



# **DIAGNÓSTICO DE LA ENFERMEDAD DE ALZHEIMER MEDIANTE ANÁLISIS DE SEÑALES ELECTROENCEFALOGRÁFICAS**

**TESIS**

Que para obtener el grado de  
**MAESTRO EN CIENCIAS EN INGENIERÍA ELÉCTRICA**  
Sistemas Computacionales

presenta

**Javier Alejandro Rodríguez Herrejón**

**Dr. José Antonio Camarena Ibarrola**

**Director de Tesis**

Universidad Michoacana de San Nicolás de Hidalgo

Julio 2018





## DIAGNÓSTICO DE LA ENFERMEDAD DE ALZHEIMER MEDIANTE EL ANÁLISIS DE SEÑALES ELECTROENCEFALOGRÁFICAS

Los Miembros del Jurado de Examen de Grado aprueban la **Tesis de Maestría en Ciencias en Ingeniería Eléctrica** de *Javier Alejandro Rodríguez Herrejón*.

Dr. Félix Calderón Solorio  
*Presidente del Jurado*

Dr. José Antonio Camarena Ibarrola  
*Director de Tesis*

Dr. Juan José Flores Romero  
*Vocal*

Dr. Jaime Cerda Jacobo  
*Vocal*

Dr. Alain Manzo Martínez  
*Revisor Externo (Universidad Autónoma de Chihuahua)*

Dr. Félix Calderón Solorio  
*Jefe de la División de Estudios de Posgrado de la Facultad de Ingeniería Eléctrica. UMSNH (Por reconocimiento de firmas).*




***Deseo expresar mis agradecimientos a:***

*Mi madre María Salud Herrejón Ruiz, y a mi padre Guillermo Antonio Rodríguez González (†), por el apoyo brindado incondicionalmente en todos los aspectos de mi vida.*

*De igual manera a mis hermanos, Guillermo Fernando y Yeslie Salud. Su apoyo y ejemplo como estudiante y persona durante todo el trayecto de mis estudios.*

*Sin las enseñanzas y el trabajo empeñado durante el proceso de la maestría este trabajo no podría haber sido elaborado, agradezco a mis profesores, en especial al Dr. José Antonio Camarena Ibarrola.*

*Al Consejo Nacional de Ciencia y Tecnología (CONACyT) por los apoyos financieros para cursar mis estudios de posgrado.*



# Resumen

La enfermedad de Alzheimer es una enfermedad neurodegenerativa que provoca pérdida de memoria gradual cada vez más notoria, se ha convertido en una de las enfermedades crónicas que más afectan a la población, sobre todo en el caso de los adultos mayores. Las señales electroencefalográficas muestran la actividad del cerebro, la cual es analizada para conocer condiciones específicas de su funcionalidad o trastornos mentales.

El tema central de este trabajo es sobre el diagnóstico de la enfermedad de Alzheimer mediante el uso de las señales electroencefalográficas, el análisis de las señales permite hacer un estudio económico y más cómodo para el paciente, además de que permite realizarlo en una etapa temprana y de ese modo, planificar mejor el tratamiento y mejorar la calidad de vida de los pacientes.

Parte de la metodología consistió en obtener la energía, entropía espectral y coeficiente de correlación de Pearson para encontrar las mejores características de la señal y que ayudan a realizar la detección de la enfermedad de Alzheimer. En la búsqueda de mejorar estas características, se aplicó la técnica del análisis de componentes independientes a las señales antes del cálculo de la energía y entropía.

Se encontró que la energía y la entropía fueron las características con mejores resultados dentro de la presente investigación, al diferenciar de mejor manera a las personas sin la enfermedad de Alzheimer, y a los pacientes que sí la tienen. También se observó que el análisis de componentes independientes no contribuyó a mejorar las características para determinar el diagnóstico, esto debido a la naturaleza de las señales utilizadas.

Palabras claves: ICA, frecuencia, energía, entropía, correlación.





# Abstract

Alzheimer's disease is a neurodegenerative disease that causes gradual memory loss every time it is notorious, it has become one of the chronic diseases that affect the population, especially in the case of older adults. The electroencephalographic signals show the activity of the brain, which is analyzed to know the specific specifications of its functionality or mental disorders.

The central theme of this work is about the diagnosis of Alzheimer's disease through the use of electroencephalographic signals, the analysis of the causes of the statistics allows an economic study and more comfortable for the patient, in addition to allowing it to be carried out in a stage early and thus, better plan the treatment and improve the quality of life of patients.

Part of the methodology consisted in obtaining the energy, spectral entropy and Pearson's correlation coefficient to find the best characteristics of the signal and help to make the detection of Alzheimer's disease. In the search to improve these characteristics, the technique of analysis of independent components was applied to the signals before the calculation of energy and entropy.

We found that energy and entropy were the characteristics with the best results within the present investigation, by differentiating better the people without Alzheimer's disease, and the patients who have the same. It was also observed that the analysis of independent components did not contribute to improve the characteristics for the diagnosis, this due to the nature of the signals used.



# Contenido

Dedicatoria . . . . .	III
Resumen . . . . .	V
Abstract . . . . .	VII
Contenido . . . . .	IX
Lista de Figuras . . . . .	XI
Lista de Tablas . . . . .	XIII
Lista de Símbolos . . . . .	XV
1. Introducción . . . . .	1
1.1. Planteamiento del problema . . . . .	1
1.2. Antecedentes . . . . .	2
1.2.1. Estado del arte . . . . .	2
1.2.2. Separación ciega de fuentes . . . . .	6
1.2.3. Deterioro cognitivo leve . . . . .	6
1.2.4. Métodos para el diagnóstico de la enfermedad de Alzheimer . . . . .	7
1.3. Objetivos de la tesis . . . . .	9
1.3.1. Objetivo general . . . . .	9
1.3.2. Objetivos particulares . . . . .	10
1.4. Descripción de capítulos . . . . .	10
2. Detección y generalidades de la enfermedad de Alzheimer . . . . .	11
2.1. La enfermedad de Alzheimer . . . . .	12
2.2. Las señales electroencefalográficas . . . . .	14
2.3. Conclusiones sobre la detección y generalidades de la EA . . . . .	19
3. Extracción de características . . . . .	21
3.1. Entropía espectral . . . . .	21
3.2. Coeficiente de correlación de Pearson . . . . .	22
3.3. Separación ciega de fuentes y selección de características . . . . .	24
3.3.1. Matrices de transformación como filtros . . . . .	28
3.3.2. Análisis de componentes principales (PCA) . . . . .	32
3.3.3. Ejemplo de descomposición de valores propios . . . . .	34
3.3.4. Análisis de componentes independientes (ICA) . . . . .	37
3.3.5. Algoritmo rápido de punto fijo para el ICA . . . . .	43

---

3.4. Conclusión sobre la extracción de características . . . . .	45
4. Sistema de extracción de características para la detección de la enfermedad de Alzheimer (SECDEA)	47
4.1. Técnica utilizada para extraer características . . . . .	47
4.1.1. Señales . . . . .	48
4.1.2. Transformada de Fourier y cálculo de energía . . . . .	49
4.1.3. Distribución de probabilidad y entropía espectral . . . . .	51
4.1.4. Coeficiente de correlación de Pearson . . . . .	52
4.2. Análisis de componentes independientes . . . . .	52
4.3. Conclusión sobre el SECDEA . . . . .	56
5. Resultados	59
5.1. Base de datos . . . . .	59
5.2. Energía . . . . .	61
5.3. Entropía espectral . . . . .	62
5.4. Coeficiente de correlación de Pearson . . . . .	63
5.5. Diagnóstico de la EA . . . . .	64
5.6. Conclusión sobre resultados . . . . .	66
6. Conclusiones y trabajos futuros	67
6.1. Conclusiones particulares . . . . .	68
6.2. Trabajos futuros . . . . .	68
Referencias	69

# Lista de Figuras

1.1.	Corte sagital que evidencia prominencia de surcos y cisuras, (a) en paciente con EA, (b) persona sana [Orellana02]. . . . .	4
2.1.	Tomografía computarizada de paciente con EA [Delgado14]. . . . .	13
2.2.	A) Medición de la atrofia del lóbulo temporal en TC. Orientación del lóbulo temporal, imagen axial de TC en la que se muestra la medición de la anchura del asta temporal del ventrículo lateral respecto al diámetro biparietal. B) Atrofia del lóbulo temporal en RM. Imagen coronal potenciada en FLAIR (Fluid Attenuated Inversion Recovery), que muestra el método de medida del ancho del asta temporal del ventrículo lateral (flecha doble) así como de la cisura coroidea (flecha gruesa) [Fernández10]. . . . .	14
2.3.	Diagrama de representación del sistema 10-20. El cerebro es mapeado por cuatro puntos: Nasal, Occipital (inión) y por preauriculares derecho e izquierdo (orejas). [Sánchez14] . . . . .	17
2.4.	Registros. Ilustración de un registro referencial (izquierda) con las referencias A1 y A2 para los electrodos del lado izquierdo y del lado derecho, respectivamente, y un registro bipolar (derecha) con la disposición de los electrodos vecinos en cadenas longitudinales [Boutros02]. . . . .	18
3.1.	Datos simulados siguiendo una relación lineal entre R y E con fluctuaciones gaussianas aleatorias agregadas a los valores R. (a) sin fluctuaciones, resultando en $r = 1$ . (b) $\sigma(R) = 100$ resultando en $r = 0.977$ . (c) $\sigma(R) = 1000$ , y $r = 0.585$ . Teniendo en cuenta el cambio de escala vertical [Hall15]. . . . .	24
3.2.	Separación de las observaciones para recuperar las fuentes [Puigt11]. . . . .	25
3.3.	Problema de la fiesta de cócteles[Clifford07]. . . . .	26
3.4.	Distribución de probabilidad normal con valor de curtosis en 3 (azul) y distribuciones de probabilidad con valores de curtosis mayor a 3 (rojo) y menor a 3 (verde). . . . .	28
3.5.	El paradigma general de la SCF para el filtrado, $\mathbf{S}_p$ es la matriz de valores singulares, y $\mathbf{V}^T$ la matriz de vectores columna propios para el PCA, se observa en $Z$ la representación de señales dentro del bloque, y en $\mathbf{X}_{filt}$ después del proceso de reconstrucción una versión filtrada de los datos [Clifford08]. . . . .	31

---

3.6. Distribuciones con tercer y cuarto momento, sesgo o asimetría (a) y curtosis (b) respectivamente, que son significativamente diferentes de la distribución normal o gaussiana [Clifford08]. . . . .	39
3.7. Algoritmo rápido de punto fijo para ICA con señales de entrada simuladas, a) Señales fuentes obtenidas, b) Señales de observación o mezcladas. . . . .	45
4.1. Diagrama de flujo de la implementación. . . . .	48
4.2. 4 señales de diferentes canales EEG tomadas de las 23 posibles para dar una representación gráfica. . . . .	49
4.3. Transformada de Fourier en el dominio de la frecuencia. . . . .	50
4.4. Bandas de frecuencia. . . . .	50
4.5. Histogramas de los valores por cada banda de frecuencia de todas las señales. . . . .	52
4.6. Señales antes (a), y después (b) del ICA, se puede observar su permutación, escalamiento y separación de componentes de una señal para agregarla a otra. . . . .	53
4.7. Transformadas de Fourier antes (a), y después (b) del ICA. . . . .	54
4.8. Bandas de frecuencia antes (a), y después (b) del ICA. . . . .	55
4.9. Histogramas de las bandas de frecuencia antes (a), y después (b) del ICA, la forma del histograma se conserva aun que los valores de amplitud se ven distintos en el eje por el escalamiento. . . . .	56
5.1. Gráfico de las características relevantes, cada sujeto corresponde a un punto, también los sujetos para la prueba 1 y 2. . . . .	65
5.2. Gráfico de las distancias entre medias y sujetos de prueba. . . . .	66

# Lista de Tablas

1.1. Métodos usados para el diagnóstico de la EA basados en señales de EEG [Elgandelwar16]. . . . .	8
2.1. Patrones de frecuencia EEG y estados de comportamiento. . . . .	16
5.1. Canales en la base de datos. . . . .	60
5.2. Energía en las bandas de frecuencia sin ICA. . . . .	61
5.3. Energía en las bandas de frecuencia con ICA. . . . .	61
5.4. Entropía en las bandas de frecuencia sin ICA. . . . .	62
5.5. Entropía en las bandas de frecuencia con ICA. . . . .	63
5.6. Coeficiente de correlación de Pearson. . . . .	64





# Lista de Símbolos

$H(f)$	Entropía espectral
$r$	Coefficiente de correlación de Pearson
$p_i$	Probabilidad para el punto $i$
$C_{xy}$	Covariancia entre $x$ y $y$
$C_{xx}$	Covariancia entre $x$ y $x$
$C_{yy}$	Covariancia entre $y$ y $y$
$\mathbf{X}$	Observaciones
$\mathbf{Z}$	Fuentes independientes
$\mathbf{A}$	Matriz de mezcla
$\mathbf{Y}$	Representación transformada de $\mathbf{X}$
$\mathbf{W}$	Inverso de $\mathbf{A}$
$\kappa$	Curtosis
$P(Z)$	Función de densidad conjunta
$\mu_X$	Media de una variable aleatoria
$\sigma_X^2$	Varianza de una variable aleatoria
$\zeta(\mathbf{X})$	Sesgo



# Glosario

EA	Enfermedad de Alzheimer
EEG	Electroencefalograma
REM	Rapid Eye movement
MEG	Magnetoencefalografía
SFC	Separacion ciega de fuentes
ICA	Análisis de componentes independientes
MCI	Deterioro cognitivo leve
PCA	Análisis de componentes principales
INEGI	Instituto nacional de estadística y geografía
CONAPO	Consejo nacional de población
RM	Resonancia magnética
TC	Tomografía computarizada
FLAIR	Fluid attenuated inversion recovery
DFT	Transformada discreta de Fourier
FFT	Transformada rápida de Fourier
SCA	Análisis de componentes dispersos
NMF	Factorización de matriz no negativa
SVD	Descomposición de valores singulares
ECG	Electrocardiograma



## Capítulo 1

# Introducción

En el presente capítulo se define el problema que se aborda en esta tesis, y se desarrolla el marco teórico. Se establecen metas y objetivos que delimitan el problema y se realiza una breve descripción de los capítulos que contiene la tesis.

### 1.1. Planteamiento del problema

La enfermedad de Alzheimer (EA) es una enfermedad neurodegenerativa que provoca pérdida de memoria gradual cada vez más notoria. Actualmente existen diversas instituciones especializadas en la detección y el tratamiento de la EA. En la lucha contra ésta, se han logrado importantes avances, entre ellos la incorporación de la tecnología, desarrollando nuevas formas de realizar estudios médicos para conocer el estado de salud del paciente. La EA se ha convertido en una de las enfermedades crónicas que más afectan a la población, sobre todo en el caso de los adultos mayores, haciendo entrar en un estado de alarma a las asociaciones de salud del mundo [AlzheimerMexOrg17]. Como todo padecimiento, si la EA es detectada en la etapa temprana de la enfermedad, el tratamiento es más eficaz, además de tener la oportunidad de planificar mejor la enfermedad, para que cause el menor impacto dentro de la vida del paciente. Los médicos especulan que la enfermedad suele iniciar a los 65 años, sin embargo, en la actualidad se han observado inicios de casos de la enfermedad en la adultez temprana, incluso en los casos más drásticos en la adolescencia. Conociendo lo anterior, es necesario diseñar métodos más rápidos y accesibles para que los pacientes

identifiquen si la EA se está desarrollando.

## 1.2. Antecedentes

En el transcurso de los años la medicina y la tecnología han ido sumando esfuerzos en los tratamientos y métodos para la detección de enfermedades [Al-Jumeily15]. Los aparatos tecnológicos son utilizados diariamente en el uso de la medicina moderna. Para el estudio del cerebro existen diferentes datos que representan su funcionamiento, una fuente de estos datos son las señales electroencefalográficas, las cuales muestran la actividad cerebral, y suelen obtenerse en distintas posiciones del cuero cabelludo [Elgandelwar16].

Se han desarrollado análisis de diversos algoritmos que tienen el propósito de encontrar dentro de las señales electroencefalográficas, características que permitan afirmar que un paciente está en una fase de la EA [Elgandelwar16].

Los algoritmos que se utilizan para conseguir estas características se especializan en el manejo de la información que se encuentra dentro de la señal; mediante las mediciones de energía en las bandas de frecuencia, reducir la complejidad de las señales y los parámetros de sincronización, se obtendrán los valores que permitirán posteriormente realizar una clasificación entre los pacientes sanos y los que están en etapa temprana de un deterioro cognitivo leve, el cual es un reflejo de la aparición de la EA [Elgandelwar16].

### 1.2.1. Estado del arte

El electroencefalograma (EEG) es una prueba no invasiva que permite estudiar la actividad eléctrica cerebral, esta actividad puede ser representada y grabada mediante señales las cuales se conocen como señales electroencefalográficas, llamadas también ondas del cerebro, emitidas por las células nerviosas de la corteza del cerebro. El EEG tiene diferentes “bandas”, definidas por la frecuencia de las ondas; banda delta (lentas) de menos de 4 Hz; banda teta de 4 a 8 Hz, la alfa de 8 a 12 Hz, la beta de aproximadamente 14 a 30 Hz y la gama de 30 a 80 Hz [Al-Jumeily15]. La banda alfa se ve mejor en el área parieto-occipital, y la banda beta son normalmente más prominentes en las regiones frontal y central. Estas bandas cuando se graban simultáneamente, se diferencian unas de otras y

reflejan diferentes procesos cognitivos. Los cambios de energía, son conocidos también como actividad. De la banda alfa se aprecia mejor o son más significativos cuando el sujeto está despierto y relajado con los ojos cerrados, y las ondas beta durante la fase de sueño REM (la fase REM comienza aproximadamente 90 minutos después de dormir). La actividad eléctrica cerebral se caracteriza también por la amplitud o la fuerza de las oscilaciones [McLaughlin10].

La desaceleración del EEG consiste en la disminución de la energía en las señales alfa, beta y gama, con un aumento de la energía en las señales delta y teta [Elgandelwar16]. Es un hallazgo frecuente en pacientes con EA y su aparición está relacionada con la etapa de la enfermedad. Se puede encontrar un aumento de energía en las bandas de frecuencias bajas en más del 90 % de los pacientes con EA moderada o grave. En las primeras etapas de la enfermedad, los cambios son más sutiles, ya que un número considerable de pacientes pueden tener un EEG normal. El análisis espectral del EEG ha contribuido en gran medida a una mejor definición de las modificaciones del EEG en las primeras etapas de la enfermedad: estos cambios consisten en el decremento de la energía de la banda alfa y en un aumento de la energía de la banda teta. La energía de la banda delta aumenta y la energía de la banda alfa disminuye solo en etapas más avanzadas; en pacientes levemente demenciales, los cambios de energía en la banda alfa pueden ser incluso mayores que en los grupos de control. Todos estos cambios de EEG no son específicos de la EA, y también se han reportado en demencia vascular [Locatelli98].

Un diagnóstico definitivo solo es posible mediante autopsia, exploración física del cerebro del cadáver, usado para conocer que el paciente que falleció tenía la enfermedad, se debe intentar un diagnóstico diferencial con otros tipos de demencia y depresión mayor mientras el paciente esté con vida [Orellana02]. Las imágenes de resonancia magnética, que se muestran en Figura 1.1, donde se muestra una resonancia magnética de un paciente con EA y la de una persona sana. La tomografía computarizada pueden ser normales en las primeras etapas de la EA, pero la principal característica de las exploraciones cerebrales es una atrofia cortical difusa, la cual se refiere a la degeneración gradual y progresiva de la capa externa del cerebro (la corteza) en la parte del cerebro ubicada en la parte posterior de la cabeza [EAAssociation17]. Las pruebas de estado mental también son útiles como apoyo

a las tecnologías para el diagnóstico.

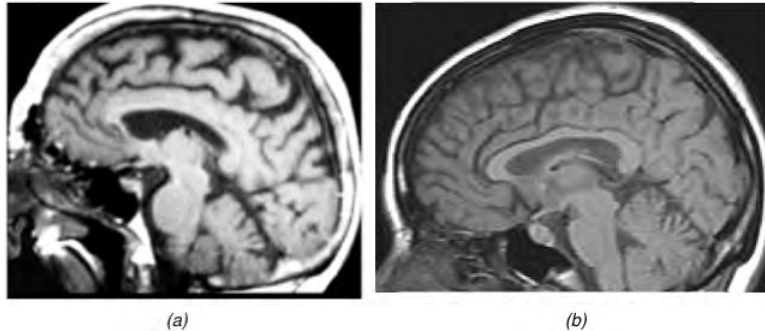


Figura 1.1: Corte sagital que evidencia prominencia de surcos y cisuras, (a) en paciente con EA, (b) persona sana [Orellana02].

El EEG se ha utilizado como una herramienta para diagnosticar demencias durante varias décadas. Las anomalías típicas de EEG en la EA se caracterizan por una desaceleración de la actividad de fondo (La desaceleración de la actividad de fondo dominante posterior está a 7 Hz, más lento que el rango de frecuencia alfa [Benbadis17]), mediante un aumento de la energía en las bandas delta y teta junto con una disminución de la energía en la banda alfa y una menor coherencia entre las áreas corticales, aunque en las primeras etapas de la enfermedad, el EEG puede exhibir frecuencias normales [Abásolo06].

El progreso reciente en la teoría de la dinámica no lineal ha proporcionado nuevos métodos para el estudio del EEG. La no linealidad en el cerebro se introduce incluso a nivel celular, por el comportamiento dinámico de las neuronas individuales. Además, la hipótesis de un cerebro completamente estocástico, entendiéndose por esto que sus acciones son completamente aleatorias, puede rechazarse debido a su capacidad para realizar tareas cognitivas sofisticadas. Considerando esto, las técnicas de análisis dinámicas no lineales pueden ser un mejor enfoque que los métodos lineales tradicionales para obtener una mejor comprensión de las dinámicas anormales en las señales de EEG. La complejidad de Lempel-Ziv es una medida de complejidad no paramétrica en una señal unidimensional relacionada con el número de subcadenas distintas y la tasa de su recurrencia. Se ha examinado la actividad de fondo del EEG en pacientes con EA con dos métodos no lineales: la complejidad de Lempel-Ziv y el cálculo de la medida de tendencia central, en este cálculo se utiliza el modelado caótico continuo que resume el grado de variabilidad en una señal [Abásolo06].



