



DIVISIÓN DE ESTUDIOS DE
POSGRADO E INVESTIGACIÓN



Tesis:

**Detección y Eliminación de Artificios Oculares en Registros
Electro-encefalográficos Neonatales**

Para obtener el grado de:

Maestro en Ciencias en Ciencias de la Computación

Presenta:

I.S.C. Valentín Delgado Lemus

Director de Tesis:

Dra. Laura Cruz Reyes

Co-director:

Dr. Claudio Castellanos Sánchez

Laboratorio de Tecnologías de Información

Cinvestav-Tamaulipas

"2011, Año del Turismo en México"



SUBSECRETARÍA DE EDUCACIÓN SUPERIOR
DIRECCIÓN GENERAL DE EDUCACIÓN SUPERIOR TECNOLÓGICA
INSTITUTO TECNOLÓGICO DE CIUDAD MADERO

Ciudad Madero, Tamps: a 5 de Mayo de 2011

OFICIO No.: US-154/11
AREA: DIVISIÓN DE ESTUDIOS
DE POSGRADO E INVESTIGACIÓN
ASUNTO: AUTORIZACIÓN DE IMPRESIÓN DE TESIS

C. ING. VALENTÍN DELGADO LEMUS
P R E S E N T E

Me es grato comunicarle que después de la revisión realizada por el Jurado designado para su examen de grado de Maestría en Ciencias en Ciencias de la Computación, se acordó autorizar la impresión de su tesis titulada:

"DETECCIÓN Y ELIMINACIÓN DE ARTIFICIOS OCULARES EN REGISTROS ELECTRO-ENCEFALOGRÁFICOS NEONATALES"

Es muy satisfactorio para la División de Estudios de Posgrado e Investigación compartir con Usted el logro de esta meta. Espero que continúe con éxito su desarrollo profesional y dedique su experiencia e inteligencia en beneficio de México.

ATENTAMENTE
"Por mi Patria y por mi Bien"

Mrs. Yolanda Chávez Cincó
M.P. MARÍA YOLANDA CHÁVEZ CINCO
JEFA DE LA DIVISIÓN



c.c.p.: Archivo
c.c.p.: Minuta

MYOHC CINCO

Ave. 10 de Mayo y San Juan S. De la Cruz, Col. Los Maderos, C.P. 24440 Cd. Madero, Tam.
Tels. (833) 332 48 20, Fax: (833) 332 48 20, Cel. 1000, email: nom@itcm.edu.mx
www.itcm.edu.mx

Declaración de Originalidad

Declaro y prometo que este documento de tesis es producto de mi trabajo original y que no infringe los derechos de terceros, tales como derecho de publicación, derechos de autor, patentes y similares.

Además, declaro que en las citas textuales que he incluido (las cuales aparecen entre comillas) y en los resúmenes que he realizado de publicaciones ajenas, indico explícitamente los datos de los autores y las publicaciones.

Además, en caso de infracción a los derechos de terceros derivados de este documento de tesis, acepto la responsabilidad de la infracción y relevo de ésta a mi director y co-director de tesis, así como al Instituto Tecnológico de Cd. Madero y sus autoridades.

Mayo de 2011, Cd. Madero, Tamaulipas.

I.S.C. Valentín Delgado Lemus

Agradecimientos

La presente Tesis es un esfuerzo en el cual, directa o indirectamente, participaron varias personas leyendo, opinando, corrigiendo, teniéndome paciencia, dando ánimo, acompañando en los momentos de crisis y en los momentos de felicidad.

Agradezco a la Dra. Laura Cruz Reyes por haber confiado en mi persona, por la paciencia y por la dirección de este trabajo. Al Dr. Claudio Castellanos Sánchez, que a pesar de sus múltiples ocupaciones aceptó ser mi codirector de tesis, además por los consejos, el apoyo y el ánimo que me brindó.

Al comité tutorial, la Dra. Claudia Gómez Santillán, la Dra. Ma. Lucila Morales Rodríguez y al Dr. Arturo Hernández Ramírez por su paciencia ante mi inconsistencia y su atenta lectura de este trabajo además de sus comentarios en todo el proceso de elaboración de la Tesis y sus atinadas correcciones.

Al Dr. Luz Noé Oliva Moreno por brindarme su confianza y transmitirme sus conocimientos durante mi estancia en el ESCOM-IPN. Así como también a la M.C. Norma Edith García Ávalos por sus invaluable consejos y seguir durante estos dos años la elaboración de este trabajo de tesis.

Gracias también a mis queridos compañeros, que me apoyaron y me permitieron entrar en su vida durante estos dos años de convivir dentro y fuera del salón de clase.

Al Consejo Nacional de Ciencia y Tecnología (CONACYT) por haberme brindado la beca para poder llevar a cabo este trabajo de investigación a través del Instituto Tecnológico de Ciudad Madero en colaboración con el LTI del Cinvestav unidad Tamaulipas.

Todo esto nunca hubiera sido posible sin el amparo incondicional de mi familia, mis padres y mis hermanas.

Gracias a todos . . .

Dedicatoria

Mi tesis la dedico con todo mi amor y cariño.

A ti Dios que me diste la oportunidad de vivir y de regalarme una familia maravillosa.

Con mucho cariño principalmente a mis padres que me dieron la vida y han estado conmigo en todo momento. Gracias por todo a mi papá **Valentín Delgado Hernández** y a mi mamá **María del Pilar Lemus Gómez** por darme la mejor herencia: una educación, para mi futuro y por creer en mí, aunque hemos pasado momentos difíciles siempre han estado apoyándome y brindándome todo su amor, por todo esto les agradezco de todo corazón el que estén conmigo a mi lado.

Y como en alguna ocasión lo dije no existen las palabras para agradecer todo lo que han hecho por mí, todo su esfuerzo y sacrificios para que pudiera realizar uno más de mis sueños, el terminar un posgrado.

A mis hermanas **Pilar, Isabel y Lériða** gracias por estar conmigo y apoyarme siempre, las quiero mucho.

Artículos publicados

Durante el desarrollo del presente trabajo de investigación se elaboraron las siguientes publicaciones:

- * Delgado Lemus Valentín, Cruz Reyes Laura, Castellanos Sánchez Claudio, "Estado del arte en Detección y/o eliminación de Crisis en Electroencefalograma Neonatal", III Encuentro de investigadores del ITCM, Cd. Madero, Tamps., 2009.
- * Delgado Lemus Valentín, Cruz Reyes Laura, Castellanos Sánchez Claudio, García Avalos Norma E., "Detección y eliminación de artificios oculares en registros EEG Neonatales", IV Encuentro de investigadores del ITCM, Cd. Madero, Tamps., 2010.
- * Delgado Lemus Valentín, Cruz Reyes Laura, Castellanos Sánchez Claudio, García Avalos Norma E., Gómez Santillan Claudia G., "Detección y eliminación de artificios oculares en registros EEG Neonatales", "XXIII Encuentro Nacional De Investigación Científica Y Tecnológica Del Golfo De México, Altamira, Tamps., 2011.

Resumen

El análisis y la interpretación de las señales electroencefalográficas (EEG) permiten la comprensión de la función cerebral, ayudando de esta manera en el diagnóstico de estados disfuncionales en neurología tanto de adultos como en la etapa neonatal. Esta última es el caso de estudio en el presente trabajo de investigación.

La presencia de artificios en el registro EEG puede enmascarar actividad de crisis y en consecuencia pueden conducir a clasificaciones erróneas de los registros. El EEG neonatal es muy diferente del EEG de niños mayores y adultos. Hallazgos que son anormales en niños mayores pueden ser normales en ciertos estados y edades de un neonato. Para una mejor interpretación de un EEG se requiere diferenciar la actividad cerebral y la presencia de artificios. Siendo el caso más crítico y relevante el producido por la contaminación debida a la actividad ocular.

Para llevar a cabo la detección y eliminación de artificios oculares se utilizó el Análisis de Componentes Independientes (ICA Independent Component Analysis), el cual es una técnica estadística multivariada utilizada para encontrar las componentes estadísticamente independientes dentro de un espacio multidimensional, en el cual no se conoce información alguna de los datos.

En este trabajo se estudia el algoritmo INFOMAX, una variante de ICA, el cual se basa en la maximización de la información por medio de la entropía relativa. Al aplicar la transformación de las componentes obtenidas por ICA y los datos de entrada, se obtiene la solución al problema conocido en la literatura como Separación Ciega de Fuentes (BSS Blind Source Separation). Cuando las fuentes son artificios oculares mezcladas con señales neurológicas de neonatos la aplicación de INFOMAX no es trivial.

La contribución de este trabajo es una metodología para la detección y eliminación de artificios oculares presentes en registros EEG neonatales. La metodología propuesta consta de las siguientes etapas: a) preparación de datos; b) pre-procesamiento ICA, c) detección de artificios oculares, d) eliminación de componentes identificados como artificios oculares y; e) reestructuración de la señal. La metodología se aplicó a un conjunto de datos reales proporcionados por el Hospital Infantil de Tamaulipas.

Abstract

The analysis and interpretation of electroencephalographic signals (EEG) allows the understanding of the brain function, is very useful for diagnosing dysfunctional neural states in neurology both adults and neonates. The latter is the case study in this research.

The presence of artifacts in EEG signals may mask seizures and therefore lead to misclassification of records. The neonatal EEG is quite different from the EEG in older children and adults. Abnormal findings in older children may be normal in certain states and age of a neonate. For a better interpretation of an EEG is required to differentiate brain activity and the presence of artifacts. Being the most critical and important the pollution produced by ocular activity.

To carry out the detection and removal of ocular artifacts used Independent Component Analysis (ICA), which is a multivariate statistical technique used to find the statistically independent components within a multidimensional space, in which information is not known any of the data.

This paper studies the INFOMAX algorithm, a variant of ICA, which is based on maximizing the information through the relative entropy. By applying the transformation of the components obtained by ICA and the input data, we obtain the solution to the problem known in the literature as Blind Source Separation (BSS). When the sources are ocular artifacts mixed with neurological signals of neonates the implementation of INFOMAX is not trivial.

The contribution of this paper is a methodology for the detection and removal of ocular artifacts present in neonatal EEG recordings. The proposed methodology comprises the following steps: a) data preparation, b) ICA preprocessing, c) detection of ocular artifacts, d) removal of components identified as ocular artifacts and; e) restructuring of the signal. The methodology was applied to a real data set provided by the Hospital Infantil de Tamaulipas.

Índice general

1. Introducción	1
1.1. Justificación	1
1.2. Hipótesis	2
1.3. Descripción de la complejidad del problema	3
1.4. Alcances y limitaciones	3
1.4.1. Alcances	3
1.4.2. Limitaciones	4
1.5. Objetivo general	4
1.6. Objetivos específicos	4
1.7. Estructura de la tesis	5
2. Marco teórico	7
2.1. Definición de señales	8
2.2. Características de las señales	8
2.3. Aplicaciones del procesamiento de señales	9
2.4. Metodología para el tratamiento de señales	10
2.5. Señales electrofisiológicas	11
2.6. Electroencefalograma	11
2.6.1. Introducción	12
2.6.2. Sistema Internacional 10-20	14
2.6.3. Crisis en los registros EEG	15
2.6.4. Artificios en los registros EEG	17
2.7. Filtros	19
2.7.1. Pasa alto	19

2.7.2.	Pasa bajo	19
2.7.3.	Pasa banda	20
3.	Estado del arte	21
3.1.	Método de detección de crisis	21
3.1.1.	Grupo de Kok-Kiong Poh et al. (2008) [34]	21
3.1.2.	Grupo de Piryatinska et al. (2009) [22]	23
3.1.3.	Grupo de Faul et al. (2007) [20]	23
3.1.4.	Grupo de Zarjam et al. (2007) [44]	24
3.1.5.	Grupo de Boashash et al. (2001) [9]	25
3.1.6.	Grupo de Wong Lisa et al. (2008) [37]	25
3.2.	Método de detección y/o eliminación de artificios	26
3.2.1.	Grupo de Deburchgraeve et al. (2008) [54]	26
3.2.2.	Grupo de Greene et al. (2008) [26]	28
3.2.3.	Grupo de Brotchie et al. (2007) [29]	29
3.2.4.	Grupo de Mukheriee et al. (2003) [41]	30
3.3.	Artificios oculares en adultos	31
3.3.1.	Grupo de Gómez Herrero et al. (2006) [47]	31
3.3.2.	Grupo de Guerrero Mosquera et al. (2009) [40]	32
3.3.3.	Grupo de Romero Lafuente et al. (2010) [47]	33
3.4.	Comentarios acerca del estado del arte	34
4.	Técnicas para el análisis de EEG	38
4.1.	Análisis de Componentes Principales (PCA)	38
4.1.1.	Definición de PCA sobre la matriz de covarianzas	39
4.1.2.	Calificación de los componentes principales	39
4.1.3.	Determinación del número de componentes principales	40
4.2.	Análisis de Componentes Independientes	40
4.2.1.	Historia	41
4.2.2.	Modelo ICA	42
4.2.3.	Procesamiento inicial	46
4.2.4.	Restricciones	46
4.2.5.	Independencia estadística	47

4.2.6. Medida de no Gaussianidad	47
4.3. ICA basado en la maximización de la información	48
4.3.1. Algoritmo INFOMAX	55
5. Metodología	57
5.1. Preparación de los registros EEG neonatales	60
5.1.1. Estandarización de los archivos EEG	60
5.1.2. Filtrado de los registros EEG neonatales	64
5.2. Pre-procesamiento INFOMAX extendido	66
5.2.1. Centrado de los datos	67
5.2.2. Blanqueado de los datos	68
5.3. Extensión propuesta del algoritmo INFOMAX para la detección de artificios oculares	69
5.3.1. Aplicación de INFOMAX extendido	69
5.3.2. Obtención de los componentes independientes	70
5.3.3. Identificación de componentes independientes asociados a ar- tificios oculares	70
5.4. Eliminación de artificios oculares	71
5.5. Reconstrucción de el registro EEG neonatal libre de artificios oculares	72
5.6. Evaluación de los resultados	73
6. Pruebas y resultados	75
6.1. Ambiente experimental	75
6.1.1. Infraestructura	75
6.1.2. Base de datos	76
6.2. Determinación de la tasa de aprendizaje	76
6.2.1. Objetivo	76
6.2.2. Método	76
6.2.3. Resultado	77
6.3. Aplicación de INFOMAX	78
6.3.1. Objetivo	78
6.3.2. Método	78
6.3.3. Resultado	79

6.4. Aplicación de INFOMAX extendido	79
6.4.1. Objetivo	79
6.4.2. Método	80
6.4.3. Resultado	80
6.5. Comentarios	81
7. Conclusiones y perspectivas	84
7.1. Aportaciones	85
7.2. Trabajos futuros	86
A. Software EDFBrowser	87
A.1. Características	87
A.2. Comentarios	88
B. Conceptos de estadística	89
B.1. Función de Distribución de Probabilidad o Función de Distribución	90
B.2. Función de Densidad de probabilidad	90
B.3. Valor medio y valor cuadrático medio	92
B.4. Varianza y Desviación estándar	93
B.5. Cumulantes	93
B.6. Variable aleatoria Gaussiana o Normal	94
B.7. Múltiples variables aleatorias	97
B.8. Independencia estadística	100
B.9. Teorema de límite central	102
B.10. Correlación	102
C. Conceptos de álgebra lineal	104
C.1. Sistema de ecuaciones lineales	104
C.2. Matrices y vectores	105
C.3. Matriz Identidad	106
C.4. Producto punto de dos vectores	106
C.5. Producto entre matrices	107
C.6. Matriz invertible	107
C.7. Determinante	108

