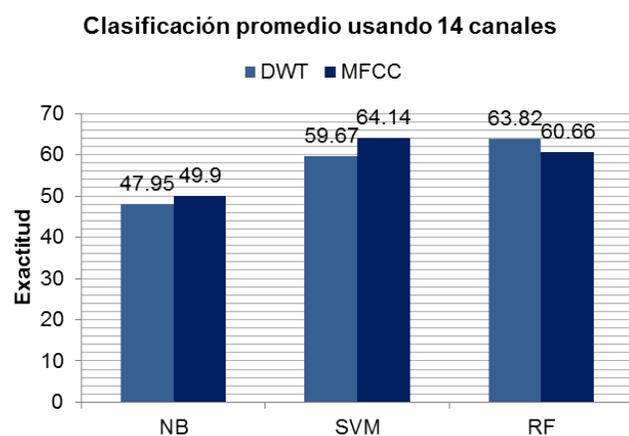


Figura C.1: Promedios en la exactitud de clasificación de los 27 sujetos usando distintos clasificadores con 4 canales.



14 canales	NB	Std.	SVM	Std.	RF	Std.
DWT	47.95	20.0	59.67	12.67	63.82	10.51
MFCC	49.90	19.94	64.14	12.23	60.66	11.42

Figura C.2: Promedios en la exactitud de clasificación de los 27 sujetos usando distintos clasificadores con 14 canales.

BIBLIOGRAFÍA

- [1] Instituto Nacional de Estadística Geografía e Informática, “Censo general de población y vivienda,” 2010.
- [2] S. Sanei and J. A. Chambers, *EEG signal processing*. John Wiley & Sons, 2008.
- [3] J. S. Brumberg, A. Nieto-Castanon, P. R. Kennedy, and F. H. Guenther, “Brain–computer interfaces for speech communication,” *Speech communication*, vol. 52, no. 4, pp. 367–379, 2010.
- [4] A. A. Torres-García, “Clasificación de palabras no pronunciadas presentes en Electroencefalogramas (EEG),” Tesis de Maestría, 2011.
- [5] H. Cecotti, “Spelling with brain-computer interfaces-current trends and prospects,” *Proc. of 5th French Conference on Computational Neuroscience (Neurocomp 2010)*, pp. 215–220, 2010.
- [6] B. Denby, T. Schultz, K. Honda, T. Hueber, J. Gilbert, and J. Brumberg, “Silent speech interfaces,” *Speech Communication*, vol. 52, no. 4, pp. 270–287, 2010.
- [7] A. A. Torres-García, C. A. Reyes-García, and L. Villaseñor Pineda, “Hacia la clasificación de habla no pronunciada mediante electroencefalogramas (EEG),” *XXXIV Congreso Nacional de Ingeniería Biomédica*, pp. 9–12, 2011.
- [8] G. Baier, T. Hermann, S. Sahle, and U. Stephani, “Sonified epileptic rhythms,” *Proceedings of 12th International Conference on Auditory Display*, pp. 148–151, 2006.
- [9] G. Baier and T. Hermann, *Sonification: listen to brain activity*. Springer, 2009.
- [10] G. Baier, T. Hermann, and U. Stephani, “Event-based sonification of EEG rhythms in real time,” *Clinical Neurophysiology*, vol. 118, no. 6, pp. 1377–1386, 2007.
- [11] T. Hermann, G. Baier, U. Stephani, and H. Ritter, “Vocal sonification of pathologic EEG features,” *Proceedings of the ICAD 2006*, vol. 6, pp. 158–163, 2006.