Se ha examinado también un método que radica en calcular la entropía del EEG [Hornero06]. La entropía es un concepto que aborda la aleatoriedad y la previsibilidad. Una mayor entropía a menudo es asociada con más aleatoriedad y menos orden del sistema [Hornero06]. Recientemente, se han introducido varios estimadores diferentes para determinar la entropía de las series de tiempo. Estos enfoques pueden clasificarse libremente en dos grupos: entropías espectrales y entropías de embebido. Las entropías espectrales estiman los cambios en la componente de amplitud del espectro de energía del EEG, el espectro de potencia se normaliza y se utiliza como función de distribución de probabilidad. Por otro lado, las entropías de embebido proporcionan información sobre cómo la señal de EEG fluctúa con el tiempo comparando las series de tiempo con una versión retrasada de sí mismas [Hornero06].

Debido a su relevancia y difícil diagnóstico, la utilidad del EEG en la detección de EA ha sido ampliamente investigada en las últimas décadas. Tanto el EEG como la magnetoencefalografía (MEG) registran las oscilaciones electromagnéticas producidas por las neuronas piramidales de forma no invasiva. En la EA, estas señales muestran anomalías que reflejan deficiencias anatómicas y funcionales de la corteza cerebral dañadas por la demencia [Escudero08]. Por lo tanto, la investigación de señales de EEG y MEG puede proporcionar información útil sobre la EA. La inspección visual de las grabaciones cerebrales electromagnéticas de pacientes con la EA en reposo con los ojos cerrados muestra una reducción en la actividad beta y alfa y una disminución del ritmo cortical posterior, entre otras anomalías. Además, diversos trabajos de investigación confirmaron esta desaceleración aplicando el análisis espectral computarizado a los datos EEG y MEG [Escudero08]. Por ejemplo, varios estudios cuantificaron los cambios que la EA produce en la energía de diferentes bandas de frecuencia. Estos hallazgos concuerdan con el hecho de que las grabaciones de EEG y MEG de los pacientes con EA tienen una frecuencia media más baja que las de individuos de control (i.e. sanos). La EA también causa otras anomalías en la actividad cerebral electromagnética, como una disminución de la conectividad funcional entre las áreas corticales. Esta reducción puede medirse realizando un análisis de coherencia de las señales EEG. Además de los cambios mencionados, la actividad en los espectros de EEG y MEG de los pacientes de la EA también disminuye. Esta reducción puede evaluarse

mediante la entropía espectral, que se calcula aplicando la entropía de Shannon a la función de densidad espectral de potencia normalizada. Además, los datos de EEG y MEG de los pacientes con EA también se han estudiado con diversas técnicas no lineales, como los estimadores de complejidad o las medidas de conectividad [Escudero08].

### 1.2.2. Separación ciega de fuentes

Existe una técnica llamada Separación Ciega de Fuentes (SCF), la cual consiste en estimar los componentes subyacentes de las señales EEG y MEG sin información a priori sobre esos componentes. Un filtrado con SCF puede mejorar la clasificación de la materia en función de las características extraídas de los datos EEG y MEG [Escudero08].

El análisis de componentes independientes (ICA) es un método para realizar SCF que recupera señales subyacentes o independientes de mezclas lineales de esas señales. El ICA es especialmente útil para eliminar ruido en el EEG, explotando criterios estadísticos independientes para separar las fuentes EEG, lo que permite limpiar la señal EEG. Por lo tanto, se ha utilizado para el análisis de diversas series temporales fisiológicas, incluido el EEG. Sin embargo, no está claro si ICA altera las distribuciones de EEG o no [Vialatte09]. Además del ruido instrumental y el ruido ambiental, el movimiento y otros ruidos fisiológicos (oculares, electromiográficos, electrodérmicos, electrovasculares y respiratorios) pueden interferir con el EEG. Los artefactos en el EEG se pueden definir como cualquier diferencia potencial debido a una fuente extra cerebral. Particularmente, los ruidos por movimientos musculares son especialmente problemáticos, porque pueden aparecer en patrones de EEG que son muy difíciles de diferenciar de las señales de EEG. Se encontró que la limpieza ICA mejoró la separación entre los pacientes con la EA en etapa temprana y los sujetos de control o pacientes sanos. Además de una distribución topográfica más consistente de la energía del EEG después de ICA [Vialatte09].

### 1.2.3. Deterioro cognitivo leve

El deterioro cognitivo leve (MCI del inglés Mild Cognitive Impairment) se caracteriza por un deterioro en el estado de la memoria del cerebro que probablemente conduzca hacia la enfermedad de Alzheimer moderada o la EA. Esta etapa de síntomas iniciales de

la EA ha estado bajo una gran influencia de la investigación durante mucho tiempo. Las estadísticas informaron que del 6 al 25 % de MCI se transforma en EA anualmente y del 0.2 al 4 % de personas sanas a EA, revelando el hecho de que MCI es un estado de transición de EA moderada y EA. La pérdida de conectividad funcional entre la zonas cortical y el hipocampo ha sido durante mucho tiempo un foco importante de las investigaciones para examinar la causa de la disfunción cognitiva en la EA [Al-Jumeily15].

El análisis estadístico de la interdependencia entre las series temporales registradas de diferentes áreas del cerebro, para estudiar la interacción funcional, se denomina conectividad funcional. Los estudios sobre MCI han mostrado una disminución de la potencia alfa y un aumento de la potencia teta (4 a 8 Hz) en las partes corticocorticales y subcorticales del cerebro. La reducción de la probabilidad de sincronización ocurre tanto en los electrodos interhemisféricos (delta-beta2) y frontoparietal (delta-gama)[Al-Jumeily15].

#### 1.2.4. Métodos para el diagnóstico de la enfermedad de Alzheimer

El objetivo de aplicar el método de análisis de componentes principales (PCA) para eliminar la información redundante de los conjuntos de datos es para obtener resultados más confiables. Los resultados en [Al-Jumeily15] revelaron que la medida de correlación cruzada muestra mayor diferencia en la sincronización de EA moderado y los sujetos de control en comparación con la sincronía de fase, mientras que la función de coherencia no funcionaba muy bien. También han indicado que las bandas alfa y teta juegan un papel importante en la identificación del cambio en la sincronización de EA moderado y sujetos control, especialmente en la región central temporal derecha (RT-C) y también en la región temporal-occipital (LT-O) [Al-Jumeily15].

La EA es caracterizada principalmente por la pérdida extensa de células neuronales, enredos neurofibrilares, y placas seniles en el hipocampo, la corteza entorrinal, en la neocórtex y otras regiones cerebrales.

Encontrar un enfoque computacional para la temprana detección en personas que serán pacientes de EA, pero que no muestran ningún signo clínico del padecimiento en el momento de las pruebas, es un desafío importante. Las señales de EEG son muy usadas para el diagnóstico de problemas neuronales. Estas señales están ayudando al diagnóstico

de enfermedades, especialmente cuando alguien considera que existe incluso después de la investigación clínica primaria. Se muestra que el análisis de datos de señales de EEG tienen características únicas para investigar el diagnóstico de las enfermedades neurológicas como el Alzheimer, la epilepsia y el Parkinson [Elgandelwar16].

Referencia	Métodos para diagnóstico	Eficacia
[Podgorelec12]	Enfoque de aprendizaje automático	86 %
[Henderson13]	Dimensión fractal fijada en cero del método EEG	78 %
[Henderson13]	Densidad de intervalos de cruce por cero	89 %
[Jutgla13]	Método sLORETA	84.4 %
[Bock05]	Análisis no lineal del EEG vía entropía de Tsallis	82 %
[Sneddon05]	Método cuantitativo para medir la variabilidad de EEG	88 %
[Dauwels11]	El poder relativo y las medidas de complejidad como características para clasificar la enfermedad	83 %
[Poil13]	Biomarcadores de EEG integradores para predecir el diagnóstico	88 %
[Dai12]	Enfoque M3 para clasificar la EA	89.4 %
[Falk12]	Análisis de modulación de amplitud	90.6 %

Tabla 1.1: Métodos usados para el diagnóstico de la EA basados en señales de EEG [Elgandelwar16].

Existen métodos, algunos de estos en Tabla 1.1, para detectar características en una señal EEG que muestran información de que un paciente está en etapa temprana de la EA, conocida también como MCI. Si el sujeto padece de la EA, algunos de los efectos principales se produjeron en la señal del EEG, como la desaceleración del EEG, la reducción de la complejidad de las señales EEG y la agitación en la sincronía EEG. Sin embargo, estas actividades mencionadas no siempre son fácilmente detectables. Una descripción breve de los principales efectos se observa a continuación:

I. Desaceleración del EEG. Se observó a partir del estudio de [Elgandelwar16] basado en la enfermedad de Alzheimer que causa que las señales de EEG se ralenticen. La EA se asocia con una disminución de la potencia en las señales alfa, beta y gama (señales de 8-30Hz de frecuencias más altas) con una potencia de aumento en las señales delta y teta.

Se han aplicado transformaciones de Fourier y mapas de tiempo-frecuencia para medir el cambio en la potencia espectral.

II. Reducción de la complejidad de las señales EEG. Diversos métodos se han usado para calcular la complejidad de las señales EEG, se han utilizado para cuantificar las señales de EEG en la MCI y los pacientes con EA. Algunos de estos métodos están basados en la teoría de la información [Sneddon05] como: la entropía de Tsallis [Bock05], la entropía aproximada [Bock05], la entropía de múltiples escalas, la entropía de la muestra y la información mutua y la complejidad de Lempel-Ziv [Costa02] se utilizan para cuantificar la complejidad del EEG [Elgandelwar16].

III. Agitación de parámetros de sincronía. Se ha empleado un gran número de parámetros de sincronía. En la biología del procesamiento de señales y las ciencias físicas, se miden estos parámetros de sincronismo. De la encuesta, se observa la disminución de la sincronía EEG en pacientes con EA. Varios parámetros de sincronía obtenidos de las señales EEG son afectados por las señales EEG. La sincronía de fases, la causalidad de Granger, la coherencia de la fase de magnitud y el coeficiente de correlación de Pearson son algunos de los parámetros de sincronía utilizados para el diagnóstico de la EA [Elgandelwar16].

La base de datos usada consiste en señales del “CHB-MIT Scalp EEG database” usadas en [Shoeb09], y señales proporcionadas por la doctora en neurociencia María Esther Olvera Corte, adscrita al Centro de Investigación Biomédica de Michoacán.

## **1.3. Objetivos de la tesis**

### **1.3.1. Objetivo general**

Investigar e implementar las técnicas para la extracción de la energía, la entropía y el coeficiente correlación de Pearson que son utilizados en las señales EEG para obtener información relevante que ayuden a establecer una conexión entre características dentro de la señal y las personas con la EA.

### 1.3.2. Objetivos particulares

- Determinar las características adecuadas que se pueden extraer de la señal EEG que permitan la identificación de la EA.
- Implementar un sistema dicotomizador que sea capaz de clasificar a un paciente como enfermo de Alzheimer o sano.
- Implementar separación ciega de fuentes para mostrar las diferencias que ofrece el método del ICA.

## 1.4. Descripción de capítulos

En el primer capítulo, se encuentran las secciones básicas del protocolo de investigación, el cual aborda el problema, la solución y sus alcances. El segundo capítulo menciona la EA y sus características médicas, también se aborda todo lo referente a las señales EEG. El tercer capítulo explica los métodos y tratamientos que se le dan a las señales EEG para la obtención de características, menciona toda la información que se debe conocer computacionalmente hablando que se maneja sobre lo aplicado en la presente tesis. También explica la técnica SCF, características, uso, métodos y matemática.

Dentro del cuarto capítulo se muestra el sistema implementado para la solución del problema, donde se pretenden encontrar las características que aporten mejor información para poder hacer la separación en dos grupos: las personas con la EA y las personas sanas. En el quinto capítulo, se muestran los resultados de los métodos desarrollados, experimentos realizados, así como interpretaciones que permitan decir que un paciente podría tener la EA. El capítulo seis, y último, menciona las conclusiones de la investigación y los resultados obtenidos en la misma, y finaliza con la mención de los trabajos futuros.

Los resultados mostraron 3 características en las cuales las diferencias entre las personas sanas y los pacientes con la EA son más notorias, siendo las siguientes: energía en la banda delta, energía en la banda gama y entropía en la banda delta. Se mostro también que en el caso específico de base de datos utilizada y con las técnicas implementadas, el algoritmo del ICA no aportaba mejoras considerables para la clasificación.

## Capítulo 2

# DetECCIÓN Y GENERALIDADES DE LA ENFERMEDAD DE ALZHEIMER

La demencia constituye la causa principal de incapacidad de largo plazo en la tercera edad, afecta al 2% de la población entre 65 y 70 años y al 20% de los mayores de 80 años. Se define como un deterioro crónico de las funciones superiores, adquirido (a diferencia del retraso mental) y en presencia de un nivel de consciencia y atención normales. La pérdida de una única función intelectual no es criterio suficiente para el diagnóstico de demencia. La demencia suele afectar a todas las funciones intelectuales, aunque en las fases iniciales se puede establecer el diagnóstico por el deterioro de tres de las siguientes áreas: lenguaje, memoria, destreza visuoespacial, afecto, personalidad o intelecto [Delgado14].

La mayor parte de las demencias se deben a procesos degenerativos diseminados y/o multifocales. Sin embargo, la masa cerebral no es un buen indicador del grado de funcionalidad intelectual y, por tanto, la existencia de una atrofia cerebral generalizada en las pruebas de imagen no siempre es indicativa de demencia. El diagnóstico de las demencias es eminentemente clínico: una historia clínica detallada es fundamental. Como, por desgracia, la mayoría de las demencias no tienen tratamiento, el esfuerzo inicial del médico debe encaminarse a intentar identificar aquel grupo de demencias en las que un tratamiento precoz puede revertir los síntomas o frenar la progresión de los mismos. En los últimos años, se han aplicado técnicas radiológicas al diagnóstico de las demencias; fundamentalmente se

han realizado estudios con resonancia magnética y tomografía computarizada de emisión monofónica. Es característica la atrofia temporal y las disfunciones temporoparietales en la fase inicial de la EA, o la atrofia y disfunción frontal en la demencia frontotemporal [Delgado14]. También el EEG puede ser de utilidad en entidades específicas como las enfermedades cerebrales de origen tóxico o metabólico, en la distinción entre delirio o depresión y demencia, así como en el diagnóstico de crisis parciales complejas. Además puede detectar patrones muy característicos de la enfermedad de Alzheimer [Sáez09].

## 2.1. La enfermedad de Alzheimer

La forma más común de demencia es la EA, en México se estima que 800,000 personas tengan un trastorno de demencia [Robledo14]. Según datos del censo de población 2010 del Instituto Nacional de Estadística y Geografía (INEGI), en nuestro país hay 10,055,379 personas, hombres y mujeres, en el grupo de edad entre 60 a 85 y más años [INEGI14]. Sumado a esto, un estudio del Consejo Nacional de Población (CONAPO) sobre la proyección de esperanza de vida de los mexicanos, muestra para el Distrito Federal un promedio de 77 y 78 años, para los años 2020 y 2030, respectivamente [CONAPO14].

La mayoría de los pacientes inicia los síntomas de la enfermedad a partir de los 65 años, aunque un debut temprano, antes de los 40 años también puede ocurrir, especialmente aquellos casos afectados de una forma hereditaria de la enfermedad. La EA se duplica cada 5 años a partir de los 60, de forma que afecta a un 1% de los pacientes a los 60 años, a un 2% a los 65 años, y a un 4% a los de 70 años [Delgado14].

Se caracteriza por una degeneración progresiva y selectiva de poblaciones neuronales en el córtex entorrinal, hipocampo, cortezas de asociación temporal, frontal y parietal, núcleos subcorticales y núcleos del tronco. No se afectan las cortezas primarias motoras y sensitivas, los ganglios basales ni el cerebelo (Figura 2.1).

A nivel macroscópico, la pérdida de neuronas se traduce en una atrofia generalizada, más grave en los lóbulos temporales, que se acompaña de dilatación secundaria del sistema ventricular [Delgado14].

La forma de diagnóstico de la EA tradicional es mediante exámenes clínicos y un



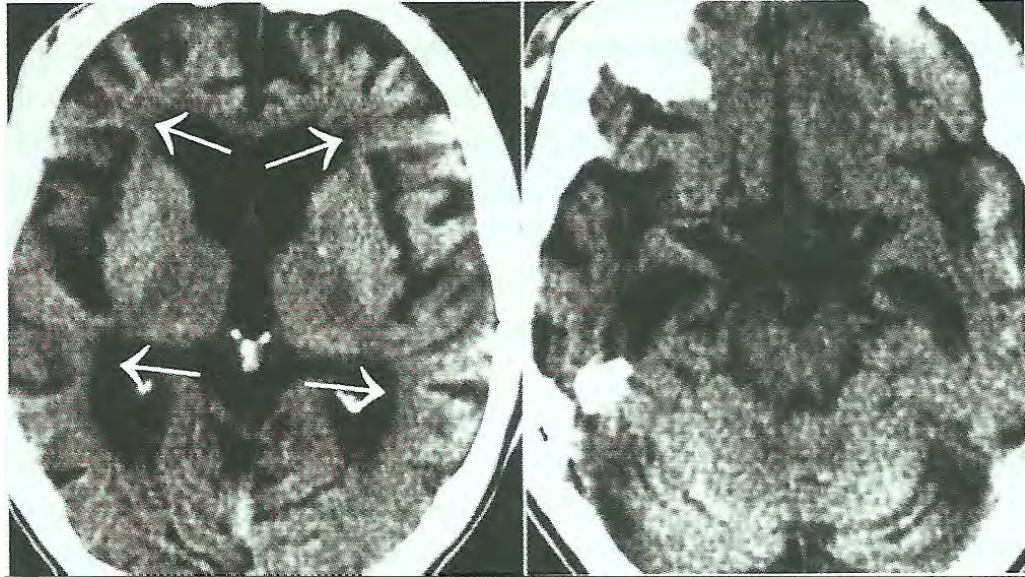


Figura 2.1: Tomografía computarizada de paciente con EA [Delgado14].

historial detallado del estado de salud del paciente, sin embargo, pruebas complementarias se realizan para apoyar el diagnóstico, entre ellas, con el uso del EEG, la resonancia magnética (RM) y la tomografía computarizada (TC).

La TC habitualmente es la primera prueba que se realiza, especialmente si hay contraindicaciones a la práctica de RM. Con TC adquirida en el plano orbitomeatal, la medida más adecuada para distinguir la enfermedad es la anchura del asta temporal del ventrículo lateral [Fernández10]. Con estos planos, se obtiene un cociente de la anchura radial del asta temporal respecto al diámetro biparietal (Figura 2.2.A). Así, se ajusta la pérdida de volumen al diámetro cefálico, siendo más útiles para seguimiento individual y diferencias intergrupos. Los sujetos normales presentan una media de 0,025, los pacientes con EA de 0,038 y aquellos con EA con extensas lesiones de sustancia blanca de 0,044 [Zhang08].

Con RM (Figura 2.2.B), el mejor plano es el coronal y se debe evaluar siempre en el mismo nivel para comprobar su evolución. Con una escala de sólo 5 grados se puede conseguir una adecuada precisión diagnóstica, entre el 83 y el 96 % [Fernández10]. Un estudio realizado con RM en población italiana por encima de 60 años, ha estimado que para EA leve el asta

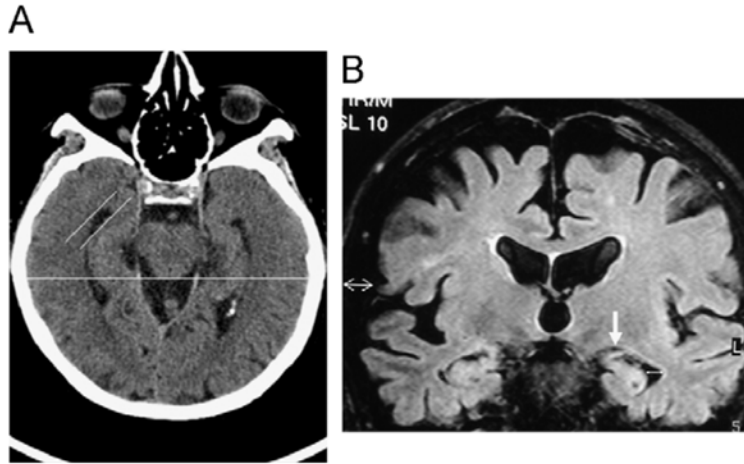


Figura 2.2: A) Medición de la atrofia del lóbulo temporal en TC. Orientación del lóbulo temporal, imagen axial de TC en la que se muestra la medición de la anchura del asta temporal del ventrículo lateral respecto al diámetro biparietal. B) Atrofia del lóbulo temporal en RM. Imagen coronal potenciada en FLAIR (Fluid Attenuated Inversion Recovery), que muestra el método de medida del ancho del asta temporal del ventrículo lateral (flecha doble) así como de la cisura coroidea (flecha gruesa) [Fernández10].

temporal mide 6.5 mm y para EA avanzada, 7.2 mm. Manteniendo una especificidad del 95 % para ambos, su sensibilidad es del 74 % para la EA leve y del 82 % para EA moderada [Frisoni96].

Las señales electroencefalográficas también son consideradas para evaluar las alteraciones cognitivas y como herramienta de diagnóstico. En los últimos años estas señales han servido para realizar el diagnóstico de la EA, debido al bajo costo y facilidad de realizar el estudio a diferencia de la TC y la RM.