C.8. Independencia lineal	109
C.9. Valores y Vectores propios (eigenvalores, eigenvectores)	110
D. Conceptos de teoría de la información	111
D.1. Entropía	111
D.2. Entropía conjunta, marginal y condicional	113
D.3. Información mutua	114
D.4. Divergencia de Kullback-Leibler	115
Bibliografía	116

Índice de figuras

2.1. Metodología estándar para el tratamiento de señales	10
2.2. Ondas cerebrales	14
2.3. Colocación de los electrodos en el sistema 10-20.	15
2.4. Tipos de crisis: A) ondas bajas; B) espigas afiladas.	16
2.5. Tipos de crisis: A) Crisis de tren de espigas; B) Crisis oscilatoria . . .	16
2.6. Filtro pasa alto	19
2.7. Filtro pasa bajo	20
2.8. Filtro pasa banda	20
3.1. Tipos de crisis: A) ondas bajas; B) espigas afiladas.	22
4.1. Arreglo de electrodos $y(t)$ y fuentes originales $x(t)$	43
4.2. Ejemplo de componentes independientes x_1 y x_2 con distribución uni- forme	44
4.3. Mezclas de las señales independientes x_1 y x_2	45
4.4. Arquitectura de la red neuronal	49
4.5. Método de la máxima entropía	50
4.6. Función logística y su derivada	51
5.1. Metodología propuesta para la eliminación de artificios oculares en EEG neonatales	59
5.2. Estandarización de los registros EEG neonatales	60
5.3. Menú File del software EDFBrowser	61
5.4. Señales a visualizar en la ventana de EDFBrowser	61
5.5. Convertir a ASCII desde el formato EDF	62

5.6. Contenido del archivo <i>header</i>	62
5.7. Contenido del archivo <i>annotations</i>	63
5.8. Contenido del archivo <i>signals</i>	63
5.9. Contenido del archivo <i>data</i>	64
5.10. La Figura (a) muestra la configuración del filtro pasa bajo y el (b) del filtro pasa alto	65
5.11. Figura superior: EEG en crudo; Figura inferior: EEG filtrado	66
5.12. Señales <i>Fp1</i> y <i>Fp2</i> antes y después del centrado	67
5.13. Función de PCA	68
5.14. Identificación de artificios oculares	72
5.15. Matriz <i>W</i>	73
5.16. Matriz <i>A</i> una vez eliminado los componentes seleccionados	74
5.17. EEG neonatal libre de artificios oculares	74
6.1. ICA en el segmento de 15 segundos de un EEG	78
6.2. EEG filtrado	79
6.3. Componentes independientes después de aplicar INFOMAX	80
6.4. EEG filtrado	81
6.5. Componentes independientes después de aplicar INFOMAX extendido	82
6.6. Mapas de actividad de los componentes independientes	82
6.7. Mapas de actividad de los componentes independientes	83
6.8. EEG libre de artificios oculares	83
A.1. Ejemplos de áreas de trabajo del software EDFBrowser	87
B.1. Ejemplos de posibles funciones de distribución de probabilidad	90
B.2. Ejemplos de posibles funciones de densidad de probabilidad	91
B.3. Variable aleatoria Gaussiana: a)función de densidad y b)función de probabilidad	96
B.4. Distribución de densidades de probabilidad, marginales, condicionales y conjunta (de juntura)	100
D.1. Entropía para el lanzamiento de una moneda	112

D.2. Relación entre la información mutua $I(X; Y)$ y las entropías $H(X)$ y $H(Y)$ 114

Índice de cuadros

3.1. Comparación de los trabajos respecto a crisis.	36
3.2. Comparación de trabajos con respecto a artificios.	37
6.1. Valores para el parámetro alfa en la regla de aprendizaje	77

Capítulo 1

Introducción

Para cumplir con los objetivos del proyecto se propuso una metodología basada en análisis estadístico multivariado de los registros electroencefalográficos neonatales en el dominio de tiempo y/o frecuencia. Además cabe mencionar que este proyecto de investigación está vinculado con un trabajo de investigación de nivel doctorado del Cinvestav-Tamaulipas, en el cual se plantea la detección automatizada de crisis epiléptica en neonatos.

1.1. Justificación

Las anormalidades o crisis, las cuales indican una disfunción en el sistema nervioso central son reflejadas en los registros EEG. A diferencia de los registros EEG adultos, la estructura de la señal de un EEG neonatal tiene una alta capacidad de diagnóstico de crisis a un plazo mayor a la del registro, [33], por lo que tempranamente pudieran detectarse patologías que estarán presentes en la vida adulta.

Algunos estudios han mostrado que ciertas crisis pueden causar cambios irreversibles en el sistema nervioso central [54]. Como una estrategia de neuroprotección, se necesita una detección efectiva de estas crisis. Sin embargo, para que sea una detección óptima se requiere que los registros EEG neonatales se encuentren libres de artificios, debido a que la presencia de estos puede conducir a un mal pronóstico de crisis. Por esta razón se propuso analizar diferentes técnicas computacionales de reducción o eliminación de artificios, que posibilitan una mejor interpretación en el

diagnóstico médico.

Otro aspecto a considerar es que durante el período de maduración el EEG neonatal debe ser interpretado en base a normas clínicas y estadísticas. Por el contrario, la interpretación de un EEG adulto se guían por una norma óptima o ideal (ausencia de patologías) [30].

Otra de las razones que dificultan el análisis EEG neonatal es que la actividad eléctrica del cerebro tiene un comportamiento aleatorio, es decir, no existe hasta la actualidad un patrón similar como ocurren en el caso de los niños mayores o adultos. Esto hace que algunos segmentos del registro EEG puedan llegar a confundirse con segmentos de crisis en los neonatos.

En el presente trabajo de investigación el método de detección y eliminación de artificios está basado en el Análisis de Componentes Independientes (ICA). El principal requisito de ICA es la independencia estadística. La señal EEG producida por los artificios es independiente, por definición, de la señal de origen cerebral y, por tanto, el ICA es un método adecuado para su separación [5].

La aplicación de ICA sobre electroencefalografía no parece sencilla, y solamente pocos estudios han separado las fuentes de actividad espontánea del cerebro aplicándola. A pesar que este método ha sido muy exitoso para encontrar artificios, a la fecha no ha sido aplicado para la detección del artefacto ocular en registros EEG neonatales.

Este trabajo de investigación está vinculado con un proyecto de tesis doctoral de CINVESTAV-Tamaulipas, en el cual se aborda el tema de extracción de crisis a partir del EEG neonatal para la detección temprana del síndrome epiléptico, y como ya se ha mencionado, los registros deben estar libres de artificios para una mejor detección de la crisis epiléptica.

1.2. Hipótesis

La detección de patrones para identificar cualquier alteración en el registro EEG neonatal que el experto realiza manualmente se basa principalmente en detección y eliminación de artificios, así como en la selección, extracción, y clasificación de características. En este trabajo de investigación se propone abordar la detección y

eliminación de artificios, para dar respuesta a la siguiente pregunta de investigación:

¿Será posible automatizar la detección y eliminación del artefacto ocular, aplicando técnicas estadísticas multivariadas en el dominio tiempo y/o frecuencia, además de aplicar técnicas de pre-procesamiento?

1.3. Descripción de la complejidad del problema

Uno de los aspectos que hace la diferencia en el tratamiento de las señales emitidas por un neonato, en contraste con las emitidas por un adulto o un niño mayor, es que la inmadurez que presenta la estructura del cerebro del neonato genera señales más difíciles de interpretar.

Por lo anterior la estructura de las señales emitidas en los adultos y niños mayores tiene trazos bien definidos. En otras palabras, existen patrones que modelan los comportamientos de las señales. Por otro lado, la estructura del cerebro de un neonato aún no tiene patrones definidos, debido a que las respuestas locales emitidas por el cerebro del neonato no alcanzan a sincronizar todo el cerebro, por esta razón, las señales EEG en neonatos generalmente son parciales y no globales como en adultos o niños mayores.

1.4. Alcances y limitaciones

En este punto se describen los aspectos que abarcará este proyecto de investigación y hasta donde se va a llegar.

1.4.1. Alcances

A continuación se listan los alcances de este proyecto de investigación:

* Detección y eliminación de artificios generados por el neonato después de la generación del registro EEG.

- * Evaluación de las técnicas seleccionadas para la detección y eliminación de artificios.
- * Encontrar las características relevantes del registro EEG con la finalidad de reducir la dimensionalidad de los datos.

1.4.2. Limitaciones

A continuación se listan las limitaciones de este proyecto de investigación:

- * El proyecto de investigación se realizará en un plan de un año.
- * Este proyecto de investigación solo se enfocará a la detección y eliminación de un tipo de artificios fisiológico: causado por los movimientos de los ojos (electro-oculograma).
- * Los registros EEG neonatales son obtenidos mediante la colaboración con Cinvestav-Tamaulipas, dichos registros EEG son proporcionados por el Hospital Infantil de Tamaulipas y el Hospital de Perinatología de la Cd. de México.
- * Para la elaboración de las técnicas computacionales, se trabajará en el ambiente del sistema operativo Windows y/o Linux; y el lenguaje de programación MATLAB, C ANSI y Java.

1.5. Objetivo general

Analizar registros EEG neonatales con la finalidad de detectar y eliminar significativamente el artificio ocular que impide encontrar patrones de crisis en registros EEG neonatales. Los registros EEG neonatales libres de artificios permitirán una mejor interpretación médica.

1.6. Objetivos específicos

Los objetivos específicos que se plantean para este trabajo de investigación son los siguientes:

- * Convertir los registros EEG que originalmente se encuentran en formato propio del equipo a formato que pueda ser interpretado por la computadora (ej. Archivo de texto).
- * Analizar técnicas existentes para eliminación de artificios.
- * Reproducir y realizar una comparación de técnicas de detección y/o eliminación de artificios.
- * Desarrollar un método de detección de artificios en registros EEG neonatales.
- * Desarrollar un método de eliminación de artificios en registros EEG neonatales.
- * Evaluar la calidad de los métodos propuestos para eliminación de artificios.
- * Analizar y aplicar algoritmos de extracción de características, con la finalidad de reducir el espacio dimensional de las variables analizadas.

1.7. Estructura de la tesis

La tesis está dividida en siete capítulos, cuyo contenido se menciona brevemente a continuación:

- * **Capítulo 1.** En este capítulo se muestra un panorama general del contenido del trabajo de investigación desarrollado. Además se puede encontrar la complejidad del problema, así como la justificación, los objetivos de la presente tesis. También se presenta la hipótesis planteada del problema a dar solución, en la cual el problema abordado es la detección y eliminación de artificios oculares en registros EEG neonatales.
- * **Capítulo 2.** Aquí se presentan los fundamentos teóricos necesarios para el entendimiento del problema abordado por la presente tesis, que se encuentra en el área de reconocimiento de patrones en señales. Por lo tanto, se puede encontrar el marco fundamental acerca de señales, características y metodología empleada en el análisis de señales. Una vez revisado los conceptos de señales,

se puede encontrar con diferentes tipos de señales como son las electrofisiológicas. El Electroencefalograma (EEG) forma parte de dichas señales. El EEG neonatal es el objeto de estudio en el presente trabajo de investigación, por lo cual se debe conocer los conceptos, historia, localización de electrodos y algunas consideraciones técnicas. Toda esta información es presentada en este capítulo. Finalmente se puede encontrar con el marco teórico básico de señales no cerebrales conocidas como artificios, de los cuales algunos de ellos pueden ser reducidos por algunos filtros.

- * **Capítulo 3.** Los trabajos abordados en el área de EEG neonatales pueden ser revisados en este capítulo. Los trabajos se presentan en dos secciones del capítulo. El primero de ellos habla acerca de la detección de crisis en registros EEG neonatales y el segundo sobre la detección y/o eliminación de artificios en registros EEG neonatales. Finalmente se presentan algunos comentarios acerca de los trabajos presentados.
- * **Capítulo 4.** El marco de referencia de las técnicas del Análisis de Componentes Principales (PCA) así como el del Análisis de Componentes Independientes se encuentra en este capítulo. Además se presenta el algoritmo ICA basado en la maximización de la información conocido como INFOMAX.
- * **Capítulo 5.** En este capítulo se presenta la metodología propuesta para la detección y eliminación de artificios oculares en registros EEG neonatales. Además se presentan las bases teóricas de la extensión del algoritmo INFO-MAX, en su versión INFOMAX extendido.
- * **Capítulo 6.** Se presentan las experimentaciones realizadas para cumplir con los objetivos del trabajo de investigación y dar respuesta a la hipótesis planteada. Además se muestran los resultados obtenidos en cada una de las experimentaciones.
- * **Capítulo 7.** Finalmente se presentan las conclusiones y perspectivas generales de este trabajo de tesis.

Capítulo 2

Marco teórico

Para desarrollar las técnicas de análisis de señales y sistemas es necesario establecer un marco teórico de referencia analítico que capture las ideas intuitivas de señales y sistemas. Las señales pueden describir una variedad amplia de fenómenos físicos. A pesar de que las señales se pueden representar de diversas maneras, en la mayoría de los casos la información dentro de una señal está contenida en un patrón de variaciones.

Existe una extensa variedad de problemas y aplicaciones de análisis de señales como son en señales de audio, de voz, en sistemas de energía, señales eléctricas como las que emiten los organismos vivos en los diferentes procesos fisiológicos que lleva a cabo a través de los sistemas por los cuales está compuesto. A las señales que emiten los organismos vivos se les denomina señales electrofisiológicas.

Existen diversos tipos de señales electrofisiológicas como las que emite el cerebro, el corazón, los músculos, entre otras. El enfoque del presente trabajo de investigación se aborda en el área de reconocimiento de patrones, específicamente en el análisis de señales electro-fisiológicas en neonatos. En particular, la señal a analizar es la emitida por el cerebro en neonatos (Electroencefalograma). Dicho análisis se refiere a la detección y eliminación de *artificios*¹, con la finalidad de facilitar la interpretación de un registro EEG neonatal. En el área de reconocimiento de patrones la detección de artificios en neonatos es un reto ya que generalmente se confunden con crisis.

¹Los *artificios* son señales emitidas por fuentes externas al cerebro y que pueden afectar al registro EEG neonatal

2.1. Definición de señales

Las señales transmiten información. El sonido y la imagen son ejemplos de señales. Matemáticamente las señales se pueden modelar como funciones. Una *señal* es una función que mapea un dominio, generalmente tiempo o espacio, en un rango, con frecuencia una medida física. Las señales pueden provenir de diferentes medios, como la radio y la TV que generan señales electromagnéticas, así como procesos fisiológicos que producen señales eléctricas [2].