- [12] M. Elgendi, B. Rebsamen, A. Cichocki, F. Vialatte, and J. Dauwels, "Real-time wireless sonification of brain signals," *Advances in Cognitive Neurodynamics (III)*, pp. 175–181, 2013.
- [13] M. Elgendi, J. Dauwels, B. Rebsamen, R. Shukla, Y. Putra, J. Gamez, N. ZePing, B. Ho, N. Prasad, D. Aggarwal *et al.*, "From auditory and visual to immersive neurofeedback: Application to diagnosis of alzheimers disease," *Neural Computation, Neural Devices, and Neural Prosthesis*, pp. 63–97, 2014.
- [14] A. Porbadnigk and T. Schultz, "EEG-based Speech Recognition: Impact of Experimental Design on Performance," Master's thesis, Institut fur Theoretische Informatik Universitat Karlsruhe (TH), 2008.
- [15] A. Porbadnigk, M. Wester, J. P. Callies, and T. Schultz, "EEG-based Speech Recognition - Impact of Temporal Effects," *International Conference on Bio-inspired Systems and Signal Processing*, pp. 376–381, 2009.
- [16] M. Wester and T. Schultz, "Unspoken Speech - Speech Recognition Based On Electroencephalography," Master's thesis, Institut fur Theoretische Informatik Universitat Karlsruhe (TH), Karlsruhe, Germany, 2006.
- [17] E. Kandel, J. Schwartz, T. Jessell, S. Siegelbaum, and A. Hudspeth, *Principles of Neural Science, Fifth Edition*. McGraw-Hill Education, 2013.
- [18] M. F. Bear, B. W. Connors, and M. A. Paradiso, *Neuroscience vol.3*. Lippincott Williams & Wilkins, 2007.
- [19] N. Geschwind, "Language and the brain." *Scientific American*, 1972.
- [20] S. Baillet, J. C. Mosher, and R. M. Leahy, "Electromagnetic brain mapping," *Signal Processing Magazine, IEEE*, vol. 18, no. 6, pp. 14–30, 2001.
- [21] P. L. Nunez and R. Srinivasan, *Electric fields of the brain: the neurophysics of EEG*. Oxford university press, 2006.
- [22] G. H. Klem, H. Lüders, H. Jasper, and C. Elger, "The ten-twenty electrode system of the international federation. the international federation of clinical neurophysiology." *Electroencephalography and clinical neurophysiology. Supplement*, vol. 52, p. 3, 1999.
- [23] Brain Dynamics Lab at the University of Calgary, Department of Psychology , "What is EEG?" *braindynamics.ca/eeg/ (consultado en Ago-2014)*, 2009.

-
- [24] Instituto Neurológico de Barcelona, "EEG/CC," *neurologicbarcelona.com/eegcc* (consultado en Ago-2014), 2012.
- [25] XWave, "Mind to mind," *siriperera.com/?p=3846* (consultado en Ago-2014), 2010.
- [26] The Cloud Lab at Columbia University Graduate School of Architecture, Planning and Preservation, "Spatial problem solving," *thecloudlab.org/projects* (consultado en Ago-2014), 2011.
- [27] N. V. Thakor, *Quantitative EEG analysis methods and clinical applications*. Artech House, 2009.
- [28] N.-H. Liu, C.-Y. Chiang, and H.-C. Chu, "Recognizing the Degree of Human Attention Using EEG Signals from Mobile Sensors," *Sensors*, vol. 13, no. 8, pp. 10 273–10 286, 2013.
- [29] A. Vallabhaneni, T. Wang, and B. He, "Brain-computer interface," *Neural Engineering*, pp. 85–121, 2005.
- [30] A. Perry and S. Bentin, "Mirror activity in the human brain while observing hand movements: A comparison between eeg desynchronization in the μ -range and previous fmri results," *Brain research*, vol. 1282, pp. 126–132, 2009.
- [31] M. Palau-Baduell, A. Valls-Santasusana, and B. Salvadó-Salvadó, "Trastornos del espectro autista y ritmo mu. una nueva perspectiva neurofisiológica," *Rev Neurol*, vol. 52, no. 1, pp. 1–4, 2011.
- [32] G. Arias and H. Felipe, "Detección y clasificación de artefactos en señales EEG," *Memorias de STSIVA*, vol. 9, 2009.
- [33] J. R. Wolpaw, N. Birbaumer, W. J. Heetderks, D. J. McFarland, P. H. Peckham, G. Schalk, E. Donchin, L. A. Quatrano, C. J. Robinson, T. M. Vaughan *et al.*, "Brain-computer interface technology: a review of the first international meeting," *IEEE transactions on rehabilitation engineering*, vol. 8, no. 2, pp. 164–173, 2000.
- [34] J. Wolpaw, N. Birbaumer, D. McFarland, G. Pfurtscheller, and T. Vaughan, "Brain-computer interfaces for communication and control," *Clinical neurophysiology*, vol. 113, no. 6, pp. 767–791, 2002.
- [35] Y. Wang, R. Wang, X. Gao, B. Hong, and S. Gao, "A practical VEP-based brain-computer interface," *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, vol. 14, no. 2, pp. 234–240, 2006.
-