## 2.2. Las señales electroencefalográficas

El cerebro viviente produce una salida continua de pequeñas señales eléctricas. El registro de estas señales, llamado EEG, es la suma de todos los potenciales postsinápticos (cambio temporal en el potencial de membrana de la célula) de las neuronas en la corteza cerebral. Las amplitudes de estas señales son tan pequeñas que se miden en microvoltios que son millonésimas de voltios o milésimas de milivoltios. Aunque son pequeñas, las señales se pueden detectar y registrar con precisión. Los electrodos que captan estas señales están

en contacto con el cuero cabelludo. Estas señales se amplifican miles de veces. Las señales amplificadas son grabadas con un electroencefalógrafo, que es un dispositivo para registrar las señales electroencefalográficas [iWorxSystems13].

El electroencefalograma es un registro continuo de ondas de frecuencia y amplitud variables. El número de ciclos de onda de los picos que se produce en un patrón de EEG en un período de tiempo determinado es su frecuencia. La amplitud del patrón de EEG es la energía eléctrica del patrón en términos de microvoltios. Hay cuatro bandas básicas de frecuencia en el EEG: gama (30 a 80 Hz), beta (14 a 30 Hz), alfa (8 a 12 Hz), teta (4 a 8 Hz) y delta (menor a 4 Hz).

Cada una de las cinco bandas de frecuencias en el EEG está asociada con un nivel diferente de activación de la corteza cerebral. La excitación cortical se refiere a los patrones de activación de las neuronas de la corteza cerebral. A medida que la frecuencia del patrón de EEG disminuye, el nivel de activación cortical disminuye, el patrón de EEG aumenta en amplitud. Por lo tanto, la frecuencia y la amplitud están inversamente relacionadas en el patrón del EEG [iWorxSystems13]. Un EEG con una gran amplitud y baja frecuencia indica un patrón de onda cerebral más sincronizado (grupos de células están actuando en conjunto), mientras que un EEG con una amplitud baja y una frecuencia alta generalmente se corresponde con una onda cerebral desincronizada (grupos de células están involucrados en actividades separadas). El nivel de activación cortical se correlaciona con varios estados psicológicos y conductuales [iWorxSystems13].

Algunos de los comportamientos o estados psicológicos a menudo asociados con las frecuencias de EEG prominentes se muestran en Tabla 2.1.

Para obtener el electroencefalograma se requiere de la colocación de electrodos, encargados de transformar las corrientes iónicas procedentes del tejido cerebral en corrientes eléctricas en la superficie craneal. Estos electrodos se encuentran conectados a un conjunto de canales independientes, permitiendo la captación y amplificación de la diferencia de potencial entre los electrodos receptores, y dirigiendo las señales recibidas a un dispositivo de registro [Boutros02].

Existen varios tipos de electrodos según la forma en la que se obtiene el registro:

- De Superficie, aplicados sobre el cuero cabelludo. Deben ser de colocación y remoción

Banda de frecuencia	Comportamiento/Estado psicológico
Alfa	Despierto, no concentrado, relajado, somnoliento; bajo nivel de estimulación ambiental (por ejemplo, ojos cerrados).
Beta	Despierto, alerta, atención enfocada y resolución de problemas; Sueño / sueño REM: alto nivel de estimulación ambiental (por ejemplo, ojos abiertos)
Delta	Imágenes mentales; sueño ligero.
Teta	Sueño profundo y reparador; estados de sueño vago.
Gama	Respuestas explosivas como las situaciones de ansiedad o terror extremas.

Tabla 2.1: Patrones de frecuencia EEG y estados de comportamiento.

fácil y rápida, indoloros y permanecer por un tiempo prolongado en la misma posición sin problemas [Argüelles09].

- Los Basales o Especiales, son colocados sobre el cráneo sin necesidad de realizar un proceso quirúrgico. Se utilizan para inspeccionar la base del cerebro (lóbulos frontales y temporales del cerebro), a través de las diferentes vías de acceso, con electrodos diseñados especialmente para ellos. Dos ejemplos serían el faríngeo y el timpánico [Navarro09].
- Neuroquirúrgicos, son utilizados a la hora de intervenciones quirúrgicas [Navarro09].

Aunque se conocen varios sistemas de posicionamiento, la disposición de los electrodos en el cuero cabelludo se ajusta a la recomendación, propuesta en 1958 por la Federación Internacional de Sociedades de Electroencefalografía y Neurofisiología Clínica, denominado sistema “diez-veinte” (10-20) y catalogado como el método de posicionamiento estandar [Taywade12]. Este sistema se basa en la relación entre la posición de un electrodo y el área subyacente de la corteza cerebral [Rowan04]. Para ello, emplea marcas craneales que son tomadas como puntos de referencia para la localización de electrodos [Navarro09]. El 10 hace referencia a que las distancias reales entre electrodos adyacentes son el 10 %, mientras que el 20 se refiere al 20 % de la distancia total delantera/trasera o derecha/izquierda, de ma-

nera que la zona que registre cada electrodo sea comparable entre los registros realizados a diferentes individuos. Con esto se consigue que haya una división del cráneo y que podamos tomar como referencia algunos puntos para cubrir la totalidad de las regiones cerebrales, se muestra en Figura 2.3 [Sánchez14].

En conjunto, una de sus principales configuraciones es con 19 electrodos superficiales distribuidos por la cabeza y 2 electrodos de referencia en los lóbulos de las orejas, aunque se admite el uso de muchos más electrodos (32, 64, 128, 256...) [Navarro09].

Cada sitio posee una letra para identificar el lóbulo y un número para definir la posición del hemisferio. Por lo tanto, tendríamos la siguiente nomenclatura: F (frontal), C (central), T (temporal), P (posterior) y O (occipital). Además, la “z” hace mención al electrodo que se coloca en la línea media del cráneo [Teplan02].

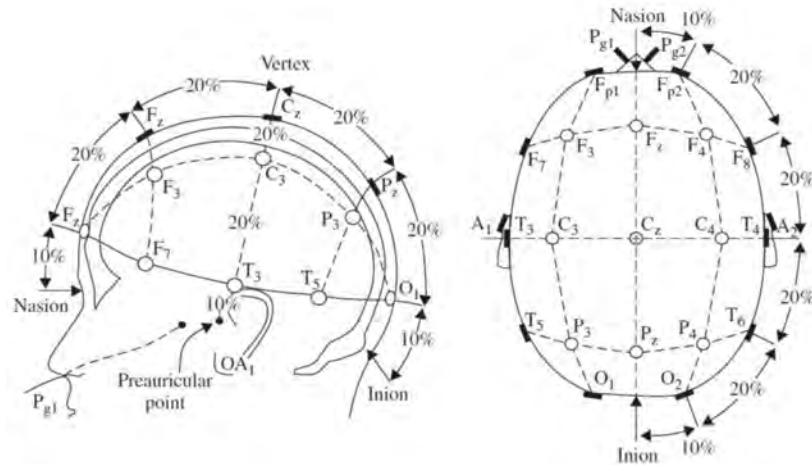


Figura 2.3: Diagrama de representación del sistema 10-20. El cerebro es mapeado por cuatro puntos: Nasal, Occipital (inión) y por preauriculares derecho e izquierdo (orejas). [Sánchez14]

Por otro lado, la presencia de números se utiliza para referirnos a la parte izquierda de la cabeza si son impares, y a la parte derecha cuando son pares [Sánchez14].

Una vez colocados los electrodos de acuerdo al sistema 10-20 es importante remarcar que cada electrodo se considera un punto de registro y que, para generar una diferencia de potencial se necesita de dos terminales, una de registro y otra de referencia [Sánchez14]. Podemos distinguir dos tipos: los registros monopolares y los bipolares (véase Figura 2.4):

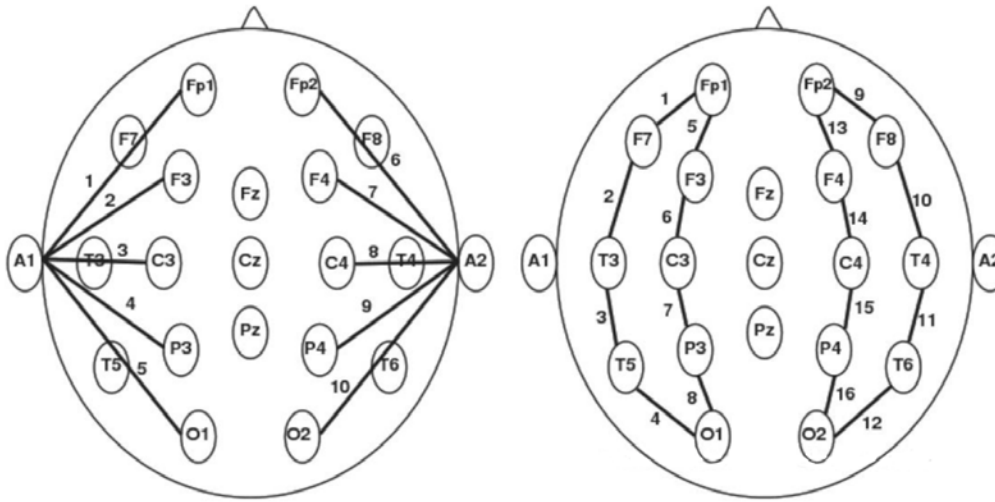


Figura 2.4: Registros. Ilustración de un registro referencial (izquierda) con las referencias A1 y A2 para los electrodos del lado izquierdo y del lado derecho, respectivamente, y un registro bipolar (derecha) con la disposición de los electrodos vecinos en cadenas longitudinales [Boutros02].

- El registro monopolar o referencial, combina dos tipos electrodos diferentes. Este registro solo considera relevante la señal de cada uno de los electrodos, independientemente de la señal de los demás [Navarro09].
- El registro bipolar o diferencial, toma las señales procedentes de parejas de electrodos activos, compara la actividad entre ellos y la elimina, registrando únicamente las diferencias de tensión entre los electrodos. Gracias a esto no se produce ninguna contaminación en los registros, pero supone una gran pérdida de información. Sin embargo, es muy útil en aplicaciones clínicas porque ayuda a localizar lugares en los que se origina una señal específica [Navarro09].

En ambos tipos la mitad de los canales son conectados al hemisferio derecho y la otra mitad al hemisferio izquierdo, de este modo la actividad eléctrica de áreas homólogas del cerebro pueden ser comparadas a simple vista [Navarro09].

## 2.3. Conclusiones sobre la detección y generalidades de la EA

La EA es una enfermedad degenerativa de difícil diagnóstico principalmente apoyado en la historia clínica del paciente, en México esta enfermedad esta teniendo un crecimiento acelerado lo cual lleva a la necesidad de buscar mejores tratamientos y formas de detección. Los cambios en la masa cerebral que se pueden observar en la RM y la TC sugieren un deterioro cognitivo es la forma más usada por los médicos junto con la historia clínica para el diagnóstico de la EA.

Las señales EEG se obtienen de forma sencilla y ofrecen mucha información de la actividad del cerebro, la cual nos sugiere que podemos ver en esta algunos trastornos mentales, tiene la ventaja de ser un estudio más barato que el de TC o RM, además de tener menos contra indicaciones. La forma de colocar los electrodos para obtener las señales es importante para obtener los resultados deseados, el sistema 10-20 aporta una estandarización en la colocación, se sugiere ser monopolar o bipolar depende el propósito que se tenga.

Las señales EEG para hacer el diagnóstico de la EA es algo que se ha llevado trabajando por los investigadores desde hace algún tiempo, se han logrado avances y el conocimiento de que es un método que funciona para el diagnóstico, pero sigue sin ser muy utilizado en las practicas de los médicos especialistas.





## Capítulo 3

# Extracción de características

En el presente capítulo se describen los métodos que obtienen las características necesarias para poder determinar dos grupos, las personas sanas y las que tienen la EA. La información necesaria que se toma como base para los métodos descritos se obtiene de las señales de EEG. Las características elegidas debido a que muestran gran aporte en el estado del arte son: entropía espectral, energía por bandas de frecuencia y coeficiente de correlación de Pearson, de estos valores se buscarán los que mejor contribuyan a la separación de los dos grupos y se formará con ellos el vector de características.

### 3.1. Entropía espectral

Uno de los problemas desafiantes para la extracción de características de datos de EEG es que la señal de EEG es compleja, no lineal, no estática y de naturaleza aleatoria. Se consideran estacionarios solo dentro de intervalos cortos, es decir, “cuasiestacionarios”, durante períodos de tiempo más largos, las características de la señal no son estacionarias [Phung14]. Como consecuencia, numerosos métodos de extracción de características lineales a menudo aplican una técnica de ventana de corto tiempo a las señales de EEG para cumplir el requisito de ser cuasiestacionarios. Sin embargo, esta suposición se mantiene durante una condición cerebral normal, pero durante las actividades mentales y físicas esta suposición no es válida. Se puede observar una señal EEG no estacionaria durante el cambio de estado de alerta y vigilia, durante el parpadeo de los ojos, entre otras estimulaciones.

La entropía es una medida de incertidumbre. En los sistemas de interfaz cerebro-computadora, la entropía puede usarse para medir el nivel de caos del sistema. Es una medida no lineal que cuantifica el grado de complejidad en una serie temporal [Phung14].

Shannon definió el concepto de información de entropía como el valor esperado, o la cantidad promedio, de la información de una distribución de probabilidad. La definición de entropía de Shannon se ha aplicado, modificado y probado como válido en una variedad de campos. En 1979, Powell y Percival introdujeron la entropía espectral, basada en los picos de la transformada de Fourier, como una medida de la regularidad [Abásolo06]. Esta definición también se extendió a la densidad espectral de potencia relativa del EEG y se puede calcular con la expresión que se define como:

$$H(f) = -\frac{1}{\ln(N)} \sum_{i=1}^N p_i \ln(p_i) \quad (3.1)$$

Donde  $p_i$  es la probabilidad para el punto  $i$  en una distribución de probabilidad, la suma de todos los  $p_i$  es igual a 1 y  $N$  es la longitud del arreglo. La entropía refleja qué tan bien uno puede predecir el comportamiento de cada parte respectiva de la trayectoria desde el otro. Básicamente, una entropía más alta indica sistemas más complejos o caóticos, por lo tanto, menos predictibilidad [Phung14].

### 3.2. Coeficiente de correlación de Pearson

El coeficiente  $r$  de Pearson, también conocido como la Correlación producto-momento de Pearson, se usa a menudo en los paquetes de software modernos disponibles para visualización de datos y ajuste de curvas [Hall15]. Para el caso del estudio de señales EEG, el coeficiente de correlación de Pearson se usa para comparar dos regiones o hemisferios cerebrales, como se menciona en el estado del arte, eso proporciona información sobre la transferencia de información en el cerebro y una voy baja correspondencia puede indicar una alteración cognitiva.

Se define como la razón de la covarianza de dos variables que representan un conjunto de datos numéricos, normalizados a la raíz cuadrada de sus varianzas [Hall15], es decir:

$$r = \frac{C_{xy}}{\sqrt{C_{xx}C_{yy}}} = \frac{C_{xy}}{\sigma_x\sigma_y} \quad (3.2)$$

Con más detalle se formula: para un conjunto de  $N$  puntos de datos bidimensionales  $[x_1, x_2, \dots, x_N]$  y  $[y_1, y_2, \dots, y_N]$ , se tiene:

$$\bar{x} = \frac{1}{N} \sum_i x_i, \quad \bar{y} = \frac{1}{N} \sum_i y_i \quad (3.3)$$

$$C_{xy} = \frac{1}{N-1} \sum_i (x_i - \bar{x})(y_i - \bar{y}) \quad (3.4)$$

$$C_{xx} = \sigma_x^2 = \frac{1}{N-1} \sum_i (x_i - \bar{x})^2 \quad (3.5)$$

$$C_{yy} = \sigma_y^2 = \frac{1}{N-1} \sum_i (y_i - \bar{y})^2 \quad (3.6)$$

Dos propiedades importantes pueden notarse inmediatamente [Hall15]:

- $r$  depende solo de los valores de los datos y la dispersión.
- $r$  no depende de errores en las cantidades medidas.

Esto nos dice inmediatamente que  $r$  no está evaluando nada más que alguna relación interna entre los valores de los datos; no está probando una hipótesis para el origen de los datos, ni está dando más peso a algunos puntos de datos, por ejemplo porque se miden con mayor precisión que otros.

De hecho, otra suposición de la estadística de Pearson es que la relación que se probará es lineal [Hall15]. En este caso, el resultado es fácil de derivar.

En otras palabras, si  $x$  y  $y$  están exactamente relacionados linealmente,  $r = \pm 1$ , dependiendo de si la pendiente es positiva o negativa (correlación o anticorrelación). Es más probable que con datos reales de cualquier tipo, haya una dispersión en los valores de  $x$  y  $y$ , en cuyo caso la correlación será menor que la máxima, es decir  $|r| < 1$ .

En la Figura 3.1 se muestran algunos ejemplos de datos simulados llamados E y R con fluctuaciones gaussianas aleatorias de diferentes magnitudes aplicadas, y los valores

de  $r$  resultantes. En este caso, dado que los datos se generaron a partir de una relación lineal, que se muestra en Figura 3.1(a), los valores más bajos de  $r$  deben entenderse como mayores fluctuaciones sobre la tendencia, no como ausencia de correlación o ajustes peores a los valores de datos.

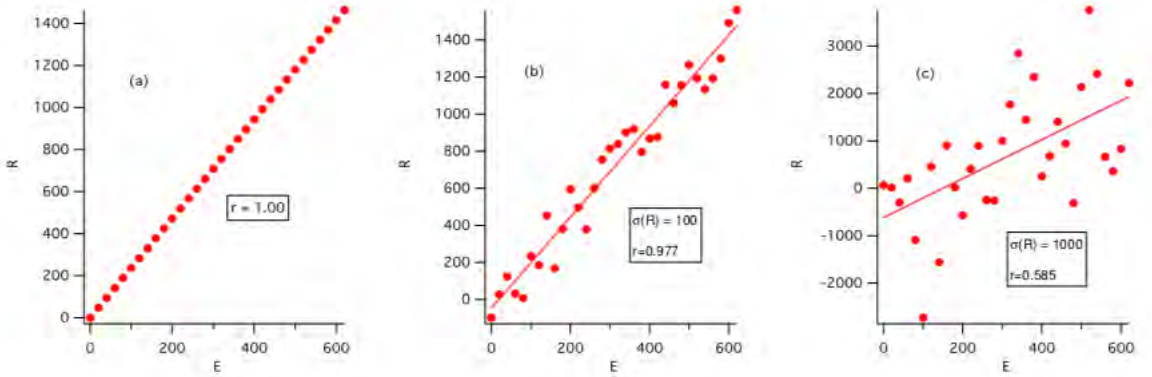


Figura 3.1: Datos simulados siguiendo una relación lineal entre  $R$  y  $E$  con fluctuaciones gaussianas aleatorias agregadas a los valores  $R$ . (a) sin fluctuaciones, resultando en  $r = 1$ . (b)  $\sigma(R) = 100$  resultando en  $r = 0.977$ . (c)  $\sigma(R) = 1000$ , y  $r = 0.585$ . Teniendo en cuenta el cambio de escala vertical [Hall15].

En vista del hecho de que  $r$  no incluye errores experimentales en los puntos de datos, que a menudo no son idénticos para cada punto de datos, es poco probable que los parámetros de ajuste calculados o sus errores basados en el coeficiente de Pearson sean confiables, excepto en el caso fortuito que la dispersión de los puntos de datos coincide estrechamente con las fluctuaciones experimentales.

### 3.3. Separación ciega de fuentes y selección de características

La separación ciega de fuentes (SCF) consiste en que las observaciones o mediciones ( $\mathbf{X}$ ) son el resultado de señales mezcladas que provienen de fuentes independientes ( $\mathbf{Z}$ ). Al establecer que las observaciones están dadas por la mezcla de las fuentes, las fuentes están mezcladas de manera natural con una matriz ( $\mathbf{A}$ ) para obtener las observaciones [Puigt11], esto se resume en un problema de sistemas lineales (3.7),  $\mathbf{X}$  y  $\mathbf{Z}$  transpuestos debido a que son vectores filas y no columnas de los canales.

$$\mathbf{X}^T = \mathbf{A}\mathbf{Z}^T \quad (3.7)$$

Se asume que se tienen el mismo número de observaciones que de fuentes. El objetivo de la SFC es poder separar las fuentes independientes solo conociendo las observaciones (solo se conoce  $\mathbf{X}$ , mientras que  $\mathbf{A}$  y  $\mathbf{Z}$  son desconocidas). En Figura 3.2 se muestra un ejemplo de SCF, donde se muestra a dos parlantes, los cuales darían los valores de las fuentes  $Z_1$  y  $Z_2$ , al estar ambos en una habitación con un micrófono destinado para cada uno se mezclan al grabar y produce las observaciones, mostrando que si estas observaciones pasan por un proceso de separación deben corresponder con las fuentes.



Figura 3.2: Separación de las observaciones para recuperar las fuentes [Puigt11].

Considerando que las señales son matrices que tienen  $N$  canales que representa la cantidad de fuentes u observaciones por el número de muestras ( $J$ ), entonces  $\mathbf{X}$  y  $\mathbf{Z}$  son de tamaño  $N \times J$ , mientras que la matriz de mezcla  $\mathbf{A}$  debe de ser de tamaño de la cantidad de fuentes por la cantidad de observaciones, siendo de tamaño  $N \times N$ .

Otro problema que ejemplifica de la forma más sencilla a (3.7), es el problema de la fiesta de cócteles, donde se observan varias personas que generan fuentes de sonido y micrófonos grabando [Clifford07], donde cada micrófono graba las señales mezcladas de las fuentes, que son las personas hablando, como se muestra en Figura 3.3.