Las señales se pueden dividir en dos tipos [56]:

- a. Una *señal analógica* es una función $s(t)$ de valores reales finitos de una variable continua t , llamada tiempo, definida por todo los valores en el intervalo $-\infty < t < +\infty$.
- b. Una *señal digital* es una secuencia de valores discretos limitados S_n con un índice simple n , llamado tiempo discreto, definido por todos los valores del intervalo $n = -\infty, \dots, +\infty$.

2.2. Características de las señales

A continuación se muestran las características que algunas señales pueden tener [56]:

- a. **Determinística.** Estas señales son generadas por algunos algoritmos no probabilísticos. Son reproducibles, predecibles (al menos en intervalos de tiempo cortos) y matemáticamente bien comportadas. Las señales determinísticas pueden ser periódicas y no periódicas. Las primeras se repiten después de un tiempo conocido como período. De lo contrario, son señales no periódicas.
- b. **Estocásticas.** Estas señales son generadas por sistemas que contienen aleatoriedad. En cualquier ventana temporal la señal se comporta como una variable aleatoria, la cual tiene una media y varianza definida.
 - **Estacionarias.** Una señal estacionaria significa que su descripción probabilística no cambia con el tiempo. Lo cual implica que todos los estadísticos de la señal, tal como la media y la varianza, son constantes.

- **No estacionaria.** Cuando una señal se hace más débil o más fuerte, o de alguna manera más variable.

2.3. Aplicaciones del procesamiento de señales

El procesamiento de señales es la disciplina de detección, manipulación y extracción de información de cantidades físicas que varían con el tiempo (señales). Algunos ejemplos de procesamiento de señales se muestran a continuación [56].

- * **Procesamiento y almacenamiento de comunicación por voz.** La idea principal de la comunicación entre humanos es el lenguaje. Un humano emite información como una señal acústica que puede ser detectada por otros humanos. Cuando no existe un nexo entre las personas que deseen conversar, se debe proveer un mecanismo para transferir la señal de un lugar a otro. Cuando no están disponibles, se necesita grabar la señal acústica para reproducirla después.
- * **Aplicaciones industriales.** Las mediciones de modos de vibraciones y el descubrimiento de sus causas y problemas estructurales mecánicos, podrían indicar una aplicación industrial de procesamiento de señales. El control de procesos químicos depende mucho de la instrumentación empleada en el procesamiento avanzado de señales.
- * **Ingeniería biomédica.** El cerebro humano se puede ver como una computadora paralela masiva que contiene cerca de 10^{15} unidades de procesamiento llamadas neuronas. Estas neuronas emiten impulsos eléctricos que no son exactamente observables, pero que colocando varios electrodos en el cuero cabelludo, son detectables. Estos registros son conocidos como electroencefalograma. La actividad eléctrica del corazón también puede ser monitoreada usando el electrocardiograma.

2.4. Metodología para el tratamiento de señales

Se distinguen cuatro etapas en la metodología estándar para el análisis de señales [3]. La Figura 2.1 muestra las cuatro etapas en forma gráfica. A continuación se detalla cada una de estas etapas.

- a. **Adquisición de señales.** Comprende la transformación de las señales, que originalmente no son eléctricas, su acondicionamiento y, su transformación de analógica a digital, para que la señal pueda ser leída por la computadora.
- b. **Pre-procesamiento.** Conlleva preparar la señal para poder extraer los parámetros importantes, aceptando además que gran parte de la información que trae originalmente es redundante.
- c. **Selección o extracción de características.** Consiste en definir, del conjunto de la señal, aquellos datos cuantificables que se estiman que son de importancia.
- d. **Clasificación de la señal.** En esta etapa se aplican algoritmos computarizados que permitan diagnosticar situaciones y otorgarles un significado.



Figura 2.1: Metodología estándar para el tratamiento de señales

2.5. Señales electrofisiológicas

Los organismos vivos están formados por muchos sistemas, el cuerpo humano, el sistema nervioso y el sistema cardiovascular, entre otros. Cada sistema está compuesto por varios subsistemas que llevan a cabo varios procesos fisiológicos. Los procesos fisiológicos son fenómenos complejos, la mayoría de los procesos son manifestados como señales eléctricas que reflejan su naturaleza y sus actividades. A estas señales se les denomina señales electro-fisiológicas [39].

Las señales electrofisiológicas representan variables fisiológicamente relevantes de las que nos interesa su curso temporal. La variable puede ser un voltaje muy pequeño, como en el EEG, o uno de mayor magnitud, como en el electrocardiograma (ECG). En todos estos casos las señales varían con el tiempo [3].

Existen varios tipos de señales electro-fisiológicas, algunas de ellas se enlistan a continuación [39]:

- * **Electroencefalograma (EEG)**. Registra la actividad eléctrica del cerebro.
- * **Electromiograma (EMG)**. Registra la actividad eléctrica de los músculos.
- * **Electrocardiograma (ECG ó EKG)**. Registra la actividad eléctrica del corazón.
- * **Electrogastrograma (EGG)**. Registra la actividad eléctrica del estómago.

Una de las señales que aportan mayor información sobre el estado clínico de un paciente es el EEG, debido a que registra la actividad eléctrica del cerebro. Dicho registro nos puede ayudar a evaluar traumatismos, identificar tumores, infecciones, Alzheimer, crisis epiléptica, entre otros. Por lo tanto, como el presente trabajo forma parte de un proyecto de investigación en el cual se aborda el área de detección automatizada de crisis epiléptica, se enfatizará la señal EEG.

2.6. Electroencefalograma

El electroencefalograma (EEG) es una medida neurofisiológica de la actividad eléctrica espontánea del cerebro que es registrada a través de electrodos colocados

en el cuero cabelludo, dichas señales contienen información acerca del estado de salud de un paciente. Además, es una herramienta principal de análisis para el estudio de enfermedades o crisis neurológicas como Parkinson, esquizofrenia o epilepsia [11, 12, 24, 33].

2.6.1. Introducción

El EEG fue inventado por Hans Berger en 1924. Tomando en cuenta que el EEG es una manifestación de la actividad eléctrica del cerebro, se debe remontar al año de 1770 cuando el italiano Luigi Galvani publicó sus observaciones sobre la electricidad del animal. Sin embargo, sus observaciones no tuvieron impacto por más de cincuenta años.

Las primeras descripciones sobre la existencia de una actividad eléctrica del cerebro, fueron realizadas por el fisiólogo inglés Richard Caton, quien demostró gran interés en los trabajos de Edouard Hitzig y Gustav Theodor quienes habían demostrado la evidencia de respuestas motoras locales después de la estimulación eléctrica en varias áreas del cerebro en perros.

Sin embargo, se considera que el padre de la EEG humana fue Hans Berger, jefe de la unidad de psiquiatría de la Universidad de Jena (Alemania), quien después de una gran serie de estudios, realizó el 6 de julio de 1924 el primer registro de las oscilaciones rítmicas del cerebro en un joven de 17 años.

El primer registro de actividad eléctrica cerebral colocando electrodos en contacto con la corteza cerebral, fue realizado por Forester y Altenburger en 1935 [42].

En los adultos o niños mayores las células del cerebro se comunican entre sí produciendo pequeños impulsos eléctricos. En un EEG, esta actividad eléctrica tenue se mide colocando electrodos en el cuero cabelludo. Para fines de este trabajo de investigación se trabajará en la etapa neonatal.

El período *neonatal*² constituye una época muy especial del desarrollo de un recién nacido con cambios significativos en la maduración cerebral. Estos cambios se ven reflejados en el EEG [25].

El EEG neonatal presenta unos de los retos más difíciles en la interpretación de un EEG, el cual es significativamente diferente al de un niño mayor y de un adulto.

²Un *neonato* es un recién nacido con 27 días o menos de nacido.

Técnicamente, la adquisición de un EEG neonatal es más difícil y diferente a la de un EEG adulto. La incapacidad de los pacientes para cooperar, el tamaño pequeño de la cabeza y su naturaleza frágil, hacen que la aplicación de los electrodos y la ejecución de la prueba se dificulten [6].

El registro EEG neonatal, se ha convertido en una herramienta poderosa debido a que combinan los principios generales de la electroencefalografía con el conocimiento clínico del período neonatal. Además se debe tener en cuenta el acelerado ritmo del desarrollo cerebral durante el período neonatal. El EEG puede proporcionar mayor información si se registra en períodos que incluyan sueño y vigilia [11].

Algunos patrones que son anormales en niños mayores podrían ser normales en ciertos estados y *edad concepcional*³ [25] del neonato.

Se pueden detectar cuatro ritmos distintos en función del estado en el que se encuentra la persona, como se muestra en la Figura 2.2 [1]:

- a. **Ritmo beta.** Cuando la persona está despierta y atenta; aparece una frecuencia de ondas más altas y la frecuencia oscila entre 13-50 Hz. Las ondas tienen una menor amplitud o voltaje y una mayor frecuencia.
- b. **Ritmo alfa.** Cuando la persona está despierta y en reposo con los ojos cerrados, la actividad dominante es en la banda alfa de 8-12 Hz. Con una mayor amplitud en la región occipital donde en promedio y de acuerdo a la edad alcanza los 50 hasta 120 micro voltios. Característicamente se atenúa o llega a bloquearse al abrir los ojos o con estímulos auditivos.
- c. **Ritmo theta.** Esta actividad es de bajo a moderado voltaje y no supera los 100 microvoltios. Es usual en menores de 15 años y en el adulto somnoliento.
- d. **Ritmo delta.** Son ondas de sueño profundo, ondas amplias y de alto voltaje, entre 0 y 5-3 Hz.

³La *edad concepcional* de un neonato es la edad desde su nacimiento más el tiempo del último día del período menstrual de la madre [6].

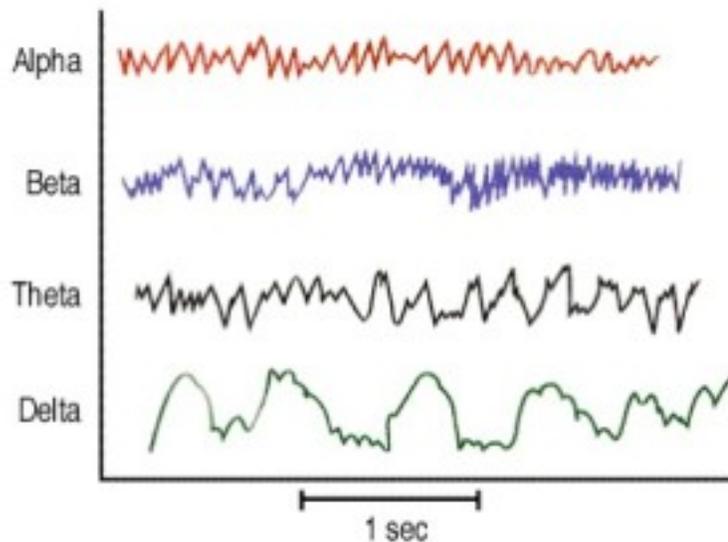


Figura 2.2: Ondas cerebrales

2.6.2. Sistema Internacional 10-20

El sistema de colocación de electrodos más utilizados para la obtención de un EEG es el sistema 10-20 de la Federación Internacional de Sociedades de Electroencefalografía [7], el cual permite asegurar con bastante confiabilidad la colocación de los electrodos en todas las partes de la cabeza.

Las diferentes posiciones de los electrodos se determinan a través de ciertas medidas entre unas marcas estandarizadas sobre el cuero cabelludo. Estas medidas permiten establecer una red de líneas cuyas intersecciones permiten posicionar los electrodos, dando lugar a una separación entre ellos de un 10 o 20 por ciento de la longitud total de la línea sobre la cual están colocados los electrodos, de ahí su nombre de sistema 10-20 [7].

Cinco puntos son establecidos a lo largo de la línea marcada entre la parte frontal (nasion) y la parte trasera (inion): polo frontal (Fp), frontal (F), central (C), parietal (P), occipital (O). El primer punto (Fp) está a 10 % de la distancia nasion-inion, por encima de la nasion, el segundo punto (F) está a 20 % de la distancia del punto Fp, y así en cada 20 % se coloca el punto central, parietal y occipital y el último punto se coloca a 10 % de la distancia sobre el inion. En la Figura 2.3, se muestra claramente

los puntos descritos anteriormente [15].

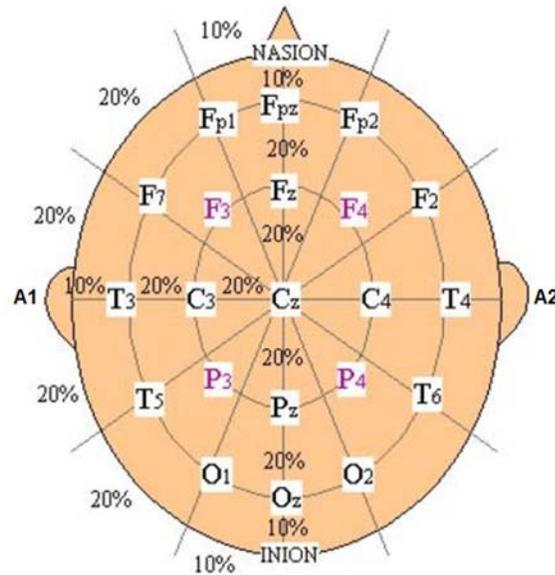


Figura 2.3: Colocación de los electrodos en el sistema 10-20.

2.6.3. Crisis en los registros EEG

El EEG normal consiste de actividad de baja frecuencia, actividad aleatoria, o puede contener espigas aleatorias [33]. Por otro lado, las crisis son los signos más comunes de disfunción neurológica en el neonato y requiere atención médica inmediata. Además, son notoriamente difíciles de diagnosticar como señales clínicas, debido a que pueden ser muy sutiles o completamente ausentes [26]. La actividad de crisis en neonatos puede localizarse en una parte, o esparcida a lo largo de las áreas del cerebro. En general, las ondas de crisis neonatales pueden consistir de descargas rítmicas, espigas, ondas Sharp y ondas bajas [20].

En el trabajo de Kok-Kiong et. al. [33] se identificaron dos tipos de crisis: ondas bajas (slow waves) y espigas afiladas (Sharp spikes); las Figuras 3.1A y 3.1B muestran las crisis respectivas. El eje Y, corresponde al voltaje (V), en este caso está dado en μV y el eje X, indica el tiempo.