- [36] J. Gutiérrez-Martínez, J. Cantillo-Negrete, R. I. Cariño-Escobar, and D. Elías-Viñas, "Los sistemas de interfaz cerebro-computadora: una herramienta para apoyar la rehabilitación de pacientes con discapacidad motora," *Investigación en Discapacidad*, vol. 2, no. 2, 2013.
- [37] C. G. P. Jr, E. L. Naves, P. Pino, E. Losson, A. O. Andrade, and G. Bourhis, "Alternative communication systems for people with severe motor disabilities: a survey," *Biomedical Engineering OnLine*, vol. 10, no. 30, 2011.
- [38] C. Hintermüller, C. Kapeller, G. Edlinger, and C. Guger, *BCI Integration: Application Interfaces, Brain-Computer Interface Systems - Recent Progress and Future Prospects*. Dr. Reza Fazel-Rezai (Ed.), 2013.
- [39] G. Kramer, B. Walker, T. Bonebright, P. Cook, J. Flowers, N. Miner, J. Neuhoff, R. Bargar, S. Barrass, J. Berger *et al.*, "The sonification report: Status of the field and research agenda," *International Community for Auditory Display (ICAD)*, Santa Fe, NM, 1999.
- [40] T. Hermann and A. Hunt, "The discipline of interactive sonification," *Proceedings of the Int. Workshop on Interactive Sonification*, pp. 1–9, 2004.
- [41] F. El-Azm, "Sonification and augmented data sets in binary classification," Phd. Thesis, Technical University of Denmark, Lyngby, Denmark, 2005.
- [42] T. Hermann, A. Hunt, and J. G. Neuhoff, *The sonification handbook*. Logos Verlag, 2011.
- [43] T. Jing, *Sensing the Environment: Sound Visualization in Garment Design*. Parsons school, 2013.
- [44] T. Hermann, "Sonification for exploratory data analysis," Phd. Thesis, Bielefeld University, Germany, 2002.
- [45] T. Hermann and H. Ritter, "Listen to your data: Model-based sonification for data analysis," *Advances in intelligent computing and multimedia systems*, vol. 8, pp. 189–194, 1999.
- [46] T. Hinterberger, "The sensorium: a multimodal neurofeedback environment," *Advances in Human-Computer Interaction*, vol. 2011, p. 3, 2011.
- [47] A. A. Trevisan and L. Jones, "Brain music system: Brain music therapy based on real-time sonified brain signals," *Proc. IET Seminar on Assisted Living*, 2011.

-
- [48] A. Lucier, "Statement on: music for solo performer," *Biofeedback and the Arts, Results of Early Experiments*. Vancouver: Aesthetic Research Center of Canada Publications, pp. 60–61, 1976.
- [49] D. Rosenboom, *On being invisible*. Music Gallery Editions, 1977.
- [50] G. Baier and T. Hermann, "The sonification of rhythms in human electroencephalogram." *International Conference on Auditory Display (ICAD)*, 2004.
- [51] T. Hermann, P. Meinicke, H. Bekel, H. Ritter, H. M. Müller, and S. Weiss, "Sonification for EEG data analysis," *Proceedings of the 2002 International Conference on Auditory Display*, 2002.
- [52] G. Baier, T. Hermann, and U. Stephani, "Multi-channel sonification of human EEG," *Proceedings of the 13th International Conference on Auditory Display*, 2007.
- [53] T. Hermann, G. Baier, U. Stephani, and H. Ritter, "Kernel regression mapping for vocal EEG sonification," *Proceedings of the International Conference on Auditory Display*, 2008.
- [54] T. Hermann and G. Baier, "Interpreting EEG by voice: Vocal EEG Sonification," *Proceedings of the 1st Sketching Sonic Interaction Design Workshop*, 2009.
- [55] T. M. Rutkowski, F. Vialatte, A. Cichocki, D. P. Mandic, and A. K. Barros, "Auditory feedback for brain computer interface management: An EEG data sonification approach," *Knowledge-Based Intelligent Information and Engineering Systems*, pp. 1232–1239, 2006.
- [56] F. B. Vialatte and A. Cichocki, "Sparse bump sonification: a new tool for multi-channel EEG diagnosis of mental disorders; application to the detection of the early stage of Alzheimer's disease," *Neural Information Processing*, pp. 92–101, 2006.
- [57] T. Kaniwa, H. Terasawa, M. Matsubara, S. Makino, and T. M. Rutkowski, "electroencephalogram Steady State Response Sonification Focused on the Spatial and Temporal Properties," *The 20th International Conference on Auditory Display (ICAD)*, 2014.
- [58] J. Eaton and E. Miranda, "Real-time notation using brainwave control," *Sound and Music Computing Conference*, 2013.
-