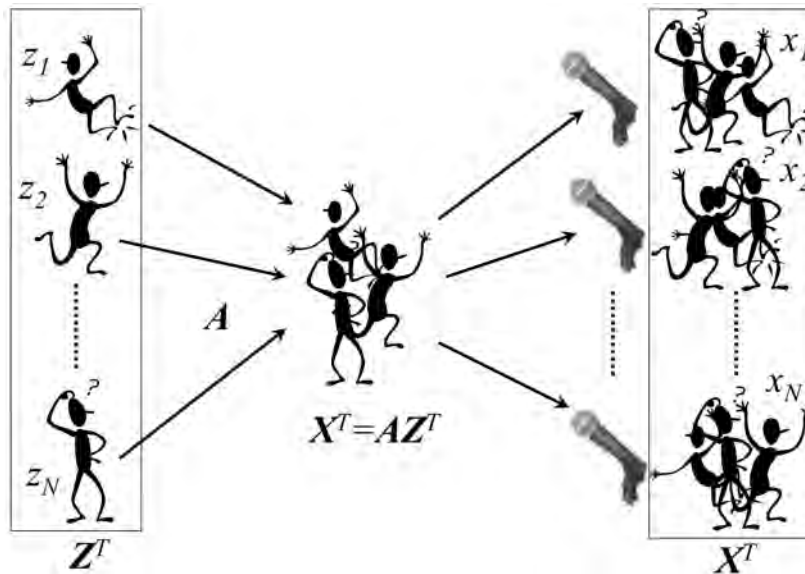


Figura 3.3: Problema de la fiesta de cócteles[Clifford07].

El efecto que produce la SCF a un conjunto de series de tiempo es transformarlas para manejarlas en otro dominio, así como la transformada de Fourier nos permite manejar los datos en el dominio de la frecuencia, la SCF nos genera nuevos datos en un dominio estadístico, dejando de lado el dominio del tiempo o la frecuencia, los datos se proyectan en nuevos ejes que cumplen algún criterio estadístico, lo que implica independencia, en lugar de ejes que representan frecuencias discretas o coeficientes de frecuencia, como ocurre en la transformada de Fourier, donde se supone independencia [Clifford08]. Existen tres familias de métodos para la SCF [Puigt11]: análisis de componentes independientes (ICA), análisis de componentes dispersos (SCA) y la factorización de matriz no negativa (NMF). En este capítulo se examinara el método de ICA y las técnicas de análisis de componentes principales (PCA) mediante la descomposición de valores singulares (SVD), que constituyen la base para comprender al ICA.

Cualquier proyección en otro conjunto de ejes o en otro espacio, es esencialmente un método para separar los datos en componentes separados o fuentes que esperamos nos permitan ver una estructura importante más claramente en una proyección particular. Esto es cierto tanto para ICA y PCA como para las técnicas basadas en Fourier. En PCA e

ICA se busca encontrar un conjunto de ejes que sean independientes entre sí en algún sentido [Clifford08]. Una diferencia importante entre estas técnicas estadísticas y las técnicas basadas en Fourier es que los componentes de Fourier en los que se proyecta un segmento de datos son fijos, relacionados con la frecuencia, mientras que las transformaciones basadas en PCA o ICA dependen de la estructura de los datos que se analizan. Por lo tanto, se descubren los ejes sobre los que se proyectan los datos. Si la estructura de los datos cambia con el tiempo, los ejes sobre los que se proyectan los datos también cambiarán [Clifford08].

Suponiendo que hay un conjunto de fuentes independientes en los datos, pero que no asumen sus propiedades exactas, se define alguna medida de independencia, como puede ser la covarianza, y se intenta descorrelacionar los datos maximizando esta medida para proyecciones en cada eje del nuevo espacio en el que hemos transformado los datos. Las fuentes son los datos proyectados en cada uno de los nuevos ejes. Es decir, no buscamos componentes predefinidos específicos, como la energía a una frecuencia específica, sino que permitimos que los datos determinen los componentes [Clifford08].

Para PCA, la medida usada para descubrir los ejes es la covarianza y conduce a un conjunto de ejes ortogonales, debido a que el producto escalar de cualquier par de los ejes recién descubiertos es cero [Clifford08]. Para ICA, esta medida se basa en la curtosis, y los ejes no son necesariamente ortogonales. La curtosis expresada en (3.8) siendo  $X$  una variable aleatoria, es el cuarto momento (la media, la varianza y la asimetría son los tres primeros momentos) y es una medida de cómo la función de distribución de probabilidad es más pronunciada en el pico [Clifford08].

$$\kappa = \frac{E[(X - \mu_x)^4]}{\sigma^4} = E[X^4] - 3(E[X^2])^2 \quad (3.8)$$

Los valores positivos grandes de curtosis indican una función de distribución de probabilidad muy puntiaguda que es mucho más estrecha que una gaussiana. Una curtosis con valores menores a 3 indica una función de distribución de probabilidad ancha que es más amplia que una gaussiana, como se observa en Figura 3.4.

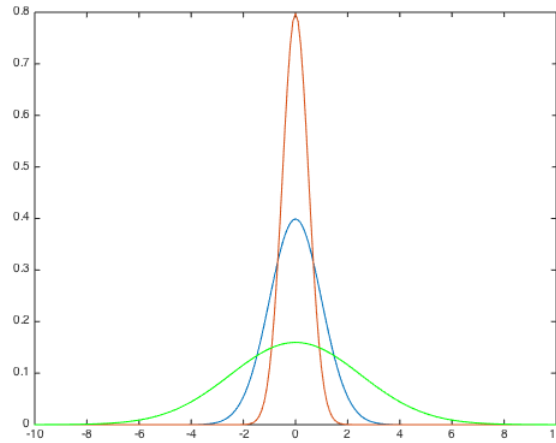


Figura 3.4: Distribución de probabilidad normal con valor de curtosis en 3 (azul) y distribuciones de probabilidad con valores de curtosis mayor a 3 (rojo) y menor a 3 (verde).

La suposición es que si se maximiza la no gaussianidad de un conjunto de señales, entonces son máximamente independientes. Esto proviene del teorema del límite central [Trotter59]; si seguimos agregando señales independientes, que tienen una función de distribución de probabilidad altamente no gaussiana, eventualmente llegaremos a una distribución gaussiana. Por el contrario, si dividimos una observación similar a una gaussiana en un conjunto de mezclas no gaussianas, cada una con distribuciones que son tan no gaussianas como sea posible, las señales individuales serán independientes. Por lo tanto, la curtosis nos permite separar las fuentes independientes no gaussianas, mientras que la varianza nos permite separar las fuentes de ruido gaussianas independientes [Clifford08].

### 3.3.1. Matrices de transformación como filtros

El filtrado más simple de una serie de tiempo implica la transformación de una serie de tiempo discreta  $\mathbf{x}[m]$  de una dimensión ( $N = 1$ ), que consta de  $M$  muestras tales que  $\mathbf{x} = (x_1, x_2, x_3 \dots x_M)^T$ , en una nueva representación,  $\mathbf{y} = (y_1, y_2, y_3 \dots y_M)^T$ . Si  $\mathbf{x}$  es un vector columna que representa un canal de EEG [Clifford08], entonces podemos generalizar esta representación para que  $N$  canales de EEG,  $\mathbf{X}$  y su representación transformada  $\mathbf{Y}$  son dados por:



$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1N} \\ x_{21} & x_{22} & \dots & x_{2N} \\ \vdots & \vdots & \dots & \vdots \\ x_{M1} & x_{M2} & \dots & x_{MN} \end{bmatrix}, \quad \mathbf{Y} = \begin{bmatrix} y_{11} & y_{12} & \dots & y_{1N} \\ y_{21} & y_{22} & \dots & y_{2N} \\ \vdots & \vdots & \dots & \vdots \\ y_{M1} & y_{M2} & \dots & y_{MN} \end{bmatrix} \quad (3.9)$$

Los  $M$  puntos de cada uno de los  $N$  canales de las señales forman matrices  $M \times N$ . Una matriz ( $N \times N$ ) de transformación  $\mathbf{W}$  puede aplicarse entonces a  $\mathbf{X}$  para crear la matriz transformada  $\mathbf{Y}$  de forma que

$$\mathbf{Y}^T = \mathbf{W}\mathbf{X}^T \quad (3.10)$$

El propósito de una transformación es mapear o proyectar los datos a otro espacio que sirve para resaltar diferentes patrones en los datos a lo largo de diferentes ejes de proyección [Clifford08]. En general, las transformaciones se pueden clasificar como transformaciones ortogonales o biortogonales. Para las transformaciones ortogonales, la señal transformada tiene la misma longitud ( $M$ ) que la original y la energía de los datos no se modifica [Clifford08]. Un ejemplo de esto es la Transformada Discreta de Fourier (DFT) donde la misma señal se mide a lo largo de un nuevo conjunto de ejes perpendiculares que corresponden a los coeficientes de la serie de Fourier. Para las transformaciones biortogonales, los ángulos entre los ejes pueden cambiar y los nuevos ejes no son necesariamente perpendiculares [Clifford08]. Tienen inversa, solo que su inversa no es su transpuesta como en las ortogonales, sin embargo no se pierde información y aún es posible la reconstrucción perfecta de la señal original, mediante:

$$\mathbf{X}^T = \mathbf{W}^{-1}\mathbf{Y}^T \quad (3.11)$$

Las transformaciones se pueden clasificar además como sin pérdida (de modo que la transformación se puede revertir y los datos originales se restauran exactamente) o como con pérdida. Cuando una señal se filtra o se comprime (por ejemplo, a través de una reducción de resolución), la información a menudo se pierde y la transformación no es invertible [Clifford08]. Por lo tanto, hay una eliminación irreversible de algunos de los datos  $N$ -dimensional y esto corresponde a mapear a un número menor de dimensiones ( $p < N$ ).

Las técnicas de PCA y de ICA intentan encontrar un conjunto independiente de vectores en el que podamos transformar los datos. Los datos que se proyectan en cada vector son las fuentes independientes. El objetivo básico en PCA es descorrelacionar la señal proyectando los datos en ejes ortogonales. Sin embargo, ICA da como resultado una transformación biortogonal de los datos y los ejes no son necesariamente ortogonales. Tanto PCA como ICA se pueden usar para realizar transformaciones sin pérdidas o con pérdidas al multiplicar los datos grabados o observaciones por una matriz de separación o desmezcla. Tanto PCA como ICA sin pérdidas implican proyectar los datos en un conjunto de ejes que están determinados por la naturaleza de los datos y, por lo tanto, son métodos de SCF [Puigt11].

Una vez que hemos descubierto los ejes de los componentes independientes en los datos y los hemos separado proyectando los datos en estos ejes, se pueden utilizar estas técnicas para filtrar los datos. Al establecer en cero las columnas de las matrices de separación de PCA e ICA que corresponden a las fuentes no deseadas, producimos matrices no invertibles [Clifford08]. Si forzamos la inversión de la matriz de separación y transformamos los datos nuevamente en el espacio de observación original, podemos eliminar la fuente no deseada de la señal original.

En Figura 3.5 se ilustra el paradigma de la SCF para el filtrado, en el cual tenemos  $N$  fuentes desconocidas en un espacio fuente desconocido que se mezclan linealmente y se transforman en un espacio de observación en el que se graban. Luego intentamos descubrir las fuentes  $\mathbf{Z}$ , o el inverso de la matriz de mezcla  $\mathbf{W} \approx \mathbf{A}^{-1}$ , y usar esto para transformar los datos en una estimación de nuestro espacio de origen.

Después de identificar las fuentes de interés y descartar aquellas que no queremos (al alterar el inverso de la matriz de desmezcla para tener columnas de ceros para las fuentes no deseadas), re proyectamos los datos nuevamente en el espacio de observación usando el inverso de la desmezcla alterada matriz,  $\mathbf{W}_p^{-1}$  en el caso de ICA. La información resultante  $\mathbf{X}_{filt}$ , es una versión filtrada de los datos originales  $\mathbf{X}$ .

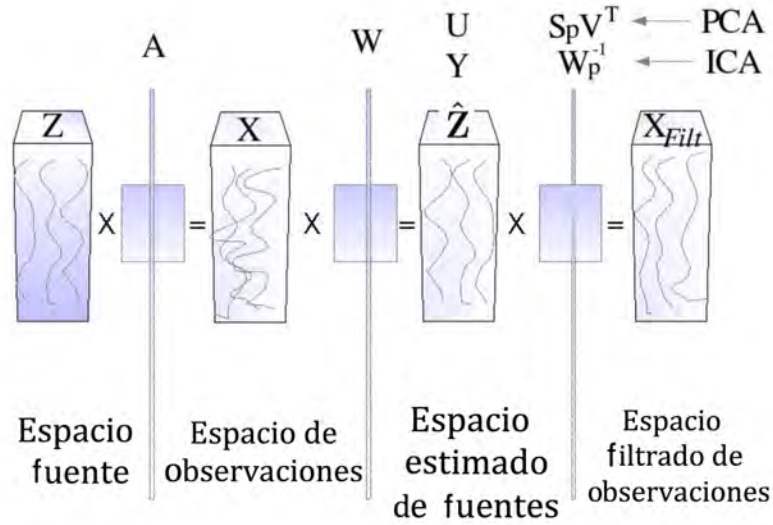


Figura 3.5: El paradigma general de la SCF para el filtrado,  $S_p$  es la matriz de valores singulares, y  $V^T$  la matriz de vectores columna propios para el PCA, se observa en  $Z$  la representación de señales dentro del bloque, y en  $X_{filt}$  después del proceso de reconstrucción una versión filtrada de los datos [Clifford08].

Formalmente, se puede describir el paradigma general de la SCF para el filtrado [Clifford08]: dada una matriz desconocida de fuentes  $Z$  que está mezclada por alguna matriz estacionaria lineal de constantes  $A$ , nuestras fuentes se proyectan desde un espacio fuente a un espacio de observación para dar las observaciones,  $X$ . Estas observaciones se mapean de nuevo a un espacio de origen estimado en el que se proyectan las estimaciones de las fuentes  $\hat{Z}$ . Luego reducimos la dimensionalidad del espacio de origen estimado, descartando las estimaciones de las fuentes que corresponden a ruido o artefactos no deseados poniendo  $N - p$  columnas de  $W^{-1}$  a cero (para dar  $W_p^{-1}$ ) y re proyectando de nuevo en el espacio de observación. La matriz resultante de observaciones filtradas es  $X_{filt}$ . El espacio de observación filtrado y el espacio de observación original son los mismos, pero los datos proyectados en ellos se filtran y no se filtran, respectivamente. En el caso de ICA,  $X$  puede multiplicarse por la matriz de desmezcla  $W$ , para revelar las estimaciones de las fuentes,  $Y = \hat{Z}$ . Las columnas de  $W^{-1}$  se pueden establecer en cero para eliminar las fuentes de “ruido” y los datos filtrados se reconstruyen con  $X_{filt} = W_p^{-1}Y$ .

### 3.3.2. Análisis de componentes principales (PCA)

En el caso de la transformada de Fourier, las funciones de base o los ejes de la nueva representación están predefinidos y se supone que son independientes, mientras que con PCA la representación, o los vectores de base, se encuentran en los datos buscando un conjunto de ejes que son independientes. Es decir, los datos se someten a una descorrelación usando la varianza como la métrica [Clifford08]. Las proyecciones en estos ejes, o vectores de base, son independientes en un sentido de segundo orden y son ortogonales (el producto punto de los vectores de base, y la correlación cruzada de las proyecciones es cercana a cero).

La idea básica en la aplicación de PCA a un conjunto de datos, es encontrar los vectores componentes  $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N$  que impliquen la cantidad máxima de varianza posible por  $N$  componentes transformados linealmente. Se puede definir PCA de una forma intuitiva utilizando una formulación recursiva [Clifford08]. La dirección del primer componente principal  $v_1$  se encuentra pasando los datos e intentando maximizar el valor de  $\mathbf{v}_1 = \underset{\|\mathbf{v}\|=1}{\operatorname{argmax}} E[(\mathbf{v}_1^T \mathbf{X})^2]$  donde  $\mathbf{v}_1$  tiene la misma longitud  $M$  que los datos  $\mathbf{X}$ . Así, el primer componente principal es la proyección sobre la dirección en la cual la varianza de la proyección se maximiza.

Cada uno de los componentes principales  $N - 1$  restantes se encuentran repitiendo este proceso en el subespacio ortogonal restante, que reduce en dimensionalidad en uno para cada nuevo componente que se descubre. Los componentes principales están dados por  $\mathbf{y}_i = \mathbf{v}_i^T \mathbf{X}$  ( $i = 1, \dots, N$ ), la proyección de  $\mathbf{X}$  en cada  $\mathbf{v}_i$ . Esta transformación de las columnas de  $\mathbf{X}$  sobre  $\mathbf{v}_i^T$ , para dar  $\mathbf{y}_i$  también se conoce como la transformada discreta de Karhunen-Loève [Dony01].

Aunque el objetivo básico en PCA es descorrelacionar los datos mediante la realización de una proyección ortogonal, a menudo reducimos la dimensión de los datos de  $N$  a  $p$  ( $p < N$ ) para eliminar componentes no deseados en la señal. Se puede demostrar que la representación de PCA es una técnica de reducción de dimensión lineal óptima en el sentido cuadrado medio [Jolliffe88]. Una aplicación importante de esta técnica es la reducción del ruido, donde se supone que los datos contenidos en los últimos componentes  $N - p$  se

deben principalmente al ruido. Otro beneficio de esta técnica es que una proyección en un subespacio de muy baja dimensión, por ejemplo dos o tres, puede ser útil para visualizar datos multidimensionales o de orden superior.

En la práctica, el cálculo de la  $\mathbf{v}_i$  se puede realizar simplemente utilizando la matriz de covarianza de la muestra  $\mathbf{C} = \mathbf{X}^T \mathbf{X}$ . Los  $\mathbf{v}_i$  son los vectores propios de  $\mathbf{C}$  (una matriz  $M \times M$ ) que corresponden a los  $N$  valores propios de  $\mathbf{C}$ . Un método para determinar los valores propios de esta manera se conoce como SVD [Abdi07].

### Método de descomposición del valor singular (SVD)

Para determinar los componentes principales de una señal multidimensional, podemos usar el método de SVD. Considere una matriz real  $\mathbf{X}$  de observaciones  $M \times N$  que puede descomponerse de la siguiente manera:

$$\mathbf{X} = \mathbf{U} \mathbf{S} \mathbf{V}^T \quad (3.12)$$

Donde  $\mathbf{S}$  es una matriz no cuadrada  $M \times N$  con cero entradas en todas partes, excepto en la diagonal con los elementos  $s_i$  dispuestos en orden descendente de magnitud, por ejemplo:

$$\mathbf{S} = \begin{bmatrix} s_1 & 0 \\ 0 & s_2 \\ 0 & 0 \end{bmatrix}$$

Cada  $s_i$  es igual a  $\sqrt{\lambda_i}$ , la raíz cuadrada de los valores propios de  $\mathbf{C} = \mathbf{X}^T \mathbf{X}$ . Una gráfica de tallos de estos valores contra su índice  $i$  se conoce como espectro singular o espectro propio [Clifford08].

Cuanto menor es el valor propio, menor es la energía total que se proyecta a lo largo del vector propio correspondiente. Por lo tanto, los valores propios más pequeños a menudo se consideran asociados con vectores propios que describen el ruido en la señal. Las columnas de  $\mathbf{V}$  forman una matriz  $N \times N$  de vectores de columna, que son los vectores propios de  $\mathbf{C}$ . La matriz  $\mathbf{U}$  de  $M \times M$  es la matriz de proyecciones de  $\mathbf{X}$  sobre los vectores propios de  $\mathbf{C}$  [Golub83]. Se puede realizar un SVD truncado de  $\mathbf{X}$  de manera que solo se

conservan los vectores propios más significativos ( $p$  más grandes). El SVD truncado viene dado por  $\mathbf{Y} = \mathbf{U}\mathbf{S}_p\mathbf{V}^T$  y las columnas de la matriz  $\mathbf{Y}$  de  $M \times N$  son la señal de ruido reducido.

Una rutina para realizar SVD [Golub83], es la siguiente:

1. Encuentre los  $N$  valores propios distintos de cero,  $\lambda_i$  de la matriz  $\mathbf{C} = \mathbf{X}^T\mathbf{X}$  y forme una matriz diagonal no cuadrada  $\mathbf{S}$  colocando las raíces cuadradas  $s_i = \sqrt{\lambda_i}$  de los  $N$  valores propios en orden descendente de magnitud en la diagonal delantera y estableciendo todos los demás elementos de  $\mathbf{S}$  a cero.
2. Encontrar los vectores propios ortogonales de la matriz  $\mathbf{X}^T\mathbf{X}$  correspondientes a los valores propios obtenidos, y organícelos en el mismo orden, esta colección ordenada de vectores de columna forma la matriz  $\mathbf{V}$ .
3. Encuentre los primeros  $N$  vectores de columna de la matriz  $\mathbf{U} : \mathbf{u}_i = s_i^{-1}\mathbf{X}\mathbf{v}_i (i = 1 : N)$ . Tenga en cuenta que  $s_i^{-1}$  son los elementos de  $\mathbf{S}^{-1}$ .
4. Agregue el resto de vectores  $M - N$  a la matriz  $\mathbf{U}$  usando el proceso de ortogonalización de Gram-Schmidt [Lambers13].