Adicionalmente, en el trabajo de Deburchgraeve et. al. [54] han identificado dos tipos de crisis más: tren de espigas y crisis oscilatorias; las Figuras 2.5A y 2.5B

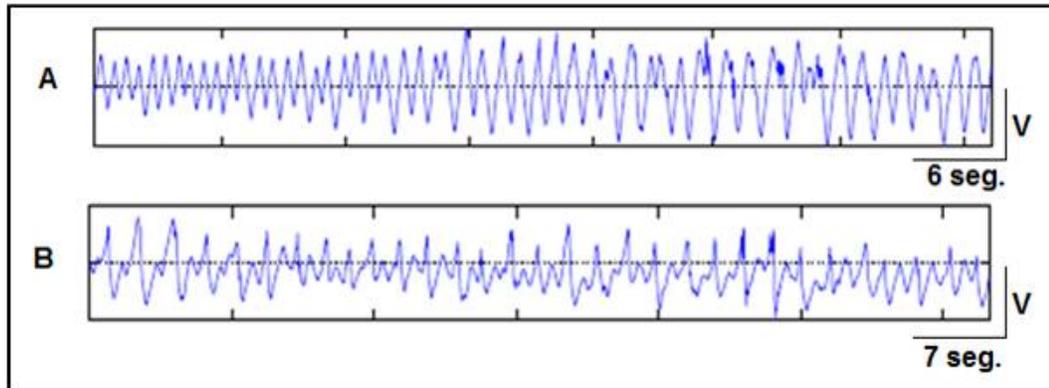


Figura 2.4: Tipos de crisis: A) ondas bajas; B) espigas afiladas.

muestran las crisis respectivas.

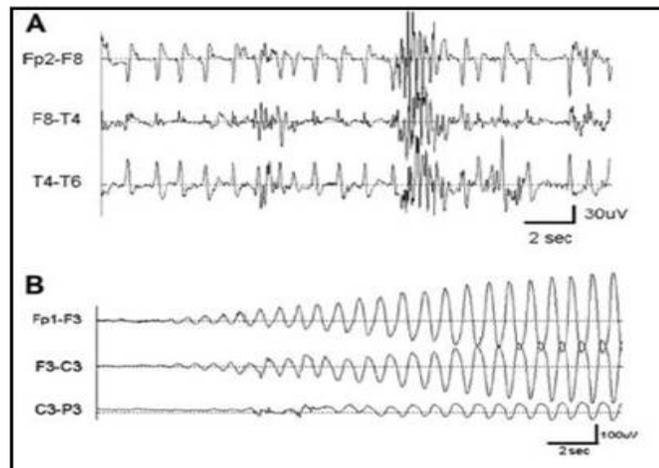


Figura 2.5: Tipos de crisis: A) Crisis de tren de espigas; B) Crisis oscilatoria

En la mayoría de los casos los registros EEG neonatales pueden contener señales no cerebrales que afectan el comportamiento de los registros EEG. Estas señales dificultan el análisis e interpretación clínica de los registros. A estas señales no cerebrales se les denomina *artificios*. Los artificios son causados por diferentes fuentes como las interferencias electricas, mala colocación de los electrodos, y los causados por el paciente, como pueden ser movimientos musculares, cardíacos y oculares. La detección y eliminación de los artificios oculares es el motivo principal de este trabajo de investigación.

2.6.4. Artificios en los registros EEG

“Por arteificio se entiende toda *deflexión del inscriptor*⁴ o de los inscriptores que no provengan del cerebro” [18], en otras palabras, son señales indeseables que pueden alterar las características de las señales electrofisiológicas, dificultando la interpretación.

Los registros EEG neonatal frecuentemente son contaminados por artificios los cuales pueden provenir de diferentes fuentes: instrumentales, fisiológicos y externos [18]. Estos artificios pueden enmascarar y llegar a parecer una crisis en el neonato y con frecuencia pueden causar confusión al aplicar los algoritmos de detección de crisis [41]. Por lo tanto, la presencia de artificios en los registros EEG puede conducir a clasificaciones erróneas de las señales [27].

De acuerdo con Walter y Parr los artificios se dividen según sus causas, en tres principales categorías [18]:

a. Artificios instrumentales

- * *Artificios por defectos del aparato*: los artificios por defecto del canal pueden manifestarse de distintas maneras, por ejemplo, con ruidos, disminución de amplitud o ausencia definitiva del registro y ondas irregulares a las ondas cerebrales. Su reconocimiento es casi siempre fácil porque generalmente afectan un solo canal.
- * *Artificios producidos por el electrodo*: cuando un electrodo ha sido colocado defectuosamente, haciendo mal contacto, esto da origen a ondas de las formas más variadas, desde agudas a lentas, irregulares, esporádicas o repetitivas. Lo mismo ocurre si el electrodo ha sido mal fabricado o se ha roto con la manipulación. El reconocimiento de los artificios producidos de esta manera, es fácil, ya que se hallan limitados a un solo canal.

b. Artificios fisiológicos

Proviene del paciente y constituyen la mayoría de los artificios encontrados en el registro EEG. Los potenciales de la retina, del corazón y de los músculos de la cara, la frente y el cuello, dan lugar a cambios de voltaje casi siempre mucho mayores que los provenientes del cerebro.

⁴Una *deflexión del inscriptor* es un cambio en las variaciones del registro de la señal eléctrica.

- * *Artificios por movimiento ocular*: los movimientos oculares son captados por los electrodos ubicados cerca de los ojos. De todos los artificios fisiológicos, éstos son los más importantes porque pueden ser confundidos con ondas lentas prefrontales o temporales, causando confusión de la existencia de una lesión cerebral o a la inversa. A veces estos artificios son difíciles de evitar, ya que la única forma para ello, es conseguir que el paciente no mueva los ojos.
- * *Potenciales musculares*: los potenciales musculares son detectados rápidamente por el electroencefalógrafo, dando lugar a un EMG. Cuando se producen potenciales musculares, estos pueden enmascarar los registros, careciendo de valor desde un punto de vista clínico.
- * *Potenciales cardíacos (ECG) y pulso*: suelen ser captados por los electrodos auriculares. Los pacientes obesos, de cuello robusto, pueden mostrar potenciales ECG en todos los electrodos del cuero cabelludo. A la fecha son difíciles de eliminar.
- * *Electrodermograma*: la transpiración produce cambios en la resistencia de la piel, que se produce por cambios en forma de ondas sumamente lentas, con actividad eléctrica del cerebro superpuesta. Estos cambios de la piel pueden ser captados por cualquier electrodo. Una forma de eliminar este artefacto es refrescando el ambiente.

c. **Artificios externos**

Los cables que suministran corriente alterna al edificio pueden producir interferencia y ser captados por el electroencefalógrafo. El encendido de la luz, o los chispazos iniciales del encendido de tubos fluorescentes, suelen ocasionar potenciales artificiales que se distribuyen en varios canales. También aparatos como ventiladores, teléfonos, desconexión accidental de la conexión a tierra producen una interferencia.

Algunos tipos de artificios pueden ser reducidos computacionalmente aplicando algunos filtros. La finalidad de algunos de los filtros es poder depurar una señal a través de un rango de frecuencias, en algunas ocasiones se puede reducir el artefacto de tipo externo.

2.7. Filtros

En algunas aplicaciones se requiere modificar las amplitudes relativas de los componentes de frecuencia de una señal o tal vez eliminar por completo algunos de ellos. A este proceso se le llama *filtrado*. En los sistemas lineales invariantes en el tiempo, como en el caso de los EEG neonatales, el espectro de salida es igual al de la entrada, multiplicado por la respuesta en frecuencia [45].

Un *filtro digital*, es un proceso computacional (algoritmo) que convierte una secuencia de números, que representa la entrada, a otra secuencia, que representa la salida. Por lo tanto, un filtro digital, además de filtrar las bandas de frecuencias deseadas, puede también ser usado como un medio computacional de soporte de otras funciones tales como la integración, diferenciación y la estimación [51]. Entre los filtros más comunes se tienen: pasa alto, pasa bajo y pasa banda [19].

2.7.1. Pasa alto

Las señales de interferencia de baja frecuencia son depuradas por un filtro pasa alto. Este filtro suprime frecuencias que son menores al límite inferior de frecuencia que se ha establecido para él (Ver Figura 2.6).

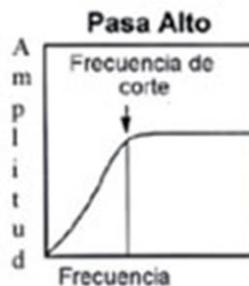


Figura 2.6: Filtro pasa alto

2.7.2. Pasa bajo

Las señales de interferencia de alta frecuencia son depuradas por un filtro pasa bajo. Este filtro suprime frecuencias que son mayores al límite superior de frecuencia que se ha establecido para él (Ver Figura 2.7).

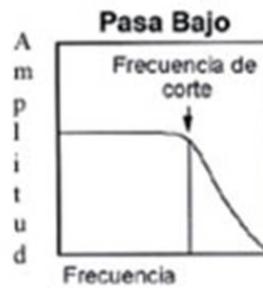


Figura 2.7: Filtro pasa bajo

2.7.3. Pasa banda

El filtro pasa banda tiene influencia en las señales con componentes de baja y alta frecuencia, con respecto del filtro de pasa banda. Un filtro pasa banda es una combinación de filtros pasa alto y pasa bajo y es determinado por la frecuencia central (Ver Figura 2.8).

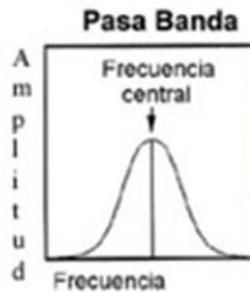


Figura 2.8: Filtro pasa banda

Capítulo 3

Estado del arte

En el marco teórico se describieron los fundamentos teóricos básicos de las áreas de estudio para este trabajo de investigación: las señales electrofisiológicas y el EEG neonatal. Dentro del tópico de EEG neonatal se puede encontrar que un registro EEG puede estar compuesto por señales cerebrales y artificios.

Las áreas abordadas en los trabajos recientes de la literatura acerca del tratamiento de registros EEG neonatales, se pueden dividir en dos: detección de crisis y detección y/o eliminación de artificios. Además, en el cuadro 3.1 y 3.2 se presenta una comparación de dichos trabajos. Se presenta también la descripción de algunos trabajos en los cuales se aborda la temática de artificios oculares en registros EEG de adultos.

3.1. Método de detección de crisis

Se describen algunos de los trabajos relacionados con la detección de crisis en EEG neonatales.

3.1.1. Grupo de Kok-Kiong Poh et al. (2008) [34]

Este grupo de trabajo considera dos clases de señales paramétricas con un número finito de grados de libertad por unidad de tiempo, lo cual es llamado tasa de innovación. A pesar de que las señales no están limitadas en banda, son muestreadas

uniformemente por debajo o arriba de la tasa de innovación usando un núcleo apropiado y posteriormente pueden ser reconstruidas resolviendo un sistema de ecuaciones lineales.

Los datos utilizados por este grupo de investigación son recolectados del Laboratorio Pediátrico EEG, del Hospital de la Universidad Nacional en Singapur. Además utilizan datos artificiales generados por un modelo basado en [49]. Los registros cuentan con 12 canales de datos EEG de acuerdo al Sistema 10-20 y con una frecuencia de muestreo de 200 Hz.

En este trabajo se abordan dos tipos de crisis: ondas bajas y picos afilados; las Figuras 3.1A y 3.1B muestran las crisis respectivas. El eje Y, corresponde al voltaje (V), en este caso está dado en μV y el eje X, indica el tiempo.

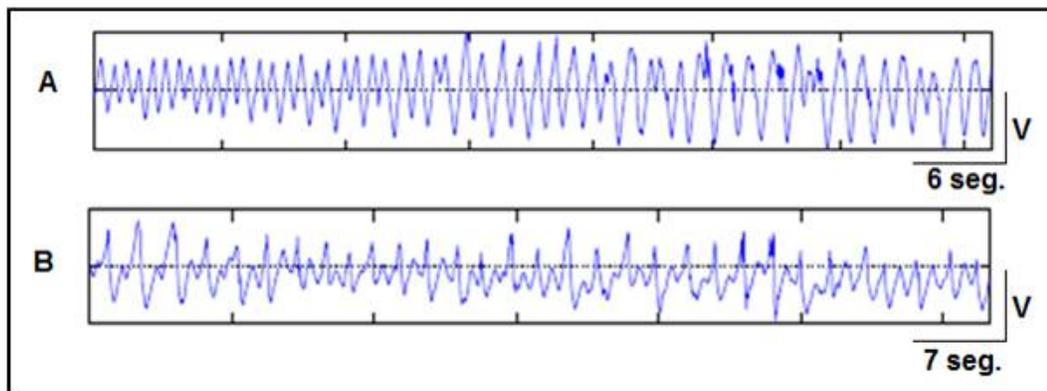


Figura 3.1: Tipos de crisis: A) ondas bajas; B) espigas afiladas.

Para la detección de crisis utilizan una técnica llamada Tasa Finita de Innovación, en la cual se modela la señal como una ranura (spline) lineal no uniforme con k piezas. Además cabe mencionar que no realiza ningún tipo de pre-procesamiento a los datos EEG y que el dominio en el cual trabaja es el frecuencia.

Con las experimentaciones realizadas se pudo determinar que modelando la señal con su tasa de innovación, la señal EEG puede ser reconstruida usando solamente $2k$ coeficientes contiguos de Fourier con un bajo porcentaje de error.

3.1.2. Grupo de Piryatinska et al. (2009) [22]

En este trabajo se integra y adapta un conjunto de herramientas estadísticas, matemáticas y computacionales avanzadas para el análisis de los estados del sueño del neonato basado en EEG.

Los datos utilizados por este grupo de investigación fueron coleccionados en la Universidad de Pittsburgh. La base de datos cuenta con 21 registros de pacientes término y 16 pretérminos, de la misma edad postconcepcional de 40 semanas, donde cada uno contiene una señal EEG muestreada a 64 Hz.

Aceptando la hipótesis de que las señales son estacionarias, proponen un método para separar las señales EEG en segmentos estacionarios estimando los puntos de transición del registro. En este enfoque se combinó el esquema de detección de puntos de cambio, análisis de grupos y una variedad de espectros tradicionales y no tradicionales y características no lineales en diferentes combinaciones. En este trabajo, se analizó el canal $Fp1 - C3$.

Los autores desarrollaron una técnica de separación automática de diferentes estados de sueño en señales EEG neonatales. La efectividad fue de 80-90 % en la detección de los estados del sueño.

3.1.3. Grupo de Faul et al. (2007) [20]

Debido a la naturaleza no lineal del EEG neonatal, Faul et al. proponen un modelo de Proceso Gaussiano (PG) para caracterizar la señal, y contrastan sus resultados con los modelos de redes neuronales. Un modelo PG provee un estimado de la varianza de su resultado predicho, el cual puede ser interpretado como un nivel de confianza del modelo. Esta medida de varianza es una ventaja principal sobre los modelos de redes neuronales.

Todos los registros EEG fueron tomados de bebés recién nacidos con crisis en la Unidad de Cuidados Intensivos Neonatales del Hospital de Maternidad Amalgamado, en Cork, Irlanda. El equipo de registro obtuvo los datos de 12 canales usando el Sistema 10-20 de localización de electrodos para neonatos ($F4 - C4, C4 - P4, P4 - O2, F3 - C3, C3 - P3, P3 - O1, T4 - C4, C4 - Cz, Cz - C3, C3 - T3, T4 - O2, T3 - O1$). Además dos canales adicionales fueron utilizados para registrar el ECG y la respira-

ción.

Todos los registros EEG fueron digitalizados y convertidos a formato EDF. El conjunto total de registros corresponden a una mujer y 3 hombres. El conjunto total de datos contiene 277 crisis identificadas.