- [59] T. Hinterberger, "Orchestral sonification of brain signals and its application to brain-computer interfaces and performing arts," *Proceedings of the 2nd International Workshop on Interactive Sonification*, 2007.
- [60] S. Mealla, M. Bosi, S. Jorda, and A. Väljamäe, "Sonification of brain and body signals in collaborative tasks using a tabletop musical interface," *Proc. International Conference on Auditory Display (ICAD)*, 2011.
- [61] M. R. Risk and A. J. Ramirez, "Análisis tiempo-frecuencia de la variabilidad de la frecuencia cardíaca y la presión arterial," *Procesamiento de Señales e Imágenes: Teoría y Aplicaciones*, pp. 205–255, 2009.
- [62] D. Ba, B. Babadi, P. L. Purdon, and E. N. Brown, "Robust spectrotemporal decomposition by iteratively reweighted least squares," *Proceedings of the National Academy of Sciences*, vol. 111, no. 50, pp. E5336–E5345, 2014.
- [63] R. Q. Quiroga and H. Garcia, "Single-trial event-related potentials with wavelet denoising," *Clinical Neurophysiology*, vol. 114, no. 2, pp. 376–390, 2003.
- [64] I. Homri, S. Yacoub, and N. Ellouze, "Optimal segments selection for EEG classification," *6th International Conference on Sciences of Electronics, Technologies of Information and Telecommunications (SETIT)*, pp. 817–821, 2012.
- [65] V. Krishnaveni, S. Jayaraman, S. Aravind, V. Hariharasudhan, and K. Ramadoss, "Automatic identification and removal of ocular artifacts from EEG using wavelet transform," *Measurement science review*, vol. 6, no. 4, pp. 45–57, 2006.
- [66] A. Subasi, "EEG signal classification using wavelet feature extraction and a mixture of expert model," *Expert Systems with Applications*, vol. 32, no. 4, pp. 1084–1093, 2007.
- [67] A. A. Torres-García, C. A. Reyes-García, and L. Villaseñor Pineda, "Análisis de Señales Electroencefalográficas para la Clasificación de Habla Imaginada," *Revista Mexicana de Ingeniería Biomédica*, vol. 34, no. 1, pp. 23–39, 2013.
- [68] E. Standard, "Speech processing, transmission and quality aspects (stq); distributed speech recognition; advanced front-end feature extraction algorithm; compression algorithms," *ETSI ES*, vol. 202, no. 050, p. v1, 2007.
- [69] J. Ye, "Speech recognition using time domain features from phase space reconstructions," Phd. Thesis, Marquette University Milwaukee, Wisconsin, 2004.

-
- [70] D. Michie, D. Spiegelhalter, and C. Taylor, *Machine Learning, Neural and Statistical Classification*. Overseas Press, 2009.
- [71] L. Rokach, *Pattern Classification Using Ensemble Methods*. World Scientific, 2009.
- [72] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.
- [73] B. G. Tabachnick, L. S. Fidell *et al.*, *Using multivariate statistics*. Allyn and Bacon Boston, 2001.
- [74] E. M. Dewan, "Occipital alpha rhythm eye position and lens accommodation," *Nature*, vol. 214, pp. 975–977, 1967.
- [75] C. S. DaSalla, H. Kambara, Y. Koike, and M. Sato, "Spatial filtering and single-trial classification of EEG during vowel speech imagery," pp. 1–4, 2009.
- [76] DZmura, M. and Deng, S. and Lappas, T. and Thorpe, S. and Srinivasan, R., "Toward EEG sensing of imagined speech," *Human-Computer Interaction. New Trends*, pp. 40–48, 2009.
- [77] P. Suppes, Z. Lu, and B. Han, "Brain wave recognition of words," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 94, no. 26, p. 14965, 1997.
- [78] P. Suppes, B. Han, and Z.-L. Lu, "Brain-wave recognition of sentences," *Proceedings of the National Academy of Sciences*, vol. 95, no. 26, pp. 15 861–15 866, 1998.
- [79] S. Deng, R. Srinivasan, T. Lappas, and M. Dzmura, "EEG classification of imagined syllable rhythm using Hilbert spectrum methods," *Journal of Neural Engineering*, vol. 7, p. 046006, 2010.
- [80] K. Brigham and B. Kumar, "Imagined speech classification with EEG signals for silent communication a preliminary investigation into synthetic telepathy," *4th International Conference on Bioinformatics and Biomedical Engineering (iCBBE)*, pp. 1–4, 2010.
- [81] Xuemin, Chi and Hagedorn, John and Schoonover, Daniel and DZmura, Michael, "EEG-based discrimination of imagined speech Phonemes," *International Journal of Bioelectromagnetism*, vol. 13, no. 4, pp. 201–206, 2011.
-