### 3.3.3. Ejemplo de descomposición de valores propios

Para encontrar la descomposición de valores singulares de la matriz en el ejemplo planteado [Clifford08],

$$\mathbf{X} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$$

primero encontramos los valores propios,  $\lambda$ , de la matriz

$$\mathbf{C} = \mathbf{X}^T\mathbf{X} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

de la manera habitual, dejando

$$\mathbf{C}\mathbf{v} = \lambda\mathbf{v} = 0$$

entonces  $(\mathbf{C} - \lambda\mathbf{I} = 0)$  y

$$\begin{vmatrix} 2 - \lambda & 1 \\ 1 & 2 - \lambda \end{vmatrix} = 0$$

Evaluando este determinante y resolviendo esta ecuación característica para  $\lambda$ , se encuentra  $(2 - \lambda)^2 - 1 = 0$ , y así  $\lambda_1 = 3$  y  $\lambda_2 = 1$ . Luego notamos el número de valores propios distintos de cero de la matriz  $\mathbf{X}^T\mathbf{X}$  (dos en este caso). Se encuentran los vectores propios ortonormales de la matriz  $\mathbf{X}^T\mathbf{X}$  correspondientes a los valores propios distintos de cero ( $\lambda_1$  y  $\lambda_2$ ) resolviendo para  $\mathbf{v}_1$  y  $\mathbf{v}_2$  usando  $\lambda_1$  y  $\lambda_2$  y  $(\mathbf{C} - \lambda\mathbf{I})\mathbf{v} = 0$ .

$$\mathbf{v}_1 = \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} \end{bmatrix}$$

Formando la matriz

$$\mathbf{V} = [\mathbf{v}_1 \mathbf{v}_2] = \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{bmatrix}$$

donde  $\mathbf{v}_1$  y  $\mathbf{v}_2$  están normalizados. A continuación, escribimos la matriz de valores singulares  $\mathbf{S}$ , que es una matriz diagonal compuesta por las raíces cuadradas de los valores propios de  $\mathbf{C} = \mathbf{X}^T\mathbf{X}$  dispuestos en orden descendente de magnitud.

$$\mathbf{S} = \begin{bmatrix} s_1 & 0 \\ 0 & s_2 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \sqrt{(\lambda_1)} & 0 \\ 0 & \sqrt{(\lambda_2)} \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \sqrt{3} & 0 \\ 0 & \sqrt{1} \\ 0 & 0 \end{bmatrix}$$

Recordando que el inverso de una matriz  $\mathbf{D}$  con elementos  $d_{ij}$  esta dado por:

$$\mathbf{D}^{-1} = \frac{1}{\det|\mathbf{D}|} \begin{bmatrix} d_{22} & -d_{21} \\ -d_{12} & d_{11} \end{bmatrix}$$

entonces:

$$\mathbf{S}^{-1} = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{3} \end{bmatrix}$$

Se encuentran las primeras dos columnas de  $\mathbf{U}$ , usando la relación

$$\mathbf{u}_1 = s_1^{-1} \mathbf{X} \mathbf{v}_1 = \frac{\sqrt{3}}{3} \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{6}}{3} \\ \frac{\sqrt{6}}{6} \\ \frac{\sqrt{6}}{6} \end{bmatrix}$$

y

$$\mathbf{u}_2 = s_2^{-1} \mathbf{X} \mathbf{v}_2 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} \end{bmatrix} = \begin{bmatrix} 0 \\ -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}$$

Usando el proceso de Gram-Schmidt [Lambers13], se calcula la tercera columna ortogonal faltante de  $\mathbf{U}$

$$\mathbf{u}_3 = s_2^{-1} \mathbf{X} \mathbf{v}_2 = \begin{bmatrix} \frac{\sqrt{3}}{3} \\ -\frac{\sqrt{3}}{3} \\ -\frac{\sqrt{3}}{3} \end{bmatrix}$$

Por lo tanto

$$\mathbf{U} = [\mathbf{u}_1 \mathbf{u}_2 \mathbf{u}_3] = \begin{bmatrix} \frac{\sqrt{6}}{3} & 0 & \frac{\sqrt{3}}{3} \\ \frac{\sqrt{6}}{6} & \frac{\sqrt{2}}{2} & -\frac{\sqrt{3}}{3} \\ \frac{\sqrt{6}}{6} & -\frac{\sqrt{2}}{2} & -\frac{\sqrt{3}}{3} \end{bmatrix}$$

y la SVD de la matriz  $\mathbf{X}$  es:

$$\mathbf{X} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} = \mathbf{U} \mathbf{S} \mathbf{V}^T = \begin{bmatrix} \frac{\sqrt{6}}{3} & 0 & \frac{\sqrt{3}}{3} \\ \frac{\sqrt{6}}{6} & \frac{\sqrt{2}}{2} & -\frac{\sqrt{3}}{3} \\ \frac{\sqrt{6}}{6} & -\frac{\sqrt{2}}{2} & -\frac{\sqrt{3}}{3} \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 \\ 0 & \sqrt{1} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{bmatrix}$$

Se puede observar en el ejemplo anterior el proceso para la descomposición para determinar los componentes principales de una señal multidimensional, los conllevan al proceso de PCA.



### 3.3.4. Análisis de componentes independientes (ICA)

En esta sección, se examinará cómo, elegir una medida de independencia distinta de la varianza (o covarianza) puede conducir a un método más efectivo para separar señales. El ICA es una búsqueda de vectores de proyección en los que las características son más independientes entre sí [Clifford08]. Considere que  $\mathbf{X}_{jn}$  es una matriz de  $J$  vectores aleatorios observados,  $\mathbf{A}$  una matriz de mezcla  $N \times N$  y  $\mathbf{Z}$ , los  $J$  vectores fuente tales que forman la ecuación (3.2), recordando que tiene la forma  $\mathbf{X}^T = \mathbf{AZ}^T$ .

Los algoritmos ICA intentan encontrar una matriz de separación o desmezcla  $\mathbf{W}$ , tal y como se observa en la ecuación (3.10), que tiene la forma  $\mathbf{Y}^T = \mathbf{WX}^T$ , donde  $\mathbf{W} = \hat{\mathbf{A}}^{-1}$ , una aproximación del inverso de la matriz de mezcla original, y  $\mathbf{Y}^T = \hat{\mathbf{Z}}^T$ .

Se supone que estas fuentes son estadísticamente independientes, generadas por procesos no relacionados, por lo tanto, la función de densidad de probabilidad conjunta es el producto de las densidades para todas las fuentes (3.13), donde  $p(z_i)$  es la función de densidad de probabilidad de la  $i$ -ésima fuente y  $P(Z)$  es la función de densidad conjunta [Clifford08].

$$P(Z) = \prod p(z_i) \quad (3.13)$$

La idea básica de ICA es aplicar operaciones a los datos observados  $\mathbf{X}^T$ , o la matriz de desmezcla  $\mathbf{W}$ , y medir la independencia entre los canales de señal de salida, (las columnas de  $\mathbf{Y}^T$ ) para obtener estimaciones de las fuentes (las columnas de  $\mathbf{Z}^T$ ). En la práctica, los métodos iterativos se utilizan para maximizar o minimizar una función de costo determinada, como la información mutua, la entropía o el momento de cuarto orden, la curtosis, una medida de no gaussianidad [Clifford08].

Veremos más adelante cómo las funciones de costo, hablando de la cantidad de iteraciones que se tienen que realizar y como avanza un valor determinado, basadas en la entropía están relacionadas con la curtosis y, por lo tanto, todas las funciones de costo son, hasta cierto punto, una medida de la no gaussianidad. Del Teorema del Límite Central ([Trotter59]), sabemos que la distribución de una suma de variables aleatorias independientes tiende hacia una distribución Gaussiana [Clifford08]. Es decir, una suma de dos variables

aleatorias independientes generalmente tiene una distribución más cercana a Gauss que las dos variables aleatorias originales. En otras palabras, la independencia no es gaussiana. En ICA, si deseamos encontrar fuentes independientes, debemos encontrar una matriz de desmezcla  $\mathbf{W}$  que maximice la no gaussianidad de cada fuente.

Ahora veremos más de cerca lo que significa la curtosis de una distribución, y cómo esto nos ayuda a separar las fuentes componentes dentro de una señal al imponer la independencia. Los primeros dos momentos de variables aleatorias son bien conocidos; la media y la varianza [Clifford08]. Si una distribución es gaussiana, entonces la media y la varianza son suficientes para caracterizar la variable. Sin embargo, si la función de densidad de probabilidad de una función no es gaussiana, funciones de densidad diferentes pueden tener la misma media y varianza [Press89].

La media (o tendencia central) de una variable aleatoria  $X$ , se define como

$$\mu_X = E[X] = \int_{-\infty}^{\infty} X p_X(X) dX \quad (3.14)$$

donde  $E[\ ]$  es el operador de expectativa,  $p_X(x)$  es la probabilidad de que  $X$  tenga un valor particular. La varianza (o segundo momento central), que cuantifica la extensión de una distribución esta dada por

$$\sigma_X^2 = E[(X - \mu_x)^2] = \int_{-\infty}^{\infty} (X - \mu_x)^2 p_X(X) dX \quad (3.15)$$

y la raíz cuadrada de la varianza es igual a la desviación estándar [Press89], de la distribución. Por extensión, podemos definir el  $n$ -ésimo momento central

$$v_n = E[(X - \mu_x)^n] = \int_{-\infty}^{\infty} (X - \mu_x)^n p_X(X) dX \quad (3.16)$$

El tercer momento de una distribución se conoce como sesgo,  $\zeta$ , y caracteriza el grado de asimetría sobre la media [Press89]. La inclinación de una variable aleatoria  $X$  viene dada por (3.12).

$$v_3 = \frac{E[(X - \mu_x)^3]}{\sigma^3} \quad (3.17)$$

Un sesgo positivo significa una distribución con una cola que se extiende hacia un valor más positivo y un sesgo negativo significa una distribución con una cola que se extiende hacia un punto más negativo [Clifford08], se observa en Figura 3.5.

El cuarto momento de una distribución se conoce como curtosis y mide el pico relativo de la planitud de una distribución con respecto a una distribución normal o gaussiana, como se observa en Figura 3.6. Se define:

$$v_4 = \kappa = \frac{E[(X - \mu_x)^4]}{\sigma^4} \quad (3.18)$$

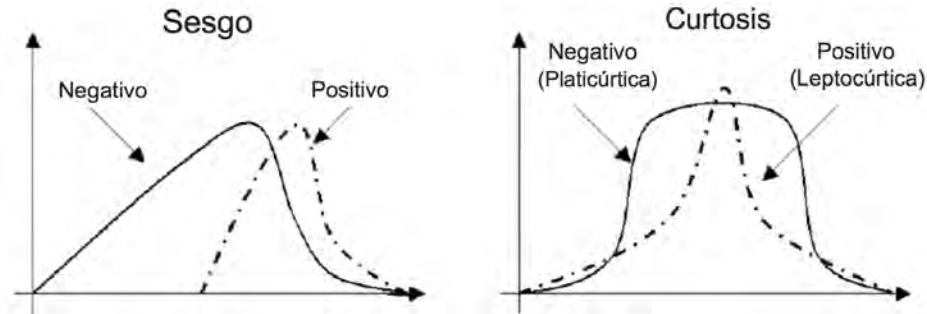


Figura 3.6: Distribuciones con tercer y cuarto momento, sesgo o asimetría (a) y curtosis (b) respectivamente, que son significativamente diferentes de la distribución normal o gaussiana [Clifford08].

Teniendo en cuenta que la curtosis de una distribución gaussiana es igual a 3, mientras que tercer momento o sesgo es igual a 0. Una distribución con una curtosis positiva, mayor a 3, se denomina leptocúrtica o supergaussiana [Clifford08]. Una distribución con una curtosis menor a 3, se denomina platycúrtica, o subgaussiana. Las distribuciones gaussianas se denominan mesocúrticas. También considerando que el sesgo o asimetría y la curtosis se normalizan dividiendo los momentos centrales por potencias apropiadas de  $\sigma$  para que sean adimensionales [Clifford08].

Sin embargo, estas definiciones son para funciones continuamente valoradas. En realidad, la función de distribución de probabilidad a menudo es difícil o imposible de calcular con precisión, por lo que debemos hacer aproximaciones empíricas de nuestras señales muestreadas. La definición estándar de la media de un vector  $\mathbf{X}$  con  $M$  valores

( $\mathbf{X} = [X_1, X_2, \dots, X_M]$ ) es:

$$\hat{\mu}_X = \frac{1}{M} \sum_{i=1}^M X_i \quad (3.19)$$

La varianza de  $\mathbf{X}$  esta dada por (3.15), la asimetría o sesgo por (3.16), y la estimación empírica de la curtosis se define de manera similar por (3.17) [Clifford08].

$$\hat{\sigma}^2(\mathbf{X}) = \frac{1}{M} \sum_{i=1}^M (X_i - \hat{\mu}_X)^2 \quad (3.20)$$

$$\hat{\zeta}(\mathbf{X}) = \frac{1}{M} \sum_{i=1}^M \left[ \frac{X_i - \hat{\mu}_X}{\hat{\sigma}} \right]^3 \quad (3.21)$$

$$\hat{\kappa}(\mathbf{X}) = \frac{1}{M} \sum_{i=1}^M \left[ \frac{X_i - \hat{\mu}_X}{\hat{\sigma}} \right]^4 \quad (3.22)$$

Aunque la idea básica detrás del ICA es muy simple, la implementación real se puede formular desde muchas perspectivas:

- Algoritmo rápido de punto fijo [Hyvärinen97].
- Maximización de la transferencia de información [Bell95].
- Maximización de la negentropía [Girolami96].
- PCA no lineal [Karhunen94].

Todos los métodos anteriores usan métricas de separación, o funciones de costo, que son esencialmente equivalentes a la medición de la no gaussianidad de las fuentes estimadas. La implementación real puede involucrar ya sea una manipulación de los datos de salida,  $\mathbf{Y}$ , o una manipulación de la matriz de desmezcla,  $\mathbf{W}$ .

### Limitaciones del ICA

Mientras que ICA es una técnica poderosa con pocas suposiciones sobre la naturaleza de las observaciones y las fuentes subyacentes, debe tenerse en cuenta que ICA tiene algunas limitaciones intrínsecas [Clifford08].

### El problema de permutación y escalado

Recondando la ecuación (3.2), se puede insertar una matriz arbitraria  $D$  y su inversa  $D^{-1}$  de manera que la ecuación (3.2) se mantiene sin cambios [Clifford08].

$$\mathbf{X}^T = \mathbf{A} \mathbf{D} \mathbf{D}^{-1} \mathbf{Z}^T \quad (3.23)$$

La matriz de mezcla ahora es  $\mathbf{A} \mathbf{D}$  y las fuentes ahora son  $\mathbf{D}^{-1} \mathbf{Z}^T$  con un orden de columna diferente y una escala diferente. Como solo conocemos  $\mathbf{X}$ , solo podemos resolver conjuntamente la matriz de mezcla y las fuentes, y es posible un número infinito de pares igualmente válidos. Por lo tanto, las estimaciones de las fuentes pueden aparecer en un orden arbitrario y con una escala arbitraria, que no tiene relación con la amplitud o la energía en las fuentes subyacentes [Clifford08].

Otra forma de pensar sobre este problema es que derivamos la estimación de la matriz de desmezcla  $\mathbf{W} \approx \mathbf{A}^{-1}$  optimizando una función de costo entre las columnas de la estimación de las fuentes  $\mathbf{Z}^T$ . Esta función de costo mide la independencia de una manera que es independiente de la amplitud [Clifford08]. Recordando que la curtosis es una cantidad adimensional. Para mitigar este problema, algunos algoritmos ICA ordenan las fuentes en términos de curtosis y las escalan para que tengan una unidad de varianza.

Para preservar la amplitud original de la fuente, se puede invertir la transformación, retener solo una fuente única, y reconstruir cada fuente en el dominio de observación. Por lo tanto, se requiere un conocimiento preciso de ciertas características o propiedades de las fuentes subyacentes, como distinguir oscilaciones morfológicas en el dominio del tiempo o el valor exacto de la curtosis, para identificar una fuente particular en las columnas de  $\mathbf{Z}^T$  [Clifford08].

### Mezcla estacionaria

El ICA supone un modelo de mezclado estacionario lineal (la matriz de mezcla es un conjunto de constantes independientes de la estructura cambiante de los datos a lo largo del tiempo). Sin embargo, para muchas aplicaciones esto solo es cierto desde ciertos puntos de observación o por periodos de tiempo muy cortos [Clifford08]. Por ejemplo, considere el

caso de tener ruido en un electrocardiograma (ECG). A medida que el sujeto inhala, el cofre se expande y el diafragma baja. Esto hace que el corazón se caiga y gire ligeramente. Si consideramos que la matriz de mezcla  $\mathbf{A}$  tiene un componente estacionario  $\mathbf{A}^s$  y un componente no estacionario  $\mathbf{A}^{ns}$  tal que  $\mathbf{A} = \mathbf{A}^s + \mathbf{A}^{ns}$  entonces  $\mathbf{A}^{ns}$  es igual a alguna constante ( $\alpha$ ) veces una de las matrices de rotación como

$$\mathbf{A}^{ns}(\theta) = \alpha \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta)t & -\text{sen}(\theta)t \\ 0 & \text{sen}(\theta)t & \cos(\theta)t \end{bmatrix} \quad (3.24)$$

Donde  $\theta = 2\pi f_{resp}$  y  $f_{resp}$  es la frecuencia de la respiración. En este caso  $\theta$  es el ángulo entre las diferentes fuentes (las fuentes son señales eléctricas de las contracciones musculares y las de la actividad cardíaca), que a su vez es una función del tiempo. Esto solo es cierto para valores pequeños de  $\alpha$ , por lo tanto, un pequeño ángulo  $\theta$ , entre cada fuente. Esta es una razón importante para las diferencias en la efectividad de ICA para la separación de fuentes para diferentes configuraciones de derivaciones.

### La suposición de la independencia

Las fuentes (columnas de  $\mathbf{Z}^T$ ) mezcladas por  $\mathbf{A}$  se suponen estadísticamente independientes. Es decir, se generan a partir de algunos procesos subyacentes que no están relacionados [Clifford08]. En el problema de la fiesta de cócteles, esto es trivialmente obvio; cada hablante no está modulando sus palabras en función de otras palabras que se hablen al mismo tiempo. Sin embargo, en el caso de tener ruido en un ECG y aplicar ICA para la eliminación de artefactos, esto a veces no es cierto. Cuando un sujeto monitoreado de repente aumenta sus niveles de actividad, los artefactos de los movimientos musculares pueden manifestarse en el ECG. A veces, habrá cambios significativos en la frecuencia cardíaca o en la morfología del latido como resultado del cambio de actividad. El artefacto muscular y el cambio de morfología del latido ya no son independientes. Si la relación es lo suficientemente fuerte, ICA no podrá separar las fuentes [Clifford08].

### 3.3.5. Algoritmo rápido de punto fijo para el ICA

El algoritmo que se presenta en esta sección utiliza un esquema de iteración de punto fijo simple pero eficiente para encontrar los extremos locales de la curtosis de una combinación lineal de las variables observadas. Dado que es sabido [Hyvärinen97] que encontrar los extremos locales de la curtosis es equivalente a estimar los componentes independientes no gaussianos.

Es necesario que las observaciones  $\mathbf{x}$  pasen por un proceso de preblanqueado [Hyvärinen97]. El blanqueamiento o blanqueado es un paso de preprocesamiento común en el análisis estadístico, una transformación lineal que convierte a un vector aleatorio  $\mathbf{x}$  de dimensión  $n$ , con una matriz de covarianza definida positiva,  $\sigma_x^2 = \mathbf{\Sigma}$ , en un nuevo vector aleatorio de la misma dimensión  $n$  y con la matriz de covarianza  $\sigma_{\hat{x}}^2 = I$  [Kessy16]. Para obtener los datos para el algoritmo, las observaciones  $\mathbf{x}$  deben pasar por alguno de los métodos existentes para lograr el blanqueado de los datos, en este caso, se usara el PCA para ayudar a descorrelacionar los datos [Kessy16], y obtener el nuevo vector de observaciones  $\hat{\mathbf{x}}$  que se usara para el algoritmo rápido de punto fijo.

Tomando en cuenta la ecuación (3.10),  $\mathbf{Y}^T = \mathbf{W}\mathbf{X}^T$ , y considerando que  $\mathbf{X}^T$  pasa por un preblanqueado, se transforma en  $\hat{\mathbf{X}}^T$ , en donde debido al blanqueamiento la matriz  $\mathbf{W}^T$  es considerada para facilitar el proceso de transformación en una matriz  $\mathbf{B}$  ortogonal, proporcionando la siguiente ecuación (3.19).

$$\mathbf{Y}^T = \mathbf{B}\hat{\mathbf{X}}^T \quad (3.25)$$

#### Estimación de un componente independiente

Tomando al vector  $v$  como una columna de la matriz ortogonal  $\mathbf{B}$ , y recordando la ecuación del cuarto momento o curtosis (3.8), el algoritmo rápido de punto fijo se define [Hyvärinen97]:

El vector final  $\mathbf{v}$  es una columna de la matriz  $\mathbf{B}$ , esto significa que  $\mathbf{v}$  separa una de las señales fuentes independientes.

---

**Algoritmo 1** Algoritmo rápido de punto fijo
 

---

Entrada: Señales de entrada después de ser aplicado un preblanqueado  $\hat{\mathbf{X}}^T$  dado por el PCA.

Salida: Vector  $\mathbf{v}$  que corresponde a una columna de la matriz de  $\mathbf{B}$

1. Toma un vector aleatorio inicial  $\mathbf{v}_0$  de norma 1. Deja  $k = 1$ .
  2. Pon  $\mathbf{v}_k = E[\hat{\mathbf{X}}\mathbf{v}_{k-1}^T\hat{\mathbf{X}}] - 3\mathbf{v}_{k-1}$ .
  3. Divide  $\mathbf{v}_k$  entre su norma ( $|\mathbf{v}_k| = \sqrt{v_{k1}^2 + v_{k2}^2 + \dots + v_{kn}^2}$ ).
  4. Si  $\mathbf{v}_k\mathbf{v}_{k-1}$  no está lo suficientemente cerca de 1, avanzamos un valor,  $k = k + 1$ , y regresamos al paso 2. De lo contrario, se termina en  $\mathbf{v}(k)$ .
- 

**Estimación de varios componentes independientes**

Para estimar  $n$  componentes independientes, ejecutamos el algoritmo rápido de punto fijo para estimar un componente independiente  $n$  veces. Para garantizar que estimamos cada vez un componente independiente diferente, solo necesitamos agregar una proyección de ortogonalización dentro del ciclo [Hyvärinen97]. Recuerde que las columnas de la matriz de mezcla  $\mathbf{B}$  son ortonormales debido al blanqueamiento, dado en este caso por PCA como preprocesamiento [Kessy16]. Por lo tanto, podemos estimar los componentes independientes uno por uno proyectando la solución actual  $v_k$  en el espacio ortogonal a las columnas de la matriz de mezcla  $\mathbf{B}$  previamente encontrada. Defina la matriz  $\hat{\mathbf{B}}$  como una matriz cuyas columnas son las columnas previamente encontradas de  $\mathbf{B}$ . Luego agregue la operación de proyección al comienzo del paso 3:

3. Dejar  $\mathbf{v}_k = \mathbf{v}_k - \hat{\mathbf{B}}\hat{\mathbf{B}}^T\mathbf{v}_k$ . Dividir  $\mathbf{v}_k$  entre su norma.