Este grupo de trabajo propone trabajar con PG para la detección de crisis epilépticas en recién nacidos, utilizando la varianza como un nivel de confianza. Los datos fueron inicialmente muestreados a 200 Hz y posteriormente reducidos a 60 Hz antes del análisis para acelerar el proceso de modelación. En este trabajo se normalizaron los datos para tener una media de cero. La ventana de análisis utilizada es de un segundo, con 1/6 segundos superpuestos. Durante los eventos de crisis la predicción de la varianza disminuyó dramáticamente, esta mostró una diferencia estadística entre los eventos de crisis y en donde hubo ausencia de crisis.

3.1.4. Grupo de Zarjam et al. (2007) [44]

En este trabajo se comparan dos enfoques diferentes para detección de crisis en señales EEG de recién nacidos. El primer enfoque es un método basado en tiempo-frecuencia (TF), en el cual la diferencia entre crisis y no crisis está basada en la distancia TF. El segundo método propuesto se basa en la Transformada Discreta Wavelet (TDW), en el cual el esquema de detección se basó en observar los cambios en el comportamiento de pocas medidas estadísticas de los coeficientes wavelet del EEG. Los resultados mostraron que TDW es ligeramente mejor en rendimiento que la TF.

Los registros EEG neonatales usados fueron registrados bajo condiciones clínicas del Hospital Infantil Royal de Brisbane en Australia. Cada registro contiene 20 canales EEG de acuerdo al Sistema 10-20. Los datos fueron grabados en forma digital con una frecuencia de muestreo de 256 Hz y una sensibilidad de $100 \mu V$.

Los datos fueron pre-procesados utilizando un filtro pasa banda en el rango de 0.5-70 Hz, el dominio que se aborda para el análisis es de TF. La presencia de crisis en todos los canales EEG fue dada por la opinión de un experto. Para la detección de crisis epilépticas se analizan segmentos de 8 segundos.

Los resultados obtenidos por este grupo de trabajo indican que la distancia medida fué confiable para la detección de crisis. Además se determinó que el método

basado en TF, en este caso la distancia Kullback-Leibler, es ligeramente mejor que la Transformada Discreta Wavelet.

3.1.5. Grupo de Boashash et al. (2001) [9]

En este trabajo se identifican los patrones iniciales mediante un análisis de TF de señales de crisis en EEG. El análisis detectó dos tipos de crisis: picos afilados y picos repetitivos, y se llevó a cabo en segmentos de 127 datos discretos. De acuerdo con los investigadores la caracterización de señales EEG no estacionarias en el dominio de TF es el primer paso hacia un método global de detección y clasificación de crisis usando procesamiento de señales en tiempo frecuencia.

Los datos utilizados contienen cinco canales de EEG de acuerdo al Sistema 10-20. Los datos fueron registrados con una frecuencia de muestreo de 256 Hz. Para la detección de artificios se utilizaron tres canales adicionales que son el EOG, ECG y la respiración. Además los registros grabados fueron obtenidos del Hospital de la Mujer Royal de Brisbane en Australia, y pertenecen a dos recién nacidos que han sido clínicamente identificados que contiene crisis.

Boashash et al. confirmaron la hipótesis de identificar segmentos de crisis en un registro EEG neonatal puede ser fácilmente caracterizado por un FM lineal.

3.1.6. Grupo de Wong Lisa et al. (2008) [37]

Se aplicaron algoritmos de segmentación de EEG en señales EEG neonatales no estacionarias. Los algoritmos evaluados fueron: medida del error espectral, tasa de verosimilitud generalizada y operador de energía no lineal. Los segmentos fueron comparados con la distribución de tiempo-frecuencia de la señal EEG. Se encontró que la tasa de verosimilitud generalizada obtuvo el mejor rendimiento. Este análisis se realizó en 10 segmentos de un minuto y dos de 20 segundos.

Los datos utilizados fueron elegidos aleatoriamente de 12 registros de 2 canales, 10 de los cuales son segmentos de un minuto y los otros son de dos segundos de longitud, los cuales fueron utilizados en la fase de evaluación, donde cada uno de los canales fue procesado por separado por cada uno de los cuatro métodos. Para llevar a cabo una evaluación apropiada este grupo de investigación se compararon

con los resultados de la distribución TF de la señal original, dando como resultado el rendimiento de la tasa de verosimilitud generalizada fue el mejor de todos los algoritmos que fueron evaluados.

Hasta el momento se ha descrito el estado del arte en detección de crisis en registros EEG neonatales. Sin embargo, para llevar a cabo con éxito la detección de dichas crisis, es necesario la implementación de un sistema de detección y eliminación de artificios con la finalidad de reducir las detecciones falsas de crisis. En la siguiente sección se aborda el estado del arte en detección y/o eliminación de artificios.

3.2. Método de detección y/o eliminación de artificios

Se describen algunos de los trabajos relacionados con la detección y/o eliminación de artificios en EEG neonatales.

3.2.1. Grupo de Deburchgraeve et al. (2008) [54]

Para la detección de crisis neonatal se desarrolla un algoritmo el cual simula un observador humano leyendo el EEG. Identificaron dos características principales de crisis neonatales las cuales conducen a su detección por un observador humano. La primera característica es que todas las crisis representan un cambio claro relativo en un EEG, la segunda y más importante es la repetitividad de la señal.

Los registros EEG neonatales fueron proporcionados por el Hospital Infantil Sofia. La base de datos consta de 26 registros de video-EEG de neonatos con asfíxia perinatal de los cuales 21 presentaban crisis electrográficas y 5 estaban libre de crisis. Los registros que contaban con 17 canales de acuerdo al Sistema 10-20 fueron 14 ($Fp1$, $Fp2$, $F3$, $F4$, $C3$, $C4$, Cz , $P3$, $P4$, $F7$, $F8$, $T3$, $T4$, $T5$, $T6$, $O1$ y $O2$). Mientras que el resto de los registros solo contaban con 13 electrodos de acuerdo al Sistema 10-20 a excepción de los canales $F3$, $F4$, $P3$ y $P4$. Los registros fueron realizados 24 horas después de su nacimiento. La frecuencia de muestreo de los registros EEG fue de 256 Hz.

En [54] se identificaron dos tipos de crisis: tren de spikes (Fig. 2.5A) y crisis

oscilatoria (Fig. 2.5B). Todas las crisis neonatales pueden ser clasificadas en una de estas crisis o en una combinación de estas. En este trabajo se aplicaron dos algoritmos para la detección de crisis, con la aplicación de los dos algoritmos se obtienen mejores resultados.

El enfoque propuesto por este grupo de investigación indicó que un algoritmo de detección de tren de espigas debería detectar un tren repetitivo de segmentos de EEG de espigas aislados. Por lo tanto, el algoritmo consta de dos pasos, el primer paso consistió en aislar los segmentos de espigas y el segundo en analizar la similaridad entre estos segmentos aislados.

El segundo algoritmo utilizado en este trabajo es el de detección de crisis de tipo oscilatoria como una actividad repetitiva continua que representa un incremento de actividad de baja frecuencia relativa a el resto del registro EEG. El algoritmo se baso en el dominio de tiempo, en el cual se utilizó la función de autocorrelación para encontrar patrones de repetición en una señal.

El artefacto cardíaco fue muy dominante y se presentó totalmente en el registro de pocos pacientes. Se implementó un algoritmo de reducción de artefactos cardíacos [53], basado en análisis de componentes independientes y se aplicó al EEG antes de que el algoritmo de detección de crisis fuera ejecutado. El análisis se llevó a cabo en segmentos de 5 segundos. El método propuesto está basado en el Análisis de Componentes Independientes (ICA), una técnica de Separación Ciega de Fuentes (BSS, Blind Source Separation).

El paso más importante del algoritmo es la selección de los componentes ICA que corresponden al artefacto ECG. Para esto primero los autores proponen calcular la correlación directa entre los componentes ICA y el registro ECG, lo cual no es óptimo debido a que la morfología del artefacto ECG es diferente de la morfología de la medida ECG. Para abordar este problema los investigadores proponen calcular la correlación entre la energía thresholded del ECG y la energía thresholded de los componentes ICA. Para calcular la energía de estas señales, los autores utilizaron un operador de energía no lineal (NLEO).

Los investigadores probaron el método propuesto como un paso de pre-procesamiento durante la detección de crisis automatizada y se ejecutaron sobre ventanas de 5 segundos. Sin la reducción del ECG el EEG completo de ambos conjunto de datos fue

detectado falsamente por el detector de crisis debido al artificio ECG.

3.2.2. Grupo de Greene et al. (2008) [26]

La base de datos de registros de EEG fueron coleccionados del Hospital de Maternidad Unificado en Cork, Irlanda. Se contó con 17 registros multicanales, los cuales presentaban 411 eventos de crisis con una duración media de crisis de 4.06 minutos. La frecuencia de muestreo fue de 256 Hz, cada registro contiene de 7-11 canales EEG y un canal adicional correspondiente a ECG

En este grupo de trabajo se realizó un pre-procesamiento filtrando el registro EEG con un pasa-banda en el rango 0.1-34 Hz, mediante el modelo de Chebyshev de quinto orden. El análisis se realizó en el dominio de frecuencia y los artificios que se analizan son los causados por movimientos y por el amplificador. La longitud del segmento analizado se basó en muestras de 2048 datos (8 segundos), utilizando segmentos no traslapados.

En un estudio previo se determinaron las características que son altamente susceptibles a artificios mientras que otras fueron determinadas que contenían información redundante. Basado en dichas consideraciones, siete de las características fueron utilizadas como son: la entropía espectral, entropía de Shannon, Frecuencia de nodo espectral, Energía no lineal, Longitud de línea, Energía Wavelet y amplitud RMS.

Una variedad de artificios pueden presentarse en las señales EEG neonatal. Greene et al. trabajó con dos tipos de artificios: por movimiento, los cuales son espigas grandes causadas por los movimientos del paciente y señal cero, causado al encender el amplificador.

Para identificar las secciones de artificio del EEG, una señal EEG con media cero fue calculada restando la media del EEG de cada muestra y procesada como sigue:

1. Se calculó la desviación estándar del valor absoluto de una señal y cualquier muestra de la señal veinte veces más grande que la desviación estándar fue etiquetado como artificio de movimiento.
2. Cualquier época de 10 muestras cuya media fue cien veces más pequeña que el 5% de la media de la señal fue etiquetado como artificio de señal cero.

Se utilizaron tres técnicas, discriminante lineal, discriminante cuadrático y discriminante regularizado. Además, se aplicaron técnicas de extracción y selección de características, con el fin de reducir la dimensión de los datos. Finalmente, un paso de post-procesamiento fue implementado para reducir el número de detecciones falsas generadas por el sistema.

Una cuestión importante que origina el desarrollo de cualquier sistema de detección de crisis es el montaje de electrodos utilizado. Por esta razón los registros fueron digitalizados en un montaje referencial ($F3, F4, Cz, C3, C4, O1, O2, T3$ y $T4$). Se realizó una comparación para determinar si un montaje referencial era más adecuado para la aplicación de detección de crisis. Dicha comparación dió como resultado que un montaje referencial supera ligeramente a montajes tradicionales bipolares.

3.2.3. Grupo de Brotchie et al. (2007) [29]

Los investigadores de este grupo señalan que las Representaciones Tiempo Frecuencia (RTF), son herramientas populares para el análisis del espectro que varía con el tiempo de señales no estacionarias, tales como EEG de recién nacido. Observaron que los artificios de corto tiempo y gran amplitud tienen un efecto en patrones EEG, particularmente en patrones de crisis en EEG. Para crear artificios de tiempo corto y gran amplitud usaron el proceso estocástico simétrico estable de Lévy.

Para llevar a cabo las experimentaciones se utilizó un modelo de generación de crisis EEG neonatales de [45]. Para la simulación de artificios se utilizó el proceso estocástico simétrico estable de Lévy. La señal contaminada fue generada agregando a la señal sintética generada por [45] el ruido generado por el proceso de simulación de artificios.

Además de utilizar datos artificiales, se utilizaron datos reales proporcionadas por el Hospital de la Mujer y de Brisbane Royal localizados en Australia. Los datos EEG en crudo fueron filtrados por un pasa banda con cortes de frecuencia de 0.5 Hz y 70 Hz, además de un filtro de rechazo de banda a 50 Hz y digitalizados a una frecuencia de muestreo de 256 Hz. Una vez recuperados los registros EEG fueron nuevamente filtrados por un pasa banda con una frecuencia de corte de 0.5 Hz y 20 Hz, antes de bajar la frecuencia de muestreo a 40 Hz.

La clasificación de los registros EEG fue determinada por un especialista en

neurología pediátrica del Hospital Infantil Royal en Brisbane, Australia.

El análisis se realizó en segmentos de 6.4 segundos. Se demostró que el mejoramiento en RTF a través de análisis visual subjetivo. La mejora fue evaluada usando conocimiento a priori de las características de tiempo frecuencia de las crisis EEG en recién nacidos.

Además en este trabajo se usaron dos técnicas de detección de artificios: la Distribución de tiempo-frecuencia robusta (DTFR) y la Distribución B Modificada (DBM). Se realizaron experimentos, uno en el que se probó que DTFR es mejor que la Distribución de TF. Se realizó una comparación con una modificación de la DBM llamada Distribución B Modificada Robusta (DBMR), siendo ésta la que mejores resultados obtuvo.

3.2.4. Grupo de Mukherjee et al. (2003) [41]

La detección de artificios en EEG neonatales se realizó en este estudio entrenando las redes neuronales para identificar artificios y distinguirlos de actividad de crisis. Este estudio emplea redes neuronales hacia adelante y función coseno de base radial (FCBR). El trabajo se enfoca en detección de artificios causados por chupar o morder, los cuales son provocados por los movimientos de la mandíbula. Para la detección de estos artificios, se analizaron segmentos de 5 segundos con 2.5 segundos de traslape.

Los registros EEG utilizados fueron proporcionados por el Hospital Infantil de Texas en Houston, Texas, los cuales están compuestos por 12 electrodos ubicados en el cuero cabelludo de acuerdo al Sistema 10-20. Los registros EEG digitales fueron recuperados con una frecuencia de muestreo de 185 Hz.

Las redes neuronales entrenadas en artificios causados por morder o chupar fueron validadas en los registros de EEG de un sujeto que contenía 13 segmentos que representaban los artificios. El evento fue clasificado como un arteificio si la mayoría de los segmentos de candidatos a crisis a través de los 12 canales fueron clasificados como artificios en el mismo instante de tiempo. Sin embargo, la tasa más alta de detección de ubicación de los artificios indica que todos los 13 segmentos de artificios fueron clasificados correctamente como artificios.

3.3. Artificios oculares en adultos

En esta sección se describen algunos de los trabajos desarrollados en el área de la detección y eliminación de artificios oculares en registros EEG de adultos con la finalidad de analizar las técnicas empleadas en dichos trabajos de investigación.

3.3.1. Grupo de Gómez Herrero et al. (2006) [47]

Este grupo de trabajo presentaron un método para la eliminación de artificios oculares en el electroencefalograma. El procedimiento estuvo basado en separación ciega de fuentes, contrastándolo con los métodos que estaban disponibles en la literatura, es un método completamente automatizado y no requiere la disponibilidad de los electrodos EOG.