- [82] M. Wand, "Wavelet-based preprocessing of electroencephalographic and electromyographic signals for speech recognition," *Studienarbeit Lehrstuhl Prof. Waibel Interactive Systems Laboratories Carnegie Mellon University, USA and Institut für Theoretische Informatik Universität Karlsruhe (TH), Germany*, 2007.
- [83] A. A. Torres-García, C. A. Reyes-García, and L. Villaseñor Pineda, "Toward a silent speech interface based on unspoken speech," *BIOSTEC - BIOSIGNALS*, pp. 370–373, 2012.
- [84] L. Wang, X. Zhang, X. Zhong, and Y. Zhang, "Analysis and classification of speech imagery EEG for BCI," *Biomedical Signal Processing and Control*, vol. 8, no. 6, pp. 901–908, 2013.
- [85] T. M. Rutkowski, A. Cichocki, and D. Mandic, "Information fusion for perceptual feedback: A brain activity sonification approach," *Signal Processing Techniques for Knowledge Extraction and Information Fusion*, pp. 261–273, 2008.
- [86] G. Marco-Zaccaria, "Sonification of EEG signals: A study on alpha band instantaneous coherence," Master thesis, 2011.
- [87] E. Miranda, A. Brouse, B. Boskamp, and H. Mullaney, "Plymouth brain-computer music interface project: Intelligent assistive technology for music-making," *International Computer Music Conference*, 2005.
- [88] C. Anderson, "EEG to tones," *Department of Computer Science, Colorado State University (cs.colostate.edu/eeg)*, 2005.
- [89] J. Calliess, "Further investigations on unspoken speech," Master's thesis, Institut für Theoretische Informatik Universität Karlsruhe (TH), Karlsruhe, Germany, 2006.
- [90] Emotiv, "Emotiv Epoc and testbench specifications," *emotiv.com/epoc.php (consultado en Oct-2014)*, 2014.
- [91] X. Pei, D. L. Barbour, E. C. Leuthardt, and G. Schalk, "Decoding vowels and consonants in spoken and imagined words using electrocorticographic signals in humans," *Journal of neural engineering*, vol. 8, no. 4, p. 046028, 2011.
- [92] J. Mourino, J. del R Millan, F. Cincotti, S. Chiappa, R. Jané, and F. Babiloni, "Spatial filtering in the training process of a brain computer interface," *Engineering in Medicine and Biology Society. Proceedings of the 23rd Annual International Conference of the IEEE*, vol. 1, pp. 639–642, 2001.

-
- [93] K. Kahle, "Optimal Spatial Filtering for BCIs Using an Emotiv Epoc EEG Headset," *Washington University Undergraduate Research Digest*, vol. 7, no. 2, pp. 104–112, 2012.
- [94] M. J. Alhaddad and K. Abdulaziz, "Common average reference (CAR) improves p300 speller," *International Journal of Engineering and Technology*, vol. 2, no. 3, p. 21, 2012.
- [95] P. Vaidyanathan, "Generalizations of the sampling theorem: Seven decades after nyquist," *Circuits and Systems I: Fundamental Theory and Applications, IEEE Transactions on*, vol. 48, no. 9, pp. 1094–1109, 2001.
- [96] Laboratorio de procesamiento de bioseñales y computación médica, "Ejemplos de sonificación de EEG," goo.gl/SVtqiQ (consultado en Dic-2014), 2014.
- [97] Lotte, F and Congedo, M and Lécuyer, A and Lamarche, F and Arnald, B, "A review of classification algorithms for EEG-based brain-computer interfaces," *Journal of Neural Engineering*, vol. 4, pp. r1–r13, 2007.
- [98] X. Hu, H. Zhang, L. Zhan, Y. Xue, W. Zhou, and G. Zhao, "Isolated word speech recognition system based on FPGA," *Journal of Computers*, vol. 8, no. 12, pp. 3216–3222, 2013.
- [99] R. Kohavi *et al.*, "A study of cross-validation and bootstrap for accuracy estimation and model selection," *IJCAI*, vol. 14, no. 2, pp. 1137–1145, 1995.
- [100] I. Omerhodzic, S. Avdakovic, A. Nuhanovic, and K. Dizdarevic, "Energy distribution of EEG signals: EEG signal wavelet-neural network classifier," *World Academy of Science, Engineering and Technology*, vol. 61, pp. 1190–1195, 2010.
- [101] D. Gutiérrez and R. Salazar-Varas, "EEG signal classification using time-varying autoregressive models and common spatial patterns," *Engineering in Medicine and Biology Society, 2011 Annual International Conference of the IEEE*, pp. 6585–6588, 2011.
- [102] D. Gutiérrez and R. Salazar-Varas, "Using eigenstructure decompositions of time-varying autoregressions in CSP-based EEG signal classification," *Biomedical Signal Processing and Control*, vol. 7, no. 6, pp. 622–631, 2012.