Además, el vector aleatorio inicial debe proyectarse de esta manera antes de comenzar las iteraciones. Para evitar que errores de estimación en  $\hat{\mathbf{B}}$  deterioren la estimación  $\mathbf{v}_k$ , este paso de proyección puede omitirse después de las primeras iteraciones: una vez que la solución  $\mathbf{v}_k$  ha ingresado a la colección como uno de los puntos fijos, permanecerá allí y convergerá a ese punto fijo.



Se puede observar en Fig.3.7 el funcionamiento del algoritmo 1, se tienen tres señales de entrada, una senoidal, una diente de tiburón y ruidos de impulsos, se observan también las limitantes del ICA habladas en la sección anterior, como la permutación.

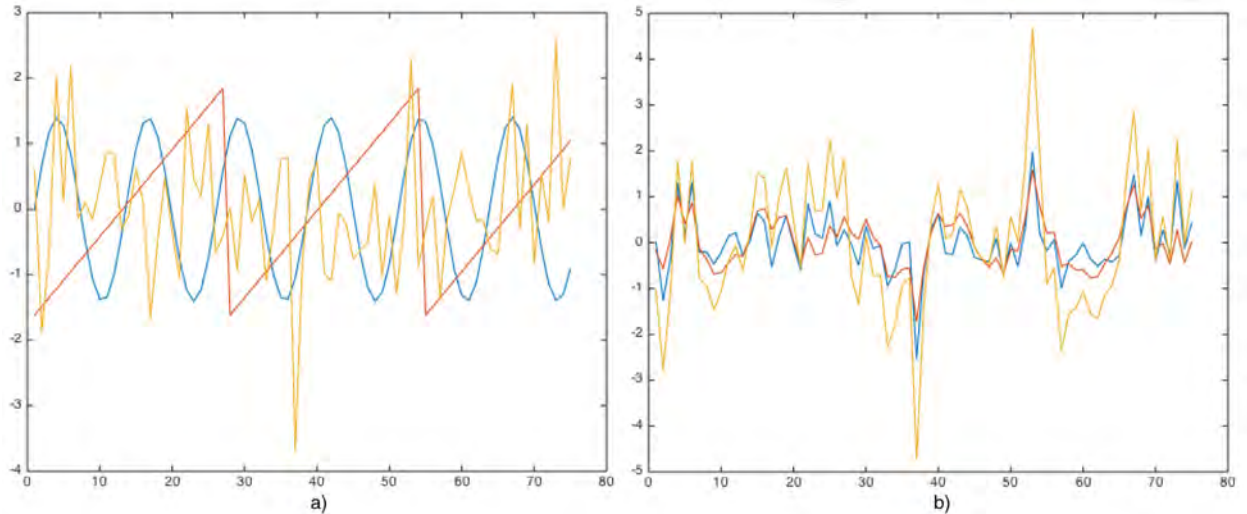


Figura 3.7: Algoritmo rápido de punto fijo para ICA con señales de entrada simuladas, a) Señales fuentes obtenidas, b) Señales de observación o mezcladas.

### 3.4. Conclusión sobre la extracción de características

Los métodos que se presentan en este capítulo proporcionan las características que aportaran información para la clasificación, para diferenciar a los individuos que tienen la EA. La entropía espectral proporcionara el nivel de complejidad de la señal y el coeficiente de correlación la sincronía que se tiene entre las señales EEG, el propósito del ICA es usarlo como preprocesamiento para mejorar las características de energía y entropía, para el caso del coeficiente de correlación no es de utilidad usar el ICA debido a que descorrelaciona los datos y esto altera de manera no deseada este valor del coeficiente.

Las características que se extraerán son 17, siendo las siguientes: energía para las 5 bandas de frecuencia (alfa, beta, delta, teta y gama), entropía espectral para las 5 bandas de frecuencia, y 7 valores de coeficiente de correlación de Pearson, se ven más a detalle en los capítulos posteriores.



## Capítulo 4

# Sistema de extracción de características para la detección de la enfermedad de Alzheimer (SECDEA)

El presente capítulo describe el sistema implementado para obtener, a partir de las señales EEG. Las características que proporcionan información acerca del estado de la persona, esto es, determinar si está es riesgo o no de padecer Alzheimer.

Se toma en cuenta en la elección de los métodos y algoritmos de este sistema aquellos dentro del marco teórico, siendo los considerados energía por bandas de frecuencia, entropía espectral y coeficiente de correlación de Pearson, gracias a esto los datos podrán comparados con investigaciones previas.

### 4.1. Técnica utilizada para extraer características

En esta sección se describe el método usado para la extracción de las características a las señales EEG, esta implementación programada en Python y validada contra los resultados en Matlab, usando las librerías de Matlab para comparar los resultados de los

métodos desarrollados en Python para de esta manera evitar errores de programación.

En Figura 4.1 se muestra el diagrama a bloques del proceso para extraer las características de las señales de EEG. Para comprender bien las partes que lo integran se analizará cada bloque por separado.

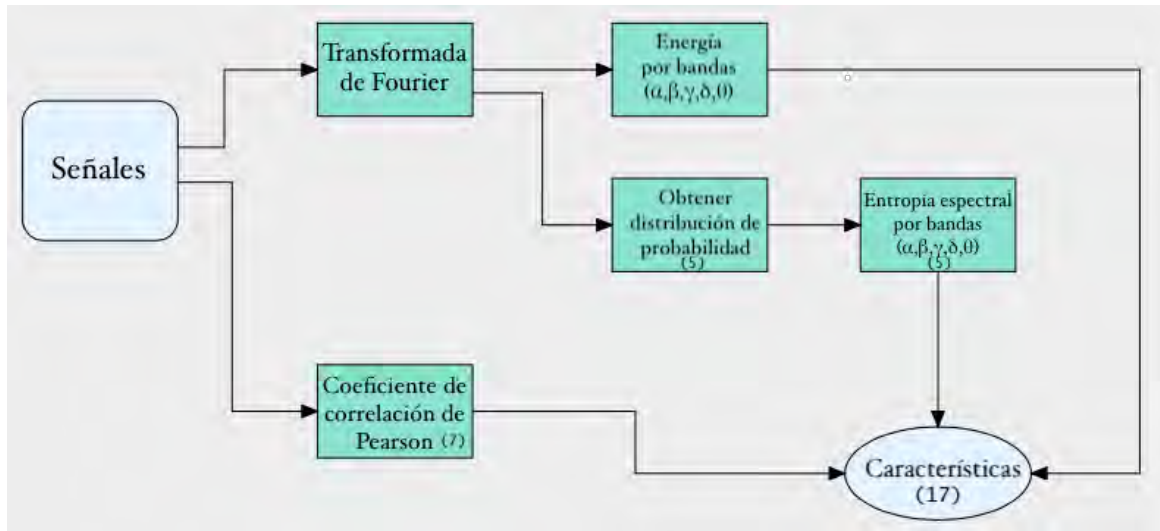


Figura 4.1: Diagrama de flujo de la implementación.

El bloque de señales corresponde a las 23 señales de los canales del EEG, para la energía por bandas y la entropía espectral por bandas se aplicó un filtro de caja sobre los componentes de frecuencia de la transformada de Fourier y las señales de las cuales se obtiene el coeficiente de correlación de Pearson están dadas una señal repetitiva por hemisferio cerebral.

#### 4.1.1. Señales

Se muestra en la Figura 4.2 cuatro de las señales EEG de un mismo sujeto, en el mismo tiempo de captura, se tomaron cuatro únicamente en lugar de todas las del registro del EEG para ejemplificar como son éstas y su comportamiento.

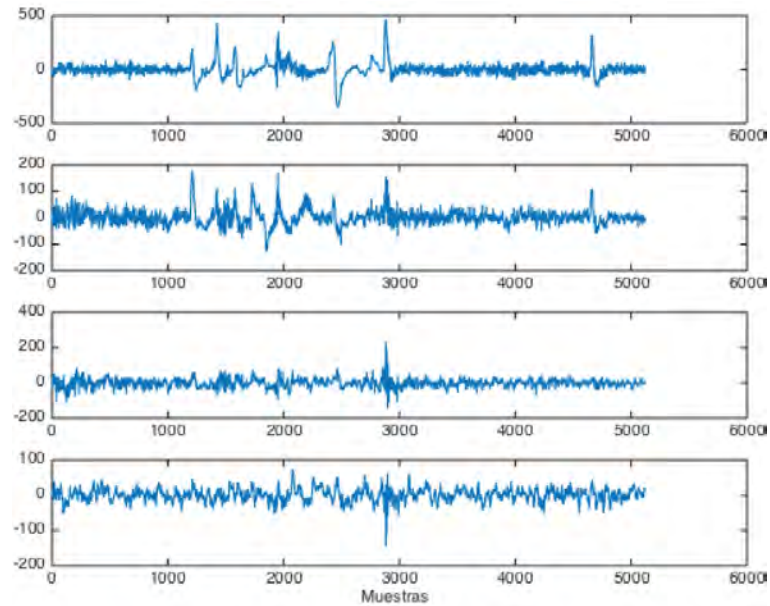


Figura 4.2: 4 señales de diferentes canales EEG tomadas de las 23 posibles para dar una representación gráfica.

Se observan en Figura 4.2 algunas características propias de las señales EEG, ejemplos de esto es que aun en distinto canal la frecuencia de la señal es similar, tienen algunos picos o impulsos en común y son series de tiempo que tienen la línea base en 0 volts.

#### 4.1.2. Transformada de Fourier y cálculo de energía

Posteriormente se realiza el cálculo de la transformada rápida de Fourier (FFT) sobre las señales EEG para conocer su contenido de frecuencia. Se observa en Figura 4.3 la transformada de Fourier correspondiente a las cuatro señales observadas en Figura 4.2. En Figura 4.3 el eje “x” está dado en valores de frecuencia en Hz, desde 0 hasta la mitad de la frecuencia de muestreo (256Hz), es correspondiente a tomar los valores de las muestras en la FFT de 0 a  $\frac{N}{2}$ .

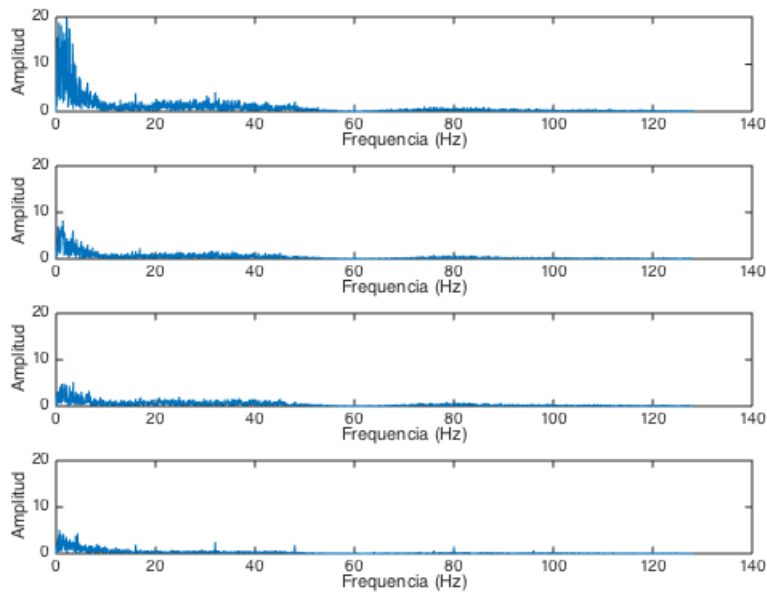


Figura 4.3: Transformada de Fourier en el dominio de la frecuencia.

En la Figura 4.3 se observa que la mayor contribución de los componentes de frecuencia están en las frecuencias más bajas.

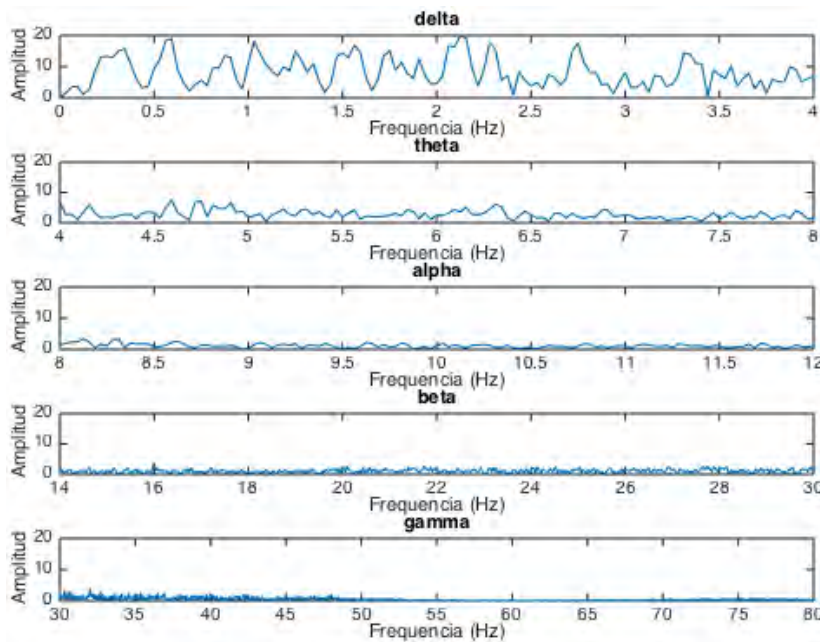


Figura 4.4: Bandas de frecuencia.

Conociendo los valores en frecuencia, se dividen en las bandas mencionadas en la Tabla 2.1, siendo estas: banda delta de menos de 4 Hz, banda teta de 4 a 8 Hz, la alfa de 8 a 12 Hz, la beta de 14 a 30 Hz y la gama de 30 a 80 Hz. Esto se aplica a la transformada de todas las señales, se muestra en Figura 4.4 esta separación para la primera señal mostrada en Figura 4.3 de manera representativa del proceso.

Con los valores por cada banda se calcula la energía en cada una de ellas. Las bandas de frecuencia se separan mediante un filtro de caja en frecuencia. Teniendo entonces 5 valores por cada señal uno por cada banda de frecuencia, tomando en cuenta que son 23 canales y no solo uno pero se desea dar solo la contribución por banda de frecuencia, se suman todos los valores de energía para todas las señales por cada banda para obtener un vector de 5 que represente la suma de los valores por banda, posteriormente a los valores obtenidos por la suma de todos los canales se realiza una normalización al vector para obtener la contribución de cada banda de frecuencia. Los valores de energía por banda son de 0 a 1 e indican la contribución que da cada banda a la cantidad de valores de frecuencia.

### 4.1.3. Distribución de probabilidad y entropía espectral

Para realizar el cálculo de la entropía espectral se deben obtener las distribuciones de probabilidad, debido a que son necesarias para la fórmula de Shannon, una por cada banda de frecuencia en la cual todas las señales por banda constituyen esta distribución de probabilidad. Para esto primero se obtienen los histogramas de cada banda de frecuencia como se puede observar en Figura 4.5.

Se puede ver que los valores del histograma dan una distribución de probabilidad discreta. Una vez que se conoce cada distribución, se calcula la entropía espectral (3.1) en cada una, tomando en cuenta que en (3.1) se da el valor de la entropía espectral entre la máxima entropía, por lo tanto, los valores estarán acotados de 0 a 1.

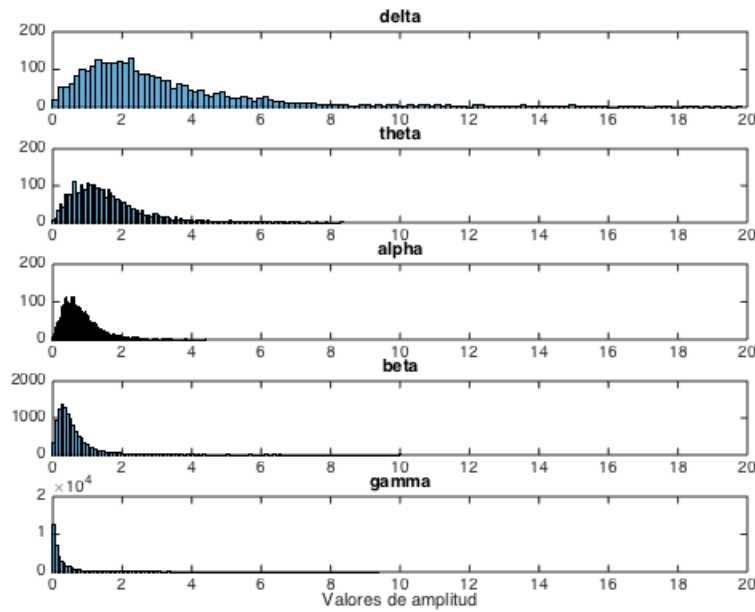


Figura 4.5: Histogramas de los valores por cada banda de frecuencia de todas las señales.

#### 4.1.4. Coeficiente de correlación de Pearson

El coeficiente de correlación de Pearson se calcula directamente con las señales, se realizan las comparaciones sugeridas por [Al-Jumeily15], las cuales son: temporal derecho con izquierdo (TD-TI), temporal derecho con frontal (TD-F), temporal derecho con central (TD-C), temporal derecho con occipital (TD-O), temporal izquierdo con frontal (TI-F), temporal izquierdo con central (TI-C), y temporal izquierdo con occipital (TI-O). Tomando una señal EEG representativa de cada zona.

## 4.2. Análisis de componentes independientes

En la implementación del ICA se obtienen los valores de energía y entropía por banda de frecuencia para poder compararlos con los valores de las mismas características sin el ICA. En el caso del coeficiente de correlación de Pearson no se plantea el uso del ICA debido a que este afecta a los valores de correlación por el preprocesamiento de blanqueado que se lleva a cabo en el ICA el cual descorrelaciona los datos.



Como se mencionó en el capítulo 3, en el ICA a las señales se les aplica un escalamiento y permutación por la naturaleza del método, tal como se puede observar en Figura 4.6, donde se muestra en la columna izquierda (a), las señales originales que se introducen al ICA, y en la columna derecha (b), las señales después del ICA.

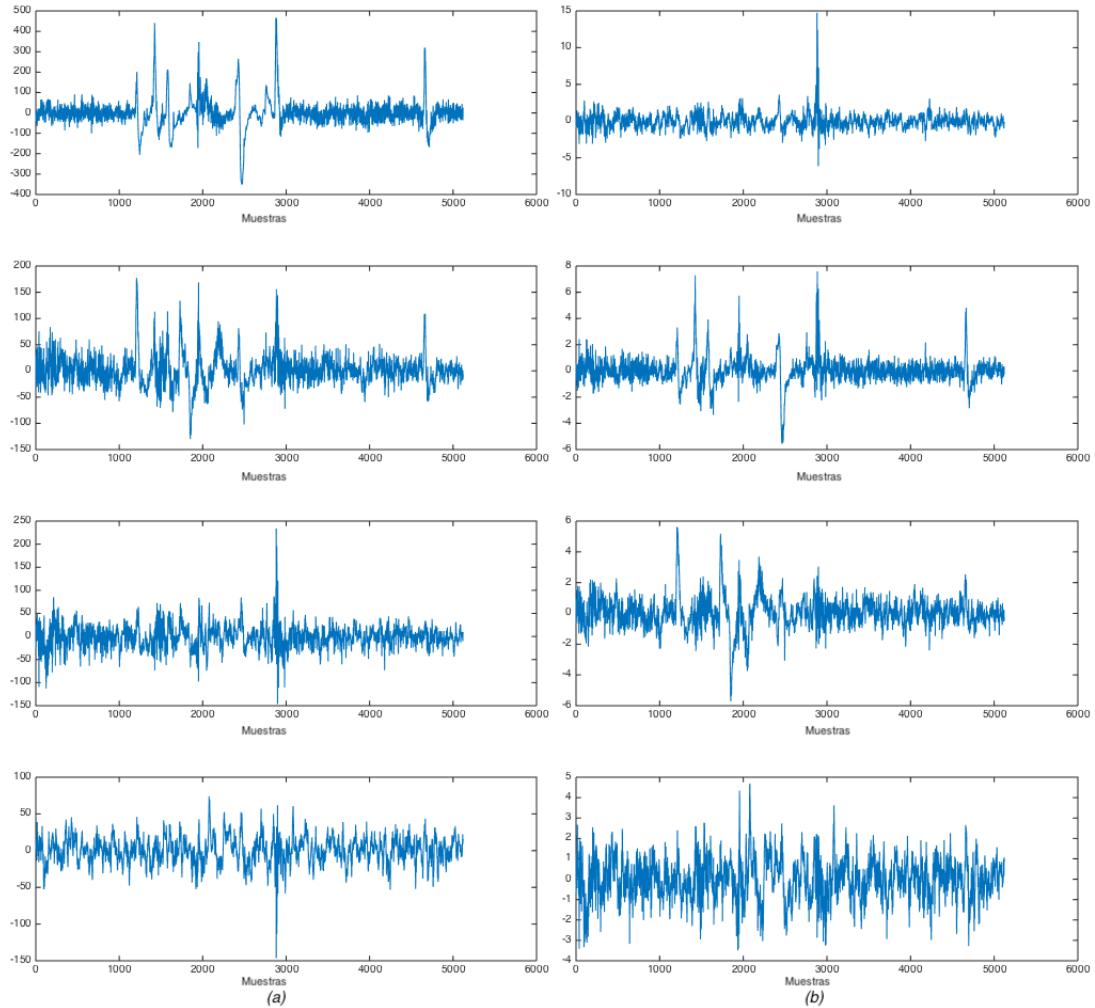


Figura 4.6: Señales antes (a), y después (b) del ICA, se puede observar su permutación, escalamiento y separación de componentes de una señal para agregarla a otra.