Las señales fueron capturadas de la salida de un arreglo de electrodos colocados en el cuero cabelludo, donde cada sensor recibía una combinación diferente de señales oculares y de señales neuronales. El problema fue modelado mediante el modelo matemático del ICA.

La corrección de los artificios EOG se realizó a través de un filtrado espacial, siguiendo los siguientes pasos:

1. Estimaron la matriz de mezcla A usando un conjunto finito de datos observados.
2. Identificaron las columnas de A correspondientes a artificios y los componentes EEG neuronales.
3. Recuperaron la actividad EEG neuronal libre de artificios por medio del filtrado espacial.

El algoritmo utilizado es SOBI (Identificación Ciega de Segundo Orden). Dicho algoritmo está basado en cantidades estadísticas de segundo orden lo cual hace la estimación más sencilla y robusta a errores. El primer paso del enfoque que propusieron consistió en realizar una descomposición SOBI de ventanas de EEG correlativas. Usando este análisis de ventanas el grupo de investigación logra trabajar con la naturaleza no estacionaria de las señales EEG.

El espectro de la actividad ocular es caracterizado por componentes de baja frecuencia. Estos componentes fueron fácilmente identificados por una dimensión fractal baja, la cual es una medida de la complejidad de una señal. Por otro lado, los trazos EEG fueron caracterizados por un espectro más plano y esparcido los cuales cuentan para dimensiones fractales más altas.

Los registros fueron grabados en la Unidad de monitoreo de Epilepsia (EMU). Incluyen las señales de 21 electrodos de acuerdo al Sistema 10-20 con electrodos adicionales $T1$ y $T2$ sobre la región temporal. La frecuencia de muestreo utilizada fue de 256 Hz y un montaje de referencia fue utilizado. Los registros EEG contienen actividad ictal de pacientes con epilepsia en lóbulo temporal.

La técnica de corrección de artificios oculares propuesto por este grupo fue aplicada a seis registros. En todos los casos, la mayoría de los artificios EOG fueron eliminados.

3.3.2. Grupo de Guerrero Mosquera et al. (2009) [40]

El algoritmo propuesto por este grupo de investigación se basó en ICA y mínimos cuadrados recursivos (RLS, por sus siglas en inglés). El algoritmo propuesto combina la efectividad de ICA para separar artificios de las ondas cerebrales, ambas con la interferencia alcanzada por los filtros adaptativos. El método utiliza electrodos separados localizados cerca de los ojos ($Fp1$, $Fp2$, $F7$ y $F8$), que registran los movimientos horizontal y vertical de los ojos, para extraer la señal de referencia.

El método para la eliminación de artificios propuesto comprende dos pasos. Primero, se aplicó ICA a los datos EEG y para los datos de referencia ($Fp1$, $Fp2$, $F7$ y $F8$). El segundo paso es el uso de cada componente ICA en un esquema de filtrado adaptativo.

Los registros utilizados fueron de tres pacientes usando el Sistema 10-20 con dos electrodos anterotemporal $T1/T2$. Los datos fueron registrados en el Hospital Universitario de Neurofisiología en Pamplona, España. Los datos en crudo fueron digitalizados con una frecuencia de muestreo de 200 Hz y fueron segmentados en ventanas de 5 segundos. Usando el sistema 10-20, los electrodos con mayor información de los movimientos de los ojos son $Fp1$ y $Fp2$ debido a que son localizados directamente arriba de los ojos. Los electrodos que registran el mayor cambio de potencial de los

ojos son registrados por $F7$ y $F8$ debido a que están cerca de los laterales de los ojos. Por lo tanto, estas señales fueron tomadas como señales de referencia.

Este grupo desarrolló un método para la cancelación de artificios usando registros EEG. Dicho método elimina eficientemente artificios producidos por los movimientos de los ojos. El método presentado esta basado en el Análisis de Componentes Independientes y un filtro adaptativo de los cuadrados mínimos recursivos. Los resultados mostraron que su método fue capaz de eliminar los artificios provocados por los movimientos de los ojos.

3.3.3. Grupo de Romero Lafuente et al. (2010) [47]

En esta tesis doctoral se estudian las diferentes situaciones de contaminación ocular en las señales EEG y su problemática bajo la situación de vigilia, mediante la exploración de bases de datos correspondientes a registros adquiridos en voluntarios sanos tras la administración de fármacos.

En el desarrollo de este trabajo de investigación se utilizaron dos bases de datos de señales EEG registradas tras la administración de fármacos. La primera base de datos comprende diferentes dosis de un fármaco ansiolítico (benzodiazepina), y la segunda evalúa el comportamiento de varios compuestos antipsicóticos clásicos y de nueva generación (neurolépticos). Ambas bases de datos se componen de registros de vigilia controlada en voluntarios sanos con ojos cerrados. Los datos provienen del Hospital de la Santa Creu i Sant Pau de Barcelona. Estas señales fueron registradas mediante 19 derivaciones de acuerdo al Sistema 10-20 ($Fp1$, $Fp2$, $F7$, $F3$, Fz , $F4$, $F8$, $T3$, $C3$, Cz , $C4$, $T4$, $T5$, $P3$, Pz , $P4$, $T6$, $O1$ y $O2$). Además se registraron dos canales de EOG: vertical y horizontal.

Las señales fueron digitalizadas con una frecuencia de muestreo de 100 Hz. Posteriormente fueron filtradas por un pasa bajo y pasa alto con frecuencias de corte de 0.5 Hz y 45 Hz, respectivamente. Además se utilizaron señales simuladas a partir de la mezcla de las actividades ocular y cerebral mediante el uso de dos estrategias: mezcla instantánea y mezcla convolutiva. La diferencia principal entre ambas radica en el modelo de propagación de las actividades bioeléctricas: independiente de la frecuencia para el caso instantáneo y dependiente para el convolutivo.

La metodología que fue propuesta se basó en ICA y los pasos se enumeran a

continuación:

1. Descomposición en señales fuente o componentes de las señales EOG y EEG registradas.
2. Identificación de las componentes fuente asociadas con artificios, en este caso de origen ocular.
3. Estimación de las señales EEG corregidas mediante la reconstrucción o deflación a partir del modelo de generación inverso sin tener en cuenta las señales fuente asociadas con actividad no cerebral.

La eficacia del método automático de filtrado ocular que propusieron ha sido validada tanto en señales simuladas como en señales reales, demostrando un mejor comportamiento al método considerado como el mejor hasta ese momento: Regresión lineal.

3.4. Comentarios acerca del estado del arte

En el Cuadro 3.1, se muestra una comparación entre los diferentes trabajos realizados en esta área que es la detección y/o eliminación de crisis en registros EEG neonatales. En ella se puede observar que cada grupo de investigadores obtienen muestras de los datos a diferentes frecuencias, lo cual indica que actualmente no hay una regla de muestreo de los registros. Además algunos de los registros fueron pre-procesados utilizando filtros para depurar en un rango de frecuencias que los autores propusieron.

Otro aspecto importante que cabe mencionar, es que todos los autores trabajan con diferentes tipos de crisis, por lo cual no sería posible realizar una comparación entre las técnicas utilizadas, debido a que cada tipo de crisis tiene diferentes características, con las cuales fueron analizadas. Además se identificó que aspectos como análisis de montajes y post-procesamientos no se han realizado en el área de detección de crisis.

En el Cuadro 3.2, se muestra una comparación entre los diferentes trabajos realizados en esta área que es la detección y/o eliminación de artificios en registros EEG

neonatales. En esta área se puede observar que existe poco esfuerzo de investigación, ya que sólo se tienen identificado cuatro trabajos disponibles que realizan detección y/o eliminación de artificios en registros EEG neonatales.

La mayoría de los trabajos presentados en la tabla muestrean los datos a 256 Hz, y sólo uno lo hace a 185 Hz, lo cual indica que no existe una regla para el muestreo de los datos. Otro punto que señalar es que tres de los trabajos realizan un pre-procesamiento a los registros antes de aplicar las técnicas necesarias para la detección y/o eliminación de los artificios. Dicho pre-procesamiento lo realizan aplicando filtros para la depuración de frecuencias que se encuentren en el rango especificado.

Finalmente se presentó algunos de los trabajos de la literatura, los cuales abordan la detección y eliminación de artificios oculares EEG de adultos.

La frecuencia de muestreo en el tratamiento de registros EEG adultos así como la de neonatos no tiene un estándar, es decir, dicha frecuencia es proporcionada por los especialistas que llevan a cabo el registro de las señales cerebrales.

Otro aspecto relevante en estos trabajos, es que no se tiene una regla para definir con cuales señales de acuerdo al Sistema 10-20 se van a utilizar para realizar el análisis de artificios oculares. Una similitud entre los trabajos presentados en esta tesis es que en la detección y eliminación de artificios oculares se llevó a cabo analizando las zonas frontales del cerebro, es decir, los electrodos ubicados en la parte frontal.

El análisis y tratamiento de los registros EEG de adultos fueron aplicados a registros EEG reales y simulados. De los trabajos presentados, Romero Lafuente et al. realiza una etapa de pre-procesamiento de los registros realizando un filtrado pasa bajo y pasa alto con una frecuencia de corte de 0.5 Hz y de 45 Hz, respectivamente.

Asimismo, la evaluación de la metodología propuesta se ha llevado a cabo de acuerdo al criterio subjetivo de cada uno de los grupos de investigación, la evaluación es llevada a cabo por un especialista en el ámbito de neurología.

Con el análisis de estos trabajos se pudo determinar que algoritmos ICA basados en BSS dan buenos resultados en el problema abordado. Por esta razón se decidió abordar el problema de la detección y eliminación de artificios oculares en registros EEG neonatales mediante la aplicación de un algoritmo ICA conocido como INFOMAX.

GPO./CARACTERÍSTICA	Frecuencia de muestreo	Pre-procesamiento	Dominio	Tipos de crisis	Técnicas	Extracción y selección de características	Tamaño Muestra	Canal
Kok-kiong et. al 2007	No especificado	No	Tiempo-Frecuencia	Ondas bajas, picos afilados	Transformada Stockwell	No	Segmentos	No especificado
Piryatinska et. al 2009	64 Hz	Pasa-banda anti-aliasing	Frecuencia	Estado del sueño	Cambios en los puntos y análisis de grupos	Si	Segmentos de 30 seg	Simple
Fault et. al 2007	200 Hz	Se normalizaron entre 0-1, y se procesaron a 66 Hz	Frecuencia	Descargas rítmicas, espigas, puntas afiladas y ondas bajas	Modelo Gaussiano	No	Segmentos de 1 seg	Multicanal
Shellhaas et. al 2007	No especificado	No	Tiempo	Patrones repetitivos y repentinos	Caracterización espacio-temporal	No	Registro	Simple
Zarjam et. al 2007	256 Hz	Pasa-banda de 0.5-70 Hz	Tiempo-Frecuencia	Crisis epiléptica	Transformada Wavelet Discreta	Si	Segmentos de 6 seg	Multicanal
Kok-kiong et. al 2008	200 Hz	No	Tiempo	Ondas bajas, picos afilados	Tasa finita de innovación	No	Segmentos de 1024 datos	Multicanal
Boashash et. al 2001	256 Hz	No	Tiempo-Frecuencia	Picos afilados y repetitivos	Análisis tiempo-frecuencia	No	Segmentos de 127 datos	Multicanal
Wong Lisa et. al 2008	No especificado	No	Tiempo-Frecuencia	Crisis	Tasa de verosimilitud generalizada	No	10 segmentos de 1 min y 2 de 20 seg	2 canales

Cuadro 3.1: Comparación de los trabajos respecto a crisis.

GPO/CARACTERÍSTICA	Frecuencia de muestreo	Pre-procesamiento	Dominio	Tipos de artificios	Técnicas	Extracción y selección de características	Tamaño Muestra	Canales	Montajes	Post-procesamiento
Deburghraeve et. al. 2008	256 Hz	Pasa-banda de 0.3-30 Hz	Tiempo	Artificio de ECG y crisis Tren de espigas y oscilatorias	ICA y algoritmos de Repetitividad y auto-correlación	No	5 seg	Multicanal con 14 reg. Con 17 electrodos; y 12 reg. Con 13 electrodos	No	No
Greene et. al. 2008	256 Hz	Pasa-banda 0.1-34 Hz	Frecuencia	Artificios de movimiento y por el amplificador	Discriminante regularizado	Si	8 seg	multicanal	Referencial	Si
Brotchie et. al. 2007	256 Hz	Pasa-banda de 0.5-20 Hz y después se muestrearon a 40 Hz	Tiempo-Frecuencia	Artificios de alta amplitud y tiempo corto	Distribución de Tiempo-Frecuencia robusto	No	6.4 seg	No especificado	No	No
Mukherjee et. al. 2003	185 Hz	No	Tiempo	Artificios por morder o chupar	Redes neuronales	Si	5 seg	multicanal	No	No

Cuadro 3.2: Comparación de trabajos con respecto a artificios.

Capítulo 4

Técnicas para el análisis de EEG

Generalmente, en los problemas de clasificación, las características que proporcionan información relevante son desconocidas, y la utilización de características que no son de importancia de acuerdo con el objetivo, genera un aumento en el tiempo de aprendizaje de los algoritmos. Los métodos para reducir la dimensionalidad tienen como objetivo extraer características que son relevantes para la clasificación planteada.

Un conjunto de datos de dimensión p se puede reducir a uno de q dimensiones, utilizando p combinaciones lineales de q variables. Las combinaciones lineales pueden ser consideradas como proyecciones lineales. Entre ellos se encuentra el Análisis de Componentes Principales (por sus siglas en inglés, Principal Component Analysis, PCA) y el Análisis de componentes Independientes [31].

4.1. Análisis de Componentes Principales (PCA)

El Análisis de Componentes Principales comprende un procedimiento matemático que transforma un conjunto de variables correlacionadas de respuesta en un conjunto menor de variables no correlacionadas las cuales son llamadas componentes principales.

En el proceso se pueden formar nuevas variables en orden decreciente de importancia, de modo que [16]:

1. No estén correlacionadas

2. La primera componente principal explique tanto de la variabilidad en los datos como sea posible
3. Cada componente siguiente tome en cuenta tanto de la variabilidad restante como sea posible

4.1.1. Definición de PCA sobre la matriz de covarianzas

Sea x el conjunto de datos y μ su media. La j -ésima ($j = 1, 2, \dots, p$) variable componente principal se define por $y_j = a_j'(x - \mu)$, en donde a_j se elige de modo que $a_j'a_j = 1$ y en tal forma que la varianza de $a_j'(x - \mu)$ sea un máximo entre todas esas combinaciones lineales de x que no estén correlacionadas con las $j - 1$ componentes principales. Se tiene que se maximiza la varianza cuando $a_j'a_j = 1$ es igual a λ_j , el j -ésimo eigenvalor más grande de la matriz de covarianzas Σ y ocurre cuando a_j es un eigenvector correspondiente a λ_j , y que satisface $a_j'a_j = 1$.

4.1.2. Calificación de los componentes principales

En el desarrollo precedente se supuso que se conocía tanto μ como Σ . Difícilmente ocurrirá esto, de modo que μ y Σ necesitarán estimarse a partir de los datos de la muestra.

Suponga que se tiene una muestra aleatoria de una población con media μ y matriz de covarianzas Σ . Entonces los estimadores de μ y Σ son $\hat{\mu}$ y $\hat{\Sigma}$, respectivamente.

Las calificaciones de cada componente principal j para la r -ésima unidad experimental, se estiman por:

$$y_{rj} = \hat{a}_j'(x_r - \hat{\mu}), \text{ para } j = 1, 2, \dots, p; r = 1, 2, \dots, N \quad (4.1)$$

Siendo N el número de unidades experimentales de la muestra y p el número de variables.

4.1.3. Determinación del número de componentes principales

De manera general, una forma muy sencilla de determinar el número de componentes principales es determinar cuántos eigenvalores son mayores a uno y éste será el número de componentes, reduciendo de ésta manera la dimension de los datos. Sin embargo, para el análisis de registros EEG el número de nuevas variables y componentes principales debe ser igual al número de canales, por lo cual no se reduce la dimensión de los datos, solo es una preparación de los datos para la aplicación del Análisis de Componentes Independientes, con la finalidad de eliminar la correlación de las variables.

4.2. Análisis de Componentes Independientes

El objetivo del análisis de componentes independientes (Independent Component Analysis, ICA) es generar una transformación de los datos tal que el resultado esté formado por atributos que son estadísticamente independientes. En PCA los atributos no son estadísticamente independientes si no que solo lo son respecto a los estadísticos de segundo orden (esto es que los componentes sean no correlacionados). Por esta razón ICA es más general que PCA, pero evidentemente más complejo.

El análisis de componentes independiente es un método para encontrar factores o componentes de datos estadísticos multivariado. La diferencia entre ICA y otros métodos es que en ICA los componentes son estadísticamente independientes, y no Gaussianos [4].

ICA es un método de separación de mezcla el cual ha sido utilizado en análisis de señales biomédicas. Dada m mezclas de n señales estadísticamente independientes, ICA puede separar cada una de las fuentes individuales. Este problema ha sido llamado Separación Ciega de Fuentes (BSS, Blind Source Separation) [48], el cual consiste en recuperar las señales de las fuentes independientes generadas por x fuentes, teniendo tan sólo acceso a las señales mezcladas y . La ponderación de la mezcla A es totalmente desconocida, y representa al medio en donde se propagan dichas señales [38].

El problema de Separación Ciega de Fuentes surge a partir de la necesidad de

recuperar, desde una mezcla de señales, las fuentes originales que la componen. Dicha extracción se caracteriza como ciega al no disponer de información tanto de la caracterización de las fuentes como del número de ellas presentes en el momento de la mezcla, ni del modo en que ésta ha sido llevada a cabo. Aunque este tipo de información sea desconocida, desde el punto de vista matemático existe solución al problema planteado siempre que las señales de entrada al proceso sean estadísticamente independientes. La idea es sencilla: Si los vectores observados son una combinación lineal de componentes independientes, estas componentes deberían poder ser separadas.

4.2.1. Historia

Hérault y Jutten parecen ser los primeros en utilizar de manera informal el concepto de ICA alrededor del año 1983, especialmente para resolver el problema de separación ciega de fuentes (BSS, blind source separation). Pocos años después, Common presentó una formulación matemática de ICA y mostró como los cumulantes de orden mayor pueden ser usados para resolver el problema de ICA. Al mismo tiempo, Cardoso y Souloumiac desarrollaron el algoritmo JADE basado en la diagonalización aproximada conjunta. Mientras estos dos enfoques usan estadísticos de segundo y cuarto orden, otros enfoques intentan explotar estadísticos de segundo orden únicamente. Fety es el primero en utilizar las matrices de covarianza, el mismo tipo de enfoque es desarrollado independientemente varios años más tarde por Tong, Belouchrani et al., y Ziehe and Müller, quienes da lugar a los métodos llamados AMUSE, SOBI y TDSEP, respectivamente.

En 1999, Müller et al. propusieron una versión modificada de JADE, la cual usa estadísticos de segundo y cuarto orden. Más recientemente Albera et al. presentaron una extensión de SOBI para estadísticos de cuarto orden, denominado FOBIUM, conduciendo a ICA especialmente bajo determinados contextos (más componentes que observaciones). Algunos autores también propusieron un método algebraico, llamado ICAR, el cual usa la matriz de redundancias del estadístico de cuarto orden de la matriz de covarianza, mejor conocida como matriz cuadri-covarianza. Como señaló Parra y Sajda, bajo el supuesto de que las fuentes son no estacionarias, no blanqueadas y no gaussianas, ICA puede ser reformulado como un problema de eigenvalor generalizado. Mientras los métodos previos identifican simultáneamente los

componentes independientes, Delfosse y Loubaton propusieron extraer un componente a la vez, el cual es llamado procedimiento de deflación. Años más tarde, Hyvarinen et al. propusieron el método FastICA, el cual iterativamente maximiza un contraste del estadístico de cuarto orden. Mientras la primera versión de este algoritmo es de tipo deflación, como el de Delfosse y Loubaton, Hyvarinen et al. propusieron después una versión simultánea de FastICA cuyo paso de ortogonalización conjunta es similar a una presentada por el autor Moreau.

En lugar de aprovechar, explícita o implícitamente, los estadísticos de segundo y cuarto orden para resolver el problema de ICA, algunos enfoques usan directamente la hipótesis de independencia. De hecho, Lee et al. presentó un enfoque de maximización de información basado en distribuciones de probabilidad parametrizadas que tienen distribuciones super y sub-gaussianas para derivar la regla de aprendizaje general, la cual es optimizada usando un algoritmo de gradiente natural propuesto por Amari et al.

Pham propuso usar estimados no paramétricos de verosimilitud o de la información mutua. Otro algoritmo, basado en la minimización de un estimador no paramétrico de la información mutua de Renyi's como un criterio para ICA es introducido por Erdogmus et al. [32].

4.2.2. Modelo ICA

El problema abordado en esta tesis consiste en separar y estimar las múltiples fuentes de señal de una mezcla, a través de un arreglo discreto de electrodos, sin un previo conocimiento de las características de las fuentes originales y de los canales de transmisión.

Como punto de partida, se supone un número igual de fuentes desconocidas y de electrodos, en donde cada electrodo por sus características y su posición, reciben una combinación lineal diferente de cada una de las señales. En la figura 4.1 se muestra un esquema del arreglo de electrodos y de fuentes originales de tamaño $n \times n$.

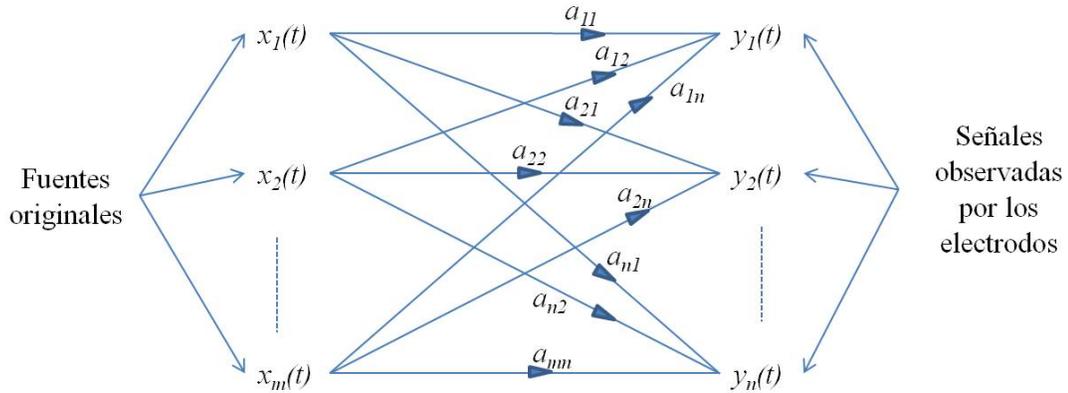


Figura 4.1: Arreglo de electrodos $y(t)$ y fuentes originales $x(t)$

El sistema mostrado es descrito por la siguiente ecuación:

$$y_i(t) = \sum_{j=1}^n a_{ij} s_j(t) \quad i, j = 1, 2, \dots, n \quad (4.2)$$

En donde: $y_i(t)$ es la señal continua en el tiempo, de la salida del i -ésimo electrodo, $s_j(t)$ es la señal desconocida, emitida por la j -ésima fuente y a_{ij} es un número escalar de la combinación lineal.

Escrito en forma matricial, la ecuación 4.2 se expresa como:

$$\mathbf{y}(t) = \mathbf{A} \times \mathbf{s}(t) \quad (4.3)$$

En donde \mathbf{A} representa la matriz de mezcla de tamaño $n \times n$. ICA es una técnica estadística de procesamiento de señal cuyo objetivo consiste en expresar un conjunto de variables aleatorias como una combinación lineal de variables estadísticamente independientes [46].

Dicho en otras palabras, ICA considera m señales mezcladas $y_1(t), y_2(t), \dots, y_m(t)$ obtenidas de la combinación lineal de n señales independientes $x_1(t), x_2(t), \dots, x_n(t)$, donde $n \geq m$.

Una manera de formular la problemática de ICA es mediante la estimación paramétrica del siguiente modelo de generación de datos. Típicamente, el problema de ICA puede ser descrito como sigue:

Suponga que $\mathbf{X} = [x_1 x_2 \dots x_n]^T$ es la señal multivariada desconocida original,

$\mathbf{Y} = [y_1 y_2 \dots y_m]^T$ es la señal observada y es transformada a través de una matriz de mezcla lineal \mathbf{A}_{mn} , tal que $y = \mathbf{A} \times \mathbf{x}$, donde:

$$\begin{aligned}
 y_1(t) &= a_{11}x_1(t) + a_{12}x_2(t) + a_{13}x_3(t) + \dots + a_{1n}x_n(t) \\
 y_2(t) &= a_{21}x_1(t) + a_{22}x_2(t) + a_{23}x_3(t) + \dots + a_{2n}x_n(t) \\
 y_3(t) &= a_{31}x_1(t) + a_{32}x_2(t) + a_{33}x_3(t) + \dots + a_{3n}x_n(t) \\
 &\dots \\
 y_m(t) &= a_{m1}x_1(t) + a_{m2}x_2(t) + a_{m3}x_3(t) + \dots + a_{mn}x_n(t)
 \end{aligned}
 \tag{4.4}$$

El objetivo de ICA es encontrar la matriz de mezcla lineal \mathbf{A} para adquirir los componentes independientes \mathbf{x} bajo las condiciones de criterios de independencia.

La función de \mathbf{A} es tal que $\mathbf{x}' = \mathbf{A} \times \mathbf{y}$, donde \mathbf{x}' es la aproximación de \mathbf{x} [17].

En la figura 4.2 se muestran dos ejemplos de señales de fuentes independientes (originales) con distribución uniforme. En la parte superior derecha de la imagen se observan los datos en un espacio bidimensional (sin la variable tiempo) mientras que en la izquierda y en la parte inferior, se muestra la distribución de cada una de las variables independientes (histogramas), así como su representación en el tiempo.

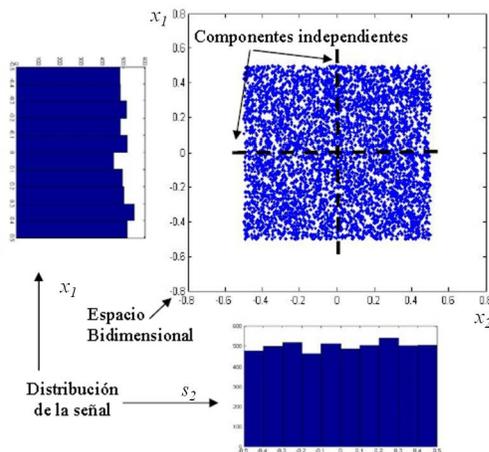


Figura 4.2: Ejemplo de componentes independientes x_1 y x_2 con distribución uniforme

Considerando la expresión 4.3 y asumiendo una matriz de mezcla \mathbf{A} , el resultado de la variable $\mathbf{y}(t)$ se muestra en la figura 4.3. En la parte superior derecha de la imagen son mostrados los datos en un plano bidimensional. A la izquierda y en la parte inferior se muestran sus distribuciones correspondientes.

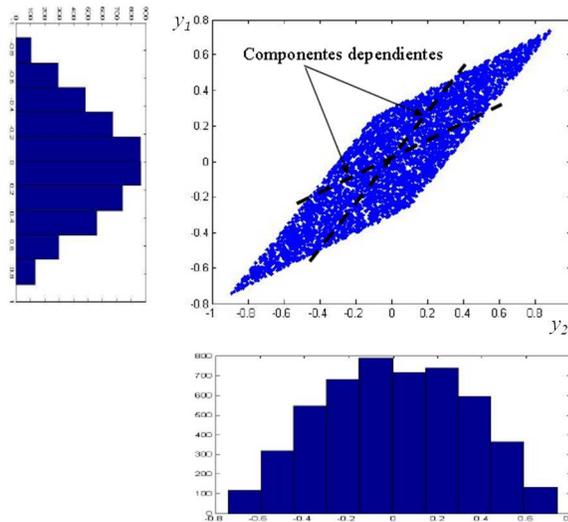


Figura 4.3: Mezclas de las señales independientes x_1 y x_2

El problema consiste en encontrar una transformación que encuentre la inversa de la matriz con la cual se mezcló. Para obtener una solución al problema, se debe cumplir lo siguiente: $\mathbf{A} \in \mathbf{R}^{n \times n}$ y $\text{Det}(\mathbf{A}) \neq 0$ (\mathbf{A} no debe ser singular o casi singular). En este sistema solamente las señales de salida $\mathbf{y}(t)$ de cada electrodo están disponibles. La matriz de mezcla \mathbf{A} y las señales de fuentes originales $\mathbf{x}(t)$ son desconocidas.

Aplicar un método convencional (como el de Gauss) para encontrar la matriz de mezcla en este problema es imposible, ya que no se cuenta con más información que las señales adquiridas por los electrodos. Sin embargo, una aproximación para resolver este problema es utilizar las propiedades estadísticas de las señales. El modelo de mezcla considerado en este problema de investigación es que tanto el medio como los sensores se comportan de forma lineal.

4.2.3. Procesamiento inicial

Antes de que las matrices de mezclas y de fuentes sean calculadas, un procesamiento inicial debe de ser aplicado a los datos observados. Estos pasos reducen la complejidad del problema ICA mientras se mantienen las características benéficas [48].

* **Centrado de los datos**

El primer paso es centrar la matriz de mezcla observada. Este paso se hace solamente para simplificar el proceso ICA y no afecta su resultado. El centrado es llevado a cabo simplemente eliminando la media de cada mezcla observada e implica que las fuentes son también centradas. La media puede ser regresada fácilmente a los datos después de aplicar el algoritmo ICA.

* **Blanqueado de los datos**

El siguiente paso es el blanqueado de los datos y es aplicado después del proceso de centrado y antes de aplicar el algoritmo ICA. El procedimiento de blanqueado es una transformación lineal de los datos a un nuevo vector de datos compuesto de elementos los cuales son no correlacionados y con varianza la unidad.

LA etapa de procesamiento inicial de ICA puede ser llevada a cabo aplicando PCA. En la ecuación de PCA 4.1 el centrado de los datos es el segundo término y el blanqueado es el resultado obtenido de la aplicación de PCA.

4.2.4. Restricciones

- * Los componentes independientes deben ser estadísticamente independientes. Es decir, que las variables x_1, x_2, \dots, x_m son independientes, si y solo si, el valor de una variable no da información acerca del valor de la otra.
- * Los componentes independientes deben tener distribuciones no Gaussianas, para validar esta restricción se pueden utilizar medidas de no Gaussianidad.
- * Por simplicidad se asume una matriz de mezcla \mathbf{A} , como matriz cuadrada.

* Este método no da información alguna sobre la varianza de la señal original ni del signo de la misma, sin embargo, este no es un problema de trascendencia; en consecuencia las características iniciales de y_n se estandarizan de modo que su varianza sea unitaria [31].

4.2.5. Independencia estadística

Técnicamente la independencia estadística puede ser definida por la probabilidad. Sea $p(x_1, x_2, \dots, x_m)$ la función de probabilidad conjunta de x_i , y $p_i(x_i)$ la función de probabilidad marginal de x_i . Entonces se dice que los x_i son independientes si y solo si la función de probabilidad conjunta se puede factorizar de la siguiente manera [4]:

$$p(x_1, x_2, \dots, x_m) = p_1(x_1)p_2(x_2) \dots p_m(x_m) \quad (4.5)$$

4.2.6. Medida de no Gaussianidad

Para usar la no gaussianidad como herramienta para calcular las fuentes originales, una medida útil debe ser elegida. La medida clásica de Gaussianidad es la Kurtosis:

$$kurt(x) = \frac{E(x - \mu)^4}{\sigma^4} \quad (4.6)$$

Donde μ es la media de \mathbf{x} y σ es la desviación estándar. El valor absoluto de la kurtosis es comúnmente usado como medida de gaussianidad. Este cálculo de la kurtosis es muy sencillo y ampliamente utilizado para propósitos estadísticos. Sin embargo, la kurtosis no provee una medida robusta de gaussianidad debido a su sensibilidad a datos aberrantes en la base de datos, esto es debido al uso de la media. Sin embargo hay otras medidas que favorecen a ICA, como puede ser la entropía.

La entropía $H(x)$, de una variable aleatoria está dada por:

$$H(x) = - \sum_{\infty} P(x) \log P(x) \quad (4.7)$$

Donde $P(x)$ es la probabilidad de que \mathbf{x} tome el valor de x y pueda ser vista como

una medida de previsibilidad la cual una señal exhibe. Una variable gaussiana tiene la entropía más grande a lo largo de todas las variables aleatorias con igual varianza y sin embargo la entropía puede ser usada como una medida de gaussianidad.

Para obtener una medida la cual es siempre no negativa y es igual a cero solamente para variables Gaussianas se utiliza la Negentropía, $J(x)$:

$$J(x) = H(x_{gauss}) - H(x) \quad (4.8)$$

Donde $H(x_{gauss})$ es la entropía de la variable x_{gauss} , la cual es una variable aleatoria Gaussiana con la misma matriz de covarianza de \mathbf{x} . La desventaja de usar esta medida con respecto a la kurtosis es que es más costosa computacionalmente [48].

4.3. ICA basado en la maximización de la información

INFOMAX es un algoritmo ICA basado en la maximización de la información. INFOMAX está basado en un gradiente que se usa para maximizar la entropía. Sin embargo, es una tarea difícil, debido a que la entropía de las mediciones es desconocida y por lo tanto nada puede ser hecho con este en la maximización. En esta situación un enfoque de sub-optimización fue propuesto, en el cual el gradiente fue evaluado basado solamente en los términos que no corresponden a la entropía de las mediciones [55].

Este método se basa en la medida de la entropía relativa de dos o más distribuciones de probabilidad de las señales. En la figura 4.4 se muestra la arquitectura de la red neuronal, en la que se efectúa la maximización de la transferencia de información y la minimización de la información mutua entre sus salidas [35].

Donde cada entrada $y_i(t)$ esta formada por la mezcla de las señales, cada salida $u_i(t)$ es la señal estimada de la fuente independiente y $g(u)$ es la función de transferencia o de activación de la red neuronal, en el caso de INFOMAX se utiliza una función logística ($\tanh(u)$). El objetivo es encontrar una matriz \mathbf{W} inversa a la matriz de mezcla \mathbf{A} , que recupere las señales de las fuentes independientes.

La razón de información sobre el cual se basa INFOMAX es la entropía y se

encuentra definida como:

$$H(X) = \sum_{x \in A_x} P(x) \log \frac{1}{P(x)} \quad (4.9)$$

Donde \mathbf{X} es el conjunto de variables aleatorias de x_i . La regla de aprendizaje para obtener la matriz \mathbf{W} se deriva de la maximización de la entropía de las salidas de la red, además $m = n$, es decir, se tiene el mismo número de electrodos y de señales de salida. Esto fue propuesto por Bell y Sejnowski [8].

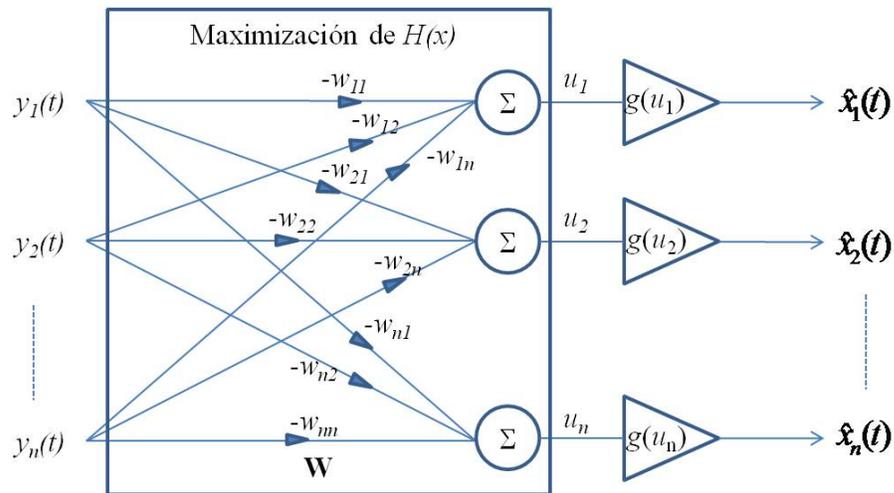


Figura 4.4: Arquitectura de la red neuronal

Bell y Sejnowski muestran que maximizando la entropía conjunta $H(\hat{x})$ de la salida de una red neuronal aproximadamente se puede minimizar la cantidad de información mutua de las salidas de los componentes $x_i = g_i(u_i)$, donde $g_i(u_i)$ es una función no lineal e invertible y $\mathbf{u} = \mathbf{W} \times \mathbf{x}$.

La entropía conjunta de las salidas de la red es:

$$H(\hat{x}_1, \dots, \hat{x}_n) = H(\hat{x}_1) + \dots + H(\hat{x}_n) - I(\hat{x}_1, \dots, \hat{x}_n) \quad (4.10)$$

donde $H(\hat{x}_i)$ es la entropía marginal de las salidas e $I(\hat{x}_1, \dots, \hat{x}_n)$ es la información mutua entre sus salidas. Maximizar la entropía conjunta equivale a maximizar las entropías marginales y a minimizar la información mutua.

El valor máximo para la entropía conjunta se alcanza cuando la información mutua llega a ser cero, es decir que no contienen información acerca de la otra salida. Bell y Sejnowski eligieron una función no lineal $g(u)$ (en este caso se usa una función logística), de la cual su derivada es una función de distribución de probabilidad similar a una función *super-gaussiana*, gaussiana con distribución más concentrada en la media y más dispersa en los extremos, asumiendo que las fuentes originales tienen distribuciones super-gaussianas. Hay dos parámetros que determinan la máxima entropía conjunta: la no linealidad $y_i = g_i(u_i)$ y la efectividad de \mathbf{W} .

El valor óptimo de los pesos sinápticos se puede encontrar maximizando la entropía conjunta con respecto a \mathbf{W} . La derivada de la ecuación D.2 con respecto a \mathbf{W} se relaciona con la divergencia Kullback-Leibler (ver Apéndice D) entre el estimado de la distribución de la fuente $p(\hat{x})$ y la fuente $p(x)$.

$$\frac{\partial H(\hat{x})}{\partial \mathbf{W}} = \frac{\partial}{\partial \mathbf{W}} (-D(p(x)||p(\hat{x}))) \tag{4.11}$$

Cuando la función de transferencia $g_i(u_i)$ y \mathbf{W} son optimizados la entropía conjunta $H(\hat{x})$ es máxima y el estimado de la distribución de la fuente y la fuente son iguales, es decir, $p(x) = p(\hat{x})$. Puesto que la función no lineal $g_i(u_i)$ a la salida de la red no introduce dependencia entre las señales, la información mutua seguirá siendo igual a cero $I(\hat{x}_1, \hat{x}_2, \dots, \hat{x}_n) = 0$.

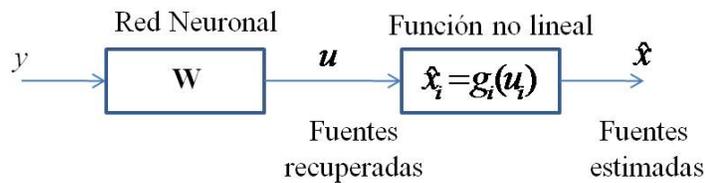


Figura 4.5: Método de la máxima entropía

De acuerdo a la figura 4.5, la función de transferencia entre \hat{x}_i y u_i se denota como:

$$\frac{p(u_i)}{p(\hat{x}_i)} = \left| \frac{\partial g_i(u_i)}{\partial u_i} \right| \tag{4.12}$$

La relación entre \hat{x}_i , u_i y la función de transferencia no lineal es:

$$p(\hat{x}_i) = \frac{p(u_i)}{\left| \frac{\partial g_i(u_i)}{\partial u_i} \right|} \quad (4.13)$$

Debido a que las fuentes independientes tienen una distribución uniforme de \hat{x}_i , se tiene:

$$p(u_i) = \left| \frac{\partial g_i(u_i)}{\partial u_i} \right| \quad (4.14)$$

Para la ecuación 4.14 se asume que u_i es una variable independiente con una distribución que es aproximadamente la forma de la derivada de la función no lineal $g(u)$. En este caso la función utilizada es una función logística. En la figura 4.6 se puede observar el comportamiento de una función logística y su derivada.

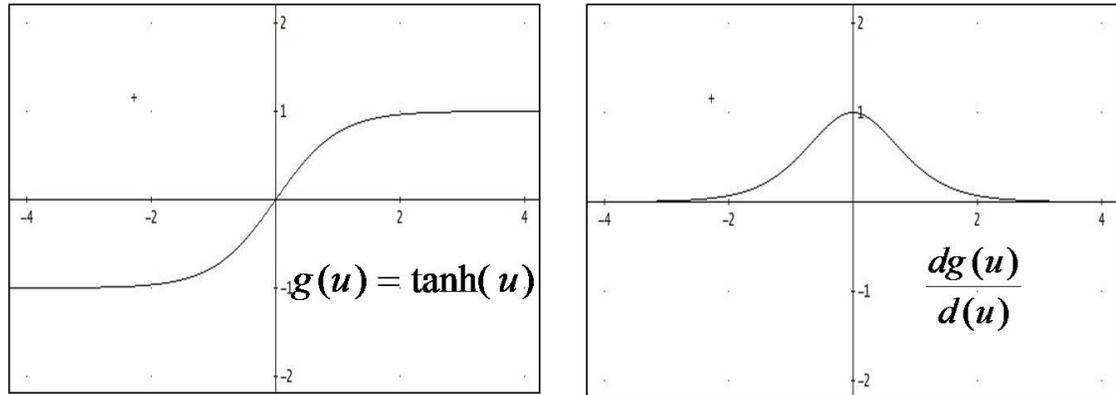


Figura 4.6: Función logística y su derivada

La función no lineal es empleada principalmente para minimizar la información mutua y así obtener salidas independientes. La independencia estadística se obtiene cuando la información mutua de la red es cero y la entropía conjunta es igual a la suma de las entropías marginales, donde cada entropía marginal se representa como:

$$H(\hat{x}_i) = -E\{\log p(\hat{x}_i)\} \quad (4.15)$$

La función de no linealidad realiza una transformación de la variable u_i a \hat{x}_i , por lo tanto la ecuación 4.13 se puede expresar como:

$$p(\hat{x}_i) = \frac{p(u_i)}{\left| \frac{\partial \hat{x}_i}{\partial u_i} \right|} \quad (4.16)$$

sustituyendo en la ecuación 4.15, la entropía marginal se obtiene:

$$H(\hat{x}_i) = -E\left\{\log\left(\frac{p(u_i)}{\left| \frac{\partial \hat{x}_i}{\partial u_i} \right|}\right)\right\} \quad (4.17)$$

Reescribiendo en la ecuación D.2 se obtiene:

$$H(\hat{x}_i) = -E\left\{\log\left(\frac{p(u_1)}{\left| \frac{\partial \hat{x}_1}{\partial u_1} \right|}\right)\right\} + \dots - E\left\{\log\left(\frac{p(u_i)}{\left| \frac{\partial \hat{x}_i}{\partial u_i} \right|}\right)\right\} - I(\hat{x}) \quad (4.18)$$

$$H(\hat{x}_i) = -\sum_{i=1}^n E\left\{\log\left(\frac{p(u_i)}{\left| \frac{\partial \hat{x}_i}{\partial u_i} \right|}\right)\right\} - I(\hat{x}) \quad (4.19)$$

Realizando la derivada de la entropía conjunta con respecto a \mathbf{W} se obtiene:

$$\frac{\partial H(\hat{x}_i)}{\partial \mathbf{W}} = \frac{\partial}{\partial \mathbf{W}}(-I(\hat{x})) - \frac{\partial}{\partial \mathbf{W}}\left(\sum_{i=1}^n E\left\{\log\left(\frac{p(u_i)}{\left| \frac{\partial \hat{x}_i}{\partial u_i} \right|}\right)\right\}\right) \quad (4.20)$$

En esta ecuación se observa la relación entre la entropía conjunta y la información mutua. La minimización directa de la información mutua genera la maximización de la entropía conjunta. Dicho de otra manera, la información mutua será minimizada ($I(\hat{x}) \cong 0$) cuando la no linealidad $\hat{x}_i = g_i(u_i)$ es aproximada a la función de densidad acumulativa de la fuente estimada $p(u_i)$.

En caso que la fuente estimada y la función no lineal resulten un valor diferente de cero para $I(\hat{x})$, esto indicará que se ha cometido un error. Por lo tanto, la ecuación 4.20 se plantea que el proceso de minimización de $I(\hat{x})$ depende de las fuentes estimadas y de la función no lineal empleada. El término de error se puede considerar despreciable si se utiliza una función logística y señales con distribución super gaussianas. En este caso, el término del error desaparece y el máximo