Se observa en Figura 4.6 que las señales tienen cambios en su forma pero se conservan algunos puntos característicos, como picos de amplitud y frecuencias, esto debido a la descomposición que realiza el ICA, envía información que estaba en una señal a otra con la suposición de que corresponde a ella. En la Figura 4.6 se observa que la primera señal

de la columna (b) corresponde con la tercera de la columna (a), la segunda señal (b) con la primera de la (a) y las dos restantes de (b), parecen una combinación de elementos de las señales (a) todo esto debido a la permutación que realiza el algoritmo de punto fijo del ICA. A estas señales se les aplica la transformada de Fourier para obtener sus valores en frecuencia, en Figura la 4.7 se muestran los gráficos de las transformadas de Fourier.

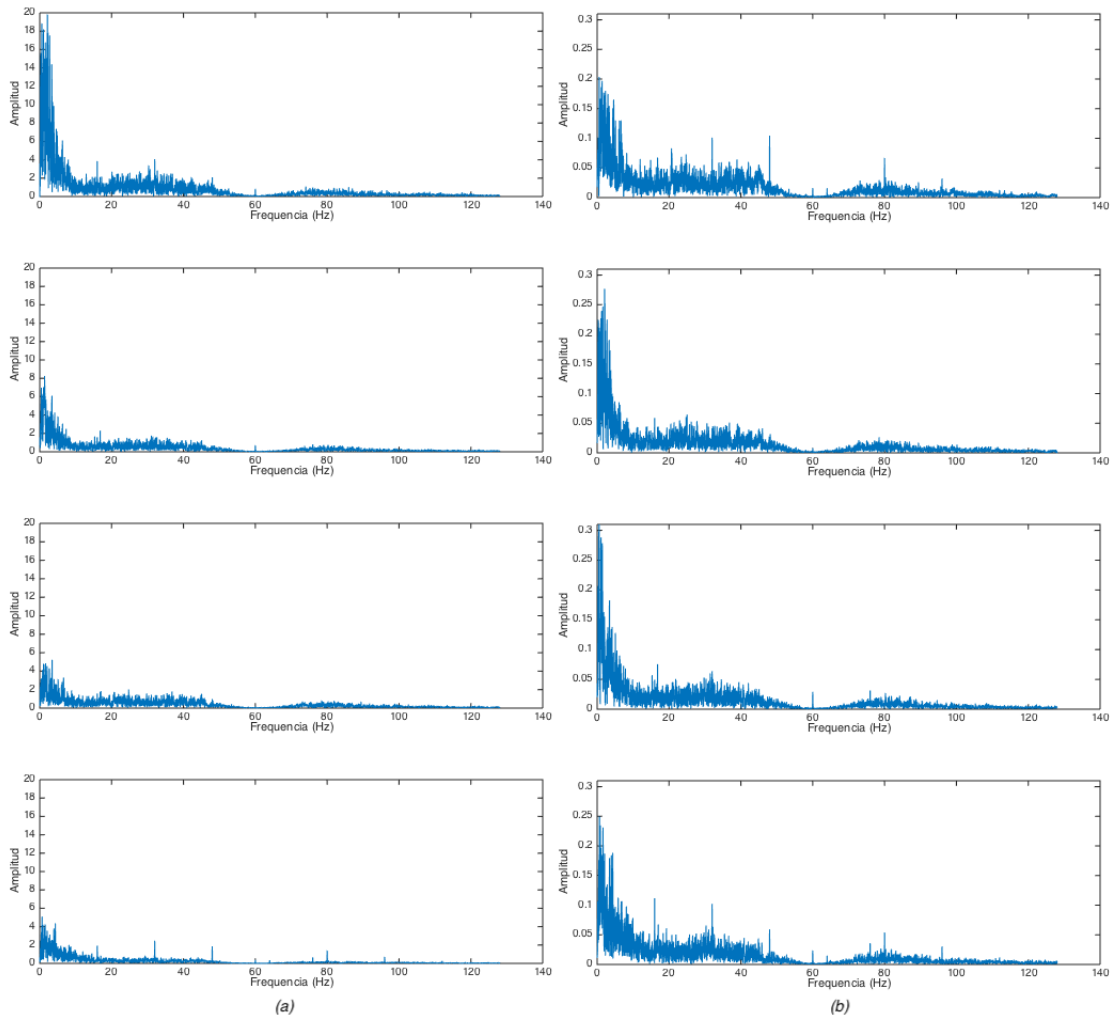


Figura 4.7: Transformadas de Fourier antes (a), y después (b) del ICA.

El siguiente paso es separar por bandas de frecuencia para hacer el cálculo de la energía, posteriormente, obtener sus distribuciones de probabilidad para obtener la entropía espectral. Tomando en cuenta que en Figura 4.7 la primera gráfica de la columna izquierda

(a), y la segunda de la columna derecha (b), representan a la misma señal antes y después del ICA, tomamos estos valores de frecuencia para ejemplificar el proceso que se realiza, en figura 4.8 se muestran los valores en las bandas de frecuencia antes y después del ICA.

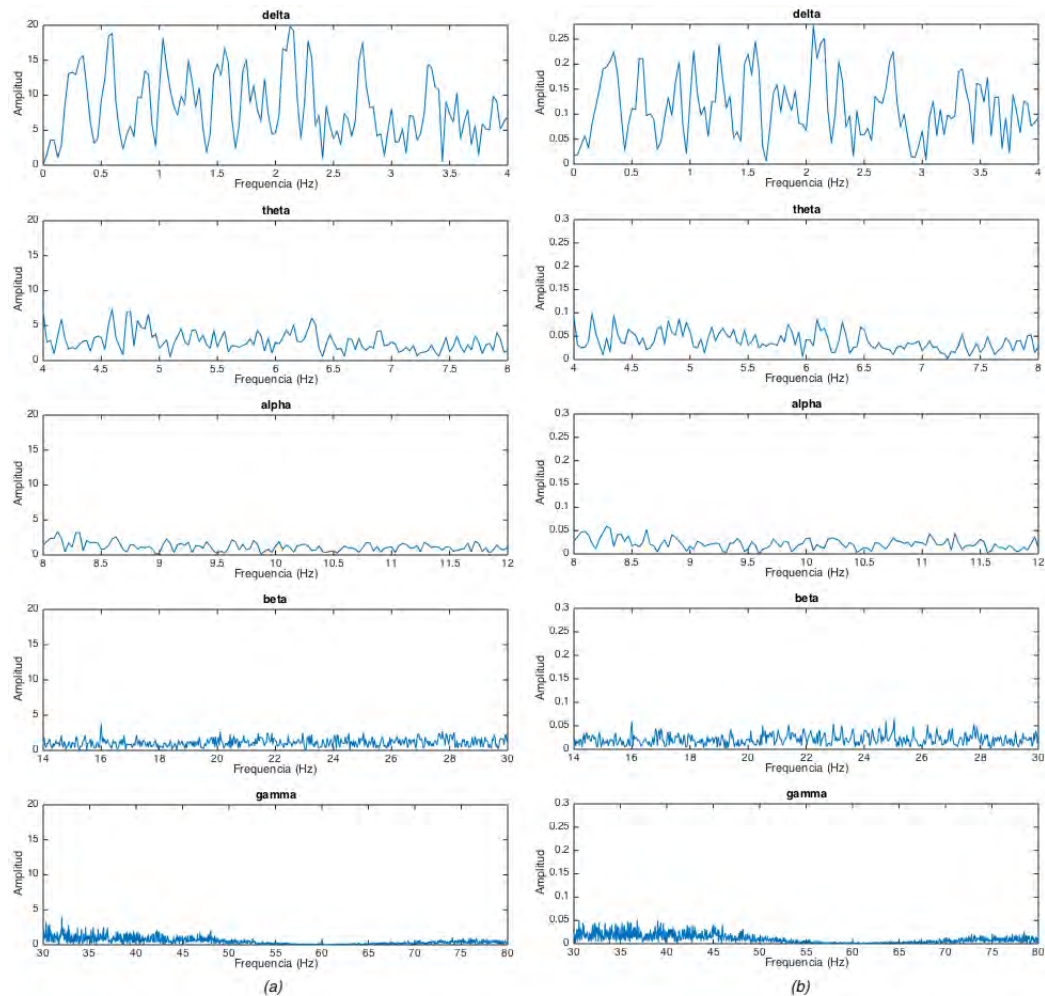


Figura 4.8: Bandas de frecuencia antes (a), y después (b) del ICA.

Se observa en Figura 4.8 que debido al escalamiento producido por el ICA, también se ven escalados los valores en las bandas de frecuencia, sin embargo, son muy similares en forma antes y después del ICA. Posteriormente se procede a calcular los histogramas para obtener las distribuciones de probabilidad de las bandas de frecuencia, se muestra en Figura 4.9 estos histogramas.

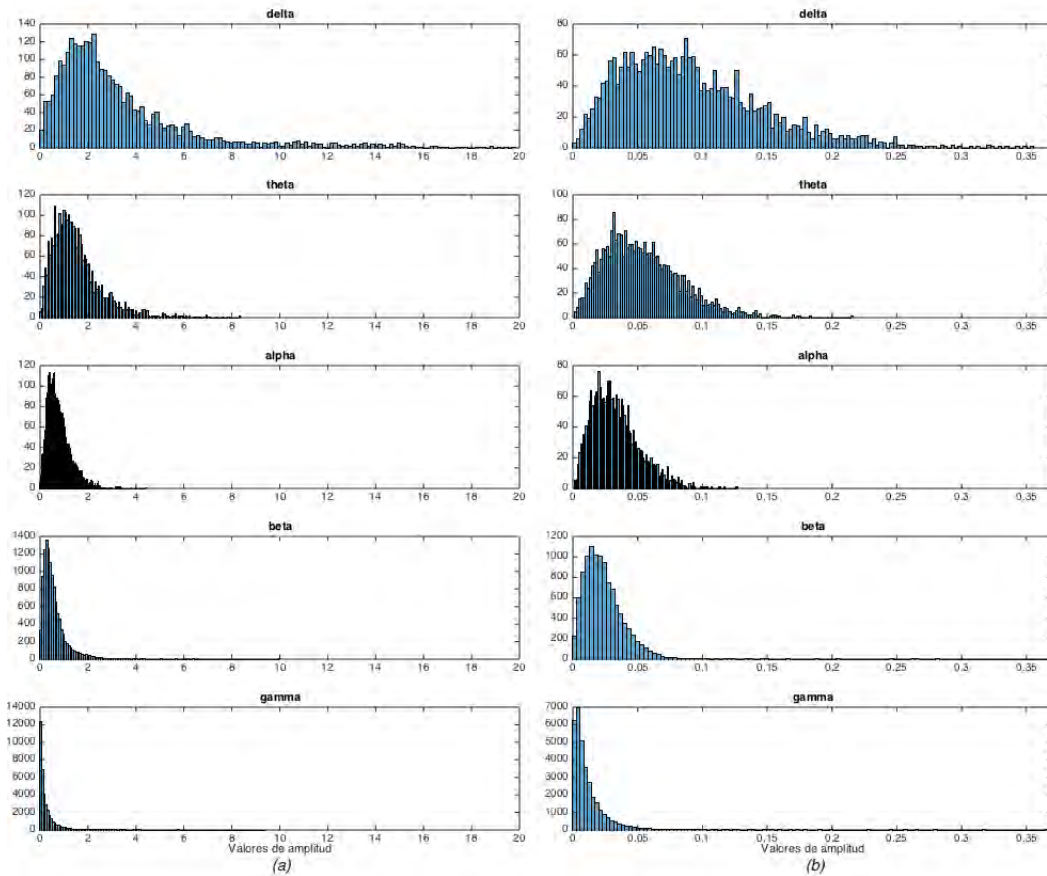


Figura 4.9: Histogramas de las bandas de frecuencia antes (a), y después (b) del ICA, la forma del histograma se conserva aun que los valores de amplitud se ven distintos en el eje por el escalamiento.

### 4.3. Conclusión sobre el SECDEA

Se puede observar en Figura 4.1 las características con las cuales se hará el análisis sin ICA, siendo estas las siguientes:

- Energía: 5 valores de energía tomados por bandas de los coeficientes de Fourier.
- Entropía: 5 valores de entropía espectral tomados de las distribuciones por banda de frecuencia.
- Coeficiente de correlación de Pearson: 7 valores de correlación tomados entre canales

representantes de región o lóbulo del cerebro.

Las características que se obtienen son:

- Energía: 5 valores de energía tomados por bandas de los coeficientes de Fourier después del ICA.
- Entropía: 5 valores de entropía espectral tomados de las distribuciones por banda de frecuencia después del ICA.

Estas características fueron escogidas para observar las 3 distintas anomalías que se presentan en las señales EEG para los pacientes con la EA que se mencionan en [Elgandelwar16], la energía para observar la desaceleración del EEG, la entropía espectral para poder observar la complejidad de las señales, y el coeficiente de correlación de Pearson para realizar una medición de sincronía.

Los valores obtenidos aplicando el ICA serán comparados en la sección de resultados con los valores sin el ICA para tomar los que mejor información aporten para realizar la clasificación de los sujetos con EA.

En el siguiente capítulo, analizando los resultados se elegirán las características que aporten mejor información de las 17 totales para formar el vector de características usado en la clasificación.



## Capítulo 5

# Resultados

El presente capítulo muestra los valores obtenidos de energía, entropía y coeficiente de correlación para las señales EEG de 12 personas, así como su respectivo análisis e interpretación.

### 5.1. Base de datos

La base de datos usada en la presente tesis consiste en 10 pacientes considerados sanos, y 2 pacientes considerados con EA. Las primeras 10 señales son del “CHB-MIT Scalp EEG database” usada en [Shoeb09], estas señales están muestreadas a 256Hz con una resolución de 16 bits, 23 canales, y manejan el sistema 10-20 de la colocación de los electrodos con una configuración de registro bipolar o diferencial.

Para el caso de los dos pacientes con la EA, las señales fueron proporcionadas por la doctora en neurociencia María Esther Olvera Corte, adscrita al Centro de Investigación Biomédica de Michoacán, las señales están muestreadas a 128Hz, tienen 16 canales y también el sistema 10-20 en configuración bipolar. Se pueden observar los canales para ambos tipos de pacientes mostrados en Tabla 5.1.

Debido a la privacidad que manejan los centros médicos o laboratorios especializados en la adquisición de las señales EEG de los pacientes con EA, fue de gran dificultad adquirir una base de datos más extensa, aun cuando se contactó a otros investigadores por medio de sus correos electrónicos o páginas de investigadores como researchgate.net, tam-

bién se buscó a la comunidad de médicos especialistas dentro de Michoacán, pero al no ser las señales EEG la herramienta con la que generalmente ellos trabajan para conocer si el paciente tiene la EA no contaban con una base de datos que pudieran compartir, y de la Alzheimer México I.A.P. nunca se recibió una respuesta. El manejo de pocas señales EEG y su distinta adquisición al ser un dispositivo de Hardware distinto el que usaron, restringe la base de datos, y por lo tanto, una mayor cantidad de resultados.

Canal	Registro bipolar (sanos)	Registro bipolar (enfermos)
1	FP1-F7	FP1-F7
2	F7-T7	F7-T3
3	T7-P7	T3-P3
4	P7-O1	FP1-F3
5	FP1-F3	F3-C3
6	F3-C3	C3-P3
7	C3-P3	P3-O1
8	P3-O1	FP2-F4
9	FP2-F4	F4-C4
10	F4-C4	C4-P4
11	C4-P4	P4-O2
12	P4-O2	FP2-F8
13	FP2-F8	F8-T4
14	F8-T8	T4-T6
15	T8-P8	T6-O2
16	P8-O2	FZ-CZ
17	FZ-CZ	
18	CZ-PZ	
19	P7-T7	
20	T7-FT9	
21	FT9-FT10	
22	FT10-T8	
23	T8-P8	

Tabla 5.1: Canales en la base de datos.



## 5.2. Energía

Los valores de la energía normalizada en las bandas de frecuencia sin ICA se muestran en Tabla 5.2.

Sujeto	Delta	Teta	Alfa	Beta	Gama
1	0.6699	0.1235	0.0298	0.0884	0.0885
2	0.7442	0.1568	0.0326	0.0343	0.0321
3	0.8709	0.0367	0.0508	0.026	0.0156
4	0.7649	0.0929	0.0278	0.0523	0.0622
5	0.7537	0.175	0.0136	0.0341	0.0237
6	0.5795	0.2743	0.0778	0.0514	0.0169
7	0.7662	0.0907	0.0299	0.0553	0.0579
8	0.882	0.0522	0.0148	0.0266	0.0245
9	0.1744	0.4647	0.1327	0.209	0.0193
10	0.8326	0.1191	0.0125	0.0155	0.0203
11 (Enfermo)	<b>0.9667</b>	0.0109	0.006	0.0105	<b>0.006</b>
12 (Enfermo)	<b>0.9299</b>	0.028	0.0104	0.0218	<b>0.0099</b>

Tabla 5.2: Energía en las bandas de frecuencia sin ICA.

Recordando que estos valores se obtuvieron de separar en bandas los coeficientes de la transformada de Fourier de las señales EEG.

Sujeto	Delta	Teta	Alfa	Beta	Gama
1	0.471	0.1572	0.0593	0.1568	0.1557
2	0.5785	0.1926	0.0541	0.0803	0.0944
3	0.5308	0.0973	0.1495	0.1283	0.0941
4	0.4448	0.0999	0.0826	0.1674	0.2052
5	0.6315	0.2105	0.0291	0.0648	0.0641
6	0.5377	0.2268	0.0761	0.0972	0.0622
7	0.3952	0.1347	0.0976	0.1831	0.1894
8	0.749	0.0693	0.0259	0.0699	0.0859
9	0.1811	0.3129	0.1335	0.2948	0.0776
10	0.7117	0.1377	0.0267	0.0297	0.0393
11 (Enfermo)	0.7363	<b>0.0281</b>	0.022	0.0933	0.1202
12 (Enfermo)	0.573	<b>0.0439</b>	0.034	0.15	0.1991

Tabla 5.3: Energía en las bandas de frecuencia con ICA.

En Tabla 5.2 se observa que la contribución en la banda delta para en los pacientes enfermos es más grande que en la de los sanos, de igual manera, se observa que la banda gama es considerablemente mayor los valores para los pacientes sanos a comparación de los enfermos, esto muestra que las señales EEG de los enfermos si tienen una desaceleración a diferencia de los sanos.

Los valores de energía cambian si a las señales se les aplica el ICA, los valores obtenidos para este caso se muestran en Tabla 5.3, se puede observar que la contribución por banda cambia y no se puede apreciar de igual manera el fenómeno de la desaceleración en las señales EEG, sin embargo, se rescata que los valores en la banda teta aportan diferencias entre los sujetos sanos, y los pacientes con la EA.

### 5.3. Entropía espectral

Se hizo el cálculo de la entropía a cada una de las bandas de frecuencia por sujeto con y sin ICA, los resultados que se muestran en Tabla 5.4 los del procedimiento sin la implementación del ICA.

Sujeto	Delta	Teta	Alfa	Beta	Gama
1	0.917	0.9145	0.9025	0.8705	0.8321
2	0.9311	0.939	0.9081	0.9104	0.8297
3	0.9078	0.902	0.8638	0.8764	0.7612
4	0.9083	0.8827	0.9168	0.9093	0.7835
5	0.9165	0.8938	0.9055	0.9046	0.8328
6	0.9304	0.9316	0.9381	0.9049	0.8057
7	0.9046	0.8996	0.9398	0.9038	0.8057
8	0.9418	0.9289	0.9478	0.912	0.817
9	0.9465	0.8824	0.933	0.8963	0.8401
10	0.9492	0.9148	0.9251	0.9308	0.895
11 (Enfermo)	<b>0.7947</b>	0.9261	0.9268	0.8562	0.8365
12 (Enfermo)	<b>0.8623</b>	0.9049	0.9245	0.9064	0.8614

Tabla 5.4: Entropía en las bandas de frecuencia sin ICA.

En Tabla 5.4 se observa que los valores en la banda de frecuencia delta presenta considerablemente más cambios respecto a los sujetos sanos con los pacientes enfermos, los

valores de entropía van relacionados a la complejidad de la señal, segunda característica considerada para observar las diferencias entre pacientes con la EA en las señales EEG dicha en [Elgandelwar16].

En Tabla 5.5 se muestran los valores de la entropía espectral al ser aplicado el ICA en las señales EEG. Los valores no cambian demasiado, y se puede observar que sin ICA o con ICA la banda que diferencia a los sanos de los enfermos sigue siendo la banda delta.

Sujeto	Delta	Teta	Alfa	Beta	Gama
1	0.9135	0.9132	0.9029	0.7503	0.7088
2	0.9349	0.8831	0.897	0.8553	0.7179
3	0.9177	0.9169	0.8389	0.8347	0.7202
4	0.8453	0.9147	0.8467	0.8916	0.7441
5	0.9309	0.8917	0.8944	0.8681	0.7287
6	0.9174	0.9223	0.8969	0.8642	0.7522
7	0.8842	0.913	0.8788	0.8773	0.6852
8	0.9133	0.9025	0.8996	0.8577	0.7256
9	0.8923	0.8858	0.9264	0.8685	0.7742
10	0.8957	0.897	0.8887	0.8186	0.7412
11 (Enfermo)	<b>0.7322</b>	0.8427	0.8287	0.8214	0.7637
12 (Enfermo)	<b>0.7507</b>	0.9061	0.8622	0.8657	0.7982

Tabla 5.5: Entropía en las bandas de frecuencia con ICA.

## 5.4. Coeficiente de correlación de Pearson

Para el cálculo del coeficiente de correlación se realizaron 7 comparaciones de los lóbulos cerebrales tal como se mencionó en el capítulo anterior, los resultados obtenidos se muestran en Tabla 5.6.

En los valores obtenidos del coeficiente de correlación, a diferencia con los valores de energía y de entropía, no se obtuvieron resultados destacados para diferenciar entre ambos tipos de pacientes, la teoría afirma [Elgandelwar16], que la correlación entre el cerebro es mejor en los pacientes sanos.

Sin embargo no se vio reflejada como se observa en Tabla 5.6, dos causas posibles para que los valores no correspondieran es la forma de adquirir las señales, siendo de registro

Sujeto	TI-TD	TI-F	TI-C	TI-O	TD-F	TD-C	TD-O
1	0.1317	0.4834	0.5245	0.3015	0.0468	0.5859	0.1279
2	0.0732	0.337	0.5151	0.0468	0.282	0.4258	0.2312
3	0.3161	0.5553	0.7077	0.3782	0.5779	0.4416	0.1579
4	0.2831	0.6444	0.5996	0.1973	0.2463	0.6517	0.3174
5	0.0326	0.1004	0.5712	0.0597	0.0519	0.5666	0.1939
6	0.3679	0.1589	0.6315	0.2593	0.2986	0.4672	0.0333
7	0.363	0.2362	0.3575	0.2622	0.4704	0.5522	0.4045
8	0.0609	0.3417	0.4205	0.3084	0.3303	0.1766	0.7567
9	0.2828	0.3013	0.5581	0.0418	0.127	0.544	0.1424
10	0.3407	0.303	0.656	0.1268	0.0284	0.6411	0.4824
11 (Enfermo)	0.3574	0.886	0.1149	0.3927	0.1053	0.3238	0.9951
12 (Enfermo)	0.0308	0.9524	0.9077	0.0359	0.9603	0.8326	0.9851

Tabla 5.6: Coeficiente de correlación de Pearson.

bipolar y no referenciada, y la naturaleza de las señales adquiridas, siendo otras personas las que se encargaron de la grabación.

## 5.5. Diagnóstico de la EA

Al observar los resultados anteriores en el presente capítulo, 3 de las características sin el ICA son las que mejor diferencian a los sujetos sanos con los pacientes que tienen la EA, estas características son: energía en la banda de frecuencia delta y gama, y entropía en la banda de frecuencia delta, y forman el vector de características. En Figura 5.1 se muestran los valores de estas tres características de los sujetos que conforman al conjunto de entrenamiento, también se le realizó la extracción de características a un sujeto más de la base de datos usada en [Shoeb09], llamado en Figura 5.1 como prueba 1, y a otra muestra que se le realizó al sujeto 11 en otro momento, que fue etiquetado por la Dr. María Esther Olvera Corte como enfermo y que en Figura 5.1 se representa por prueba 2.

Implementando un clasificador de distancia mínima euclidiana entre puntos, tomando el centroide de los sujetos sanos (rojo en Figura 5.2) y la media de los pacientes con EA (azul en Figura 5.2),

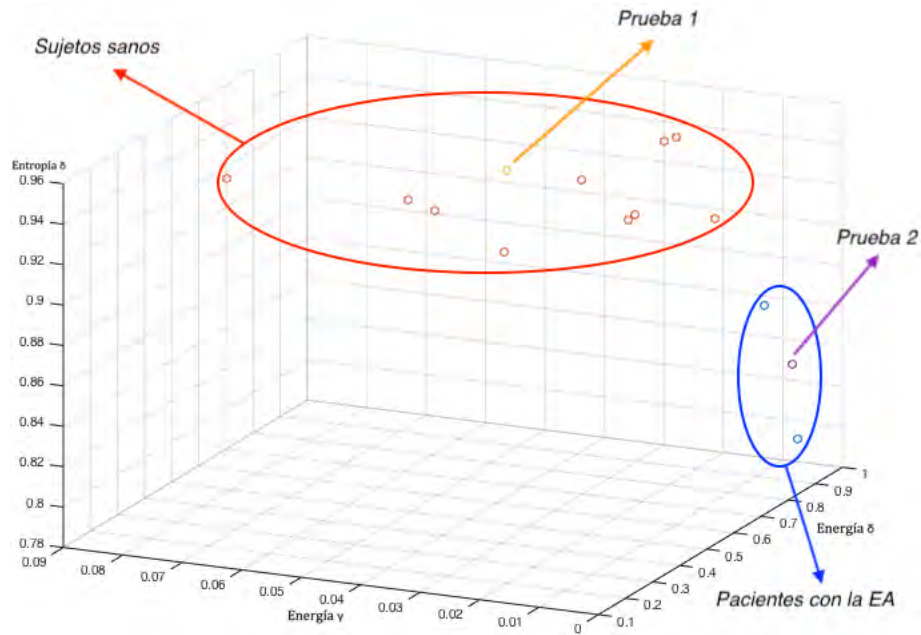


Figura 5.1: Gráfico de las características relevantes, cada sujeto corresponde a un punto, también los sujetos para la prueba 1 y 2.

Como los puntos para calcular la distancia al nuevo punto, siendo el caso de prueba 1 y prueba 2, y con el que tenga la menor distancia se considera de ese grupo (sujetos sanos o pacientes con EA) la distancia euclidiana está definida por:

$$d_E(P, Q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \dots + (p_n - q_n)^2} \quad (5.1)$$

En Figura 5.2 se muestran los centroides de las características de los pacientes sanos y de los enfermos con EA, al nuevo sujeto sano que se le extrajeron las características (prueba 1), y el otro registro del paciente enfermo (prueba 2). Se observa que la distancia menor en la prueba 1 es con el punto que representa a los sujetos sanos, la distancia menor entre los puntos con la prueba 2, es con el grupo de los enfermos de EA, lo cual sugiere que el clasificador está funcionando de manera correcta, recordando que lo ideal para construir un clasificador es tener un conjunto de entrenamiento grande, y no uno pequeño como es este caso, pero que debido a los problemas que se tuvieron no pudo ser más extenso.

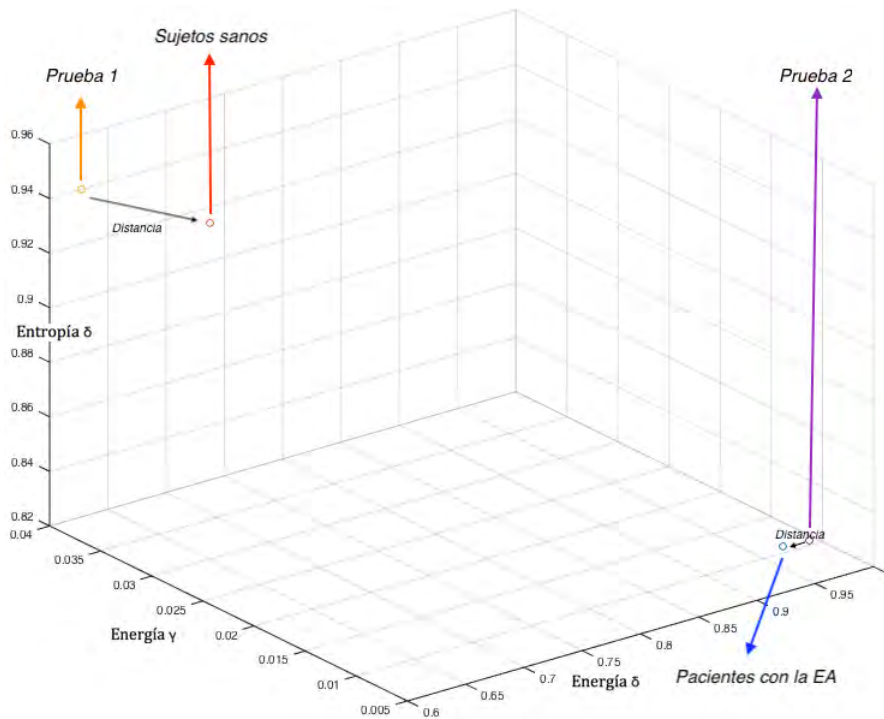


Figura 5.2: Gráfico de las distancias entre medias y sujetos de prueba.

## 5.6. Conclusión sobre resultados

Los resultados muestran características más relevantes en los valores de energía y entropía espectral, particularmente, la banda de frecuencia delta y gama en energía sin haber implementado el ICA, la banda teta con ICA, y la banda delta en entropía con o sin ICA.

Los resultados de las señales usando el método de ICA no obtuvieron una contribución relevante, en la energía sin ICA existen dos bandas de frecuencias que muestran diferencia, cuando en el ICA solo una, y en la entropía espectral sigue siendo una sola banda la que lo muestra, para la base de datos usada se demuestra con los resultados que la implementación del ICA no contribuye y por lo tanto fue preferente usar para la clasificación a los resultados sin ICA, debido a que aporta una característica más.

Con un conjunto más extenso de entrenamiento, se podrían probar clasificadores más complejos y los resultados serían más confiables.

## Capítulo 6

# Conclusiones y trabajos futuros

La investigación, desarrollo y resultados presentados en el presente trabajo han sido realizados con fines educativos, a pesar del uso de señales por parte de instituciones como el MIT y el centro de investigación Biomédica de Michoacán no se tiene ningún tipo de afiliación o asociación con estas instituciones.

El diagnóstico de la EA usando señales EEG, en términos generales, necesita más desarrollo para obtener resultados tan validados como por otros métodos, el más usado de ellos las imágenes de resonancia magnética, y las tomografías computarizadas. El algoritmo descrito en la presente, se basa en observar la desaceleración en las frecuencias de la EEG, la complejidad, y la sincronía, los métodos usados para estos propósitos, se tomaron en consideración por el estado del arte y las investigaciones previas de otros autores.

El mayor problema que se tuvo fue obtener la base de datos, ya que las señales EEG con la EA son confidenciales en la mayoría de los casos, esto por la privacidad que existen en los hospitales o laboratorios donde se realizan, al ser un método “alternativo” de diagnóstico para la EA, los especialistas con los que se pudo tener contacto no tenían las señales grabadas de EEG de los pacientes, o no quisieron compartirlas por cuestión de privacidad. Se estableció contacto con investigadores de otras universidad, las cuales su respuesta fue que tenían acuerdos firmados y no podían compartir su base de datos.

Sin embargo, se obtuvieron una limitada cantidad de señales EEG, la cual limito la extensión de los resultados como se hubieran deseado, y muy posiblemente, también la

calidad del funcionamiento del método de ICA.

Los resultados de energía y entropía espectral fueron consistentes con el estado del arte cuando no se implementaba el ICA, y a pesar de las limitaciones, se observa en la metodología implementada 3 características las cuales logran separar las señales en dos grupos.

### 6.1. Conclusiones particulares

- La energía sin el ICA en las bandas de frecuencia delta y gama ofrecen buenos resultados para diferenciar los dos grupos de pacientes.
- Los valores en la entropía espectral en la banda de frecuencia delta mostraron una diferencia entre ambos grupos.
- La calidad de la señal EEG depende de su adquisición, provocando una diferencia considerable si no se hace de manera adecuada o los parametros entre las señales a comparar son muy diferentes.
- El método de ICA no pudo comprobar dar una mejora a la obtención de las características debido posiblemente a la limitación en la base de datos.
- El coeficiente de correlación de Pearson no ofreció los resultados esperados.

### 6.2. Trabajos futuros

- Ampliar la base de datos y compartirla en redes de investigadores para solucionar el problema de no contar con los recursos necesarios para este tipo de investigaciones.
- Demostrar que el ICA realmente mejora a la obtención de características de las señales EEG.
- Aplicar distintos tipos de métodos para el análisis de sincronía en las regiones cerebrales para obtener mejores resultados de la sincronización que con coeficiente de correlación de Pearson.



# Referencias

- [Abdi07] Abdi, H. Singular value decomposition and generalized singular value decomposition. *Program in Cognition and Neurosciences, MS: Gr.4.1, Neil Salkind (Ed.) (2007). Encyclopedia of Measurement and Statistics*, págs. 1483–1492, 2007.
- [Abásolo06] Abásolo, D. Analysis of eeg background activity in alzheimer’s disease patients with lempel?ziv complexity and central tendency measure. *Medical Engineering and Physics 28*, págs. 315 – 322, 2006.
- [Al-Jumeily15] Al-Jumeily, D. A novel method of early diagnosis of alzheimer’s disease based on eeg signals. *The Scientific World Journal Volume 2015*, 2015.
- [AlzheimerMexOrg17] AlzheimerMexOrg. <http://www.alzheimermexico.org.mx>, 2017.
- [Argüelles09] Argüelles, R. Basic techniques of electroencephalography: principles and clinical applications. *Anales del Sistema Sanitario de Navarra, 32*, págs. 69–82, 2009.
- [Bell95] Bell, A. An information-maximization approach to blind separation and blind deconvolution. *Neural Computation, 7(6)*, págs. 1129–1159, 1995.
- [Benbadis17] Benbadis, S. R. Encephalopathic eeg patterns. *Medscape*, 2017.
- [Bock05] Bock, T. D. Early detection of alzheimer’s disease using nonlinear

- analysis of eeg via tsallis entropy. *International Conference, Warsaw, Poland*, págs. 11 – 15, 2005.
- [Boutros02] Boutros, N. Standard electroencephalography in clinical psychiatry: a practical handbook. *Oxford, UK*, págs. 1–184, 2002.
- [Clifford07] Clifford, G. Blind source separation: Pca and ica. Course materials for HST.582J/6.555J/16.456J, Biomedical Signal and Image Processing, 2007.
- [Clifford08] Clifford, G. Chapter 15 - blind source separation: Principal and independent component analysis. Course materials for HST-582J/6.555J/16.456J, Biomedical Signal and Image Processing, 2008.
- [CONAPO14] CONAPO. Proyecciones de la población 2010-2050. <http://www.conapo.gob.mx>, 2014.
- [Costa02] Costa, M. Multiscale entropy analysis of physiologic time series. *Phys Rev Lett, Vol.89*, 2002.
- [Dai12] Dai, Z. Discriminative analysis of early alzheimer’s disease using multi-modal imaging and multi-level characterization with multi-classifier (m3). *Journal of Neuroimage, Vol.59, No.3*, págs. 2187 – 2195, 2012.
- [Dauwels11] Dauwels, J. Slowing and loss of complexity in alzheimer’s eeg: Two sides of the same coin. *International Journal of Alzheimer’s Disease*, 2011.
- [Delgado14] Delgado, M. A. *Manual CTO de Medicina y Cirugía, Neurología*. Manual CTO de Medicina y Cirugía. CTO Editorial, 1<sup>a</sup> ed<sup>ón</sup>., 2014. ISBN 978-84-15946-05-2.
- [Dony01] Dony, R. D. The transform and data compression handbook. *Ed. K. R. Rao and P.C. Yip.*, 2001.

- [EAAssociation17] EAAssociation. Atrofia cortical. <https://www.alz.org/dementia>, 2017.
- [Elgandelwar16] Elgandelwar, S. M. Analysis of eeg signals for diagnosis of alzheimer disease. *International Journal of Scientific and Engineering Research, Volume 7*, págs. 529 – 532, 2016.
- [Escudero08] Escudero, J. Assessment of classification improvement in patients with alzheimer’s disease based on magnetoencephalogram blind source separation. *Artificial Intelligence in Medicine 43*, págs. 75 – 85, 2008.
- [Falk12] Falk, T. H. Eeg amplitude modulation analysis for semi-automated diagnosis of alzheimer’s disease. *EURASIP Journal on Advances in Signal Processing*, 2012.
- [Fernández10] Fernández, E. A. Demencias e imagen: lo básico. *Radiología No. 52*, págs. 4 – 7, 2010.
- [Frisoni96] Frisoni, G. B. Linear measures of atrophy in mild alzheimer disease. *AJNR Am J Neuroradiol No. 17*, págs. 913 – 923, 1996.
- [Girolami96] Girolami, M. Negentropy and kurtosis as projection pursuit indices provide generalised ica algorithms. *Blind Signal Separation Workshop, volume 8*, 1996.
- [Golub83] Golub, G. Matrix computations. *Oxford: North Oxford Academic, 2nd edition.*, 1983.
- [Hall15] Hall, G. Pearson’s correlation coefficient. <http://www.hep.ph.ic.ac.uk/hallg/>, 2015.
- [Heckbert95] Heckbert, P. Fourier transforms and the fast fourier transform (fft) algorithm. *Notes 3, Computer Graphics 2*, 1995.

- [Henderson13] Henderson, G. Development and assessment of methods for detecting dementia using the human electroencephalogram. *IJERT vol3*, 2013.
- [Hornero06] Hornero, R. Entropy analysis of the eeg background activity in alzheimer's disease patients. *Physiol. Meas.* 27, págs. 241 – 256, 2006.
- [Hyvärinen97] Hyvärinen, A. A fixed-point algorithm for independent component analysis. *Neural Computation*, 9, págs. 1483–1492, 1997.
- [INEGI14] INEGI. Censos de población y vivienda, 1950 - 1970, 1990, 2000 y 2010. <http://www3.inegi.org.mx/sistemas>, 2014.
- [iWorxSystems13] iWorx Systems. The electroencephalogram (eeg). <https://www.iworx.com>, 2013.
- [Jolliffe88] Jolliffe, I. Principal component analysis. *New York: Springer-Verlag*, 1988.
- [Jutgla13] Jutgla, E. G. Diagnosis of early alzheimer's disease based on eeg source localization and a standardized realistic head model. *Ieee journal of biomedical and health informatics*, vol. 17, 2013.
- [Karhunen94] Karhunen, J. Representation and separation of signals using non-linear pca type learning. *Neural Networks 1994*;7, págs. 113–127, 1994.
- [Kessy16] Kessy, A. Optimal whitening and decorrelation. 2016.
- [Lambers13] Lambers, J. Gram-schmidt orthogonalization. Course materials for MAT 415/515, 2013.
- [Locatelli98] Locatelli, T. Eeg coherence in alzheimer's disease. *Electroencephalography and clinical Neurophysiology 106*, págs. 229 – 237, 1998.
- [McLaughlin10] McLaughlin, C. Estudios eeg. <http://www.rfcom.ca/clinic>, 2010.

- [Navarro09] Navarro, B. Electroencefalografía. *Instrumentación Biomédica. Madrid, España. Universidad de Alcalá*, págs. 2–24, 2009.
- [Oppenheim98] Oppenheim, A. V. *Señales y sistemas*. Prentice Hall, 2nd edición, 1998. ISBN 970-17-0116-X.
- [Orellana02] Orellana, I. Resonancia magnética anatómica y funcional en pacientes con diagnóstico de enfermedad de alzheimer. *Anales de Radiología*, págs. 529 – 533, 2002.
- [Phung14] Phung, D. Using shannon entropy as eeg signal feature for fast person identification. *European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*, págs. 413–418, 2014.
- [Podgorelec12] Podgorelec, V. Analyzing eeg signals with machine learning for diagnosing alzheimer’s disease. *ELEKTRONIKA IR ELEKTROTEHNIKA, ISSN 1392-1215, VOL. 18, NO. 8, 2012*.
- [Poil13] Poil, J. S. S. Integrative eeg biomarkers predict progression to alzheimer’s disease at the mci stage. *Frontiers in Aging Neuroscience, Vol. 5, 2013*.
- [Press89] Press, W. H. *NUMERICAL RECIPES IN C: THE ART OF SCIENTIFIC COMPUTING*. Cambridge University Press, 1989. ISBN 0-521-43108-5.
- [Proakis07] Proakis, J. G. *TRATAMIENTO DIGITAL DE SEÑALES*. PEARSON EDUCACIÓN S.A., Madrid, 4ta edición, 2007. ISBN 978-84-8322-347-5.
- [Puigt11] Puigt, M. A very short introduction to blind source separation. Course materials, 2011.
- [Robledo14] Robledo, G. Plan de acción de alzheimer y otras demencias. <http://www.geriatria.salud.gob.mx/>, 2014.

- [Rowan04] Rowan, J. Conceptos básicos sobre eeg con mini-atlas. *Elsevier, Madrid, España*, págs. 1–187, 2004.
- [Shoeb09] Shoeb, A. H. Application of machine learning to epileptic seizure onset detection and treatment, 2009.
- [Sleigh04] Sleigh, J. W. Cortical entropy changes with general anaesthesia: theory and experiment. *Physiol. Meas.* 25, 2004.
- [Sneddon05] Sneddon, R. Eeg detection of early alzheimer disease using psychological task. *Journal of Clinical EEG and Neuro Science, Volume 36, No.3*, págs. 141 – 150, 2005.
- [Sáez09] Sáez, C. *Utilidad diagnóstica, fiabilidad y validez del Fototest en la detección de deterioro cognitivo y demencia*. Editorial de la Universidad de Granada, 1era edición, 2009. ISBN 978-84-693-1998-7.
- [Sánchez14] Sánchez, J. G. Técnicas de toma de datos y análisis de electroencefalografía. *Facultad de Biología Departamento de Fisiología Animal, Universidad de Sevilla*, págs. 1–28, 2014.
- [Taywade12] Taywade, S. A. Eeg signal analysis with different methodologies. *Proceeding on a National Conference on Innovative Paradigms in Engineering and Techonology*, 6, págs. 29–31, 2012.
- [Teplan02] Teplan, M. Fundamentals of eeg measurements. *Measurements sciences review*, 2, págs. 1–10, 2002.
- [Trotter59] Trotter, H. An elementary proof of the central limit theorem. *Arch Math*, 10, págs. 226–234, 1959.
- [Vialatte09] Vialatte, F. B. Improving the quality of eeg data in patients with alzheimers disease using ica. *ICONIP 2008, Part II*, págs. 979 – 986, 2009.

- [Walpole11] Walpole, R. E. *Probability and Statistics for Engineers and Scientists*. Prentice Hall, novena edición, 2011. ISBN 978-0-321-62911-1.
- [Zhang08] Zhang, Y. Usefulness of computed tomography linear measurements in diagnosing alzheimers disease. *Acta Radiol. No. 49*, págs. 91 – 97, 2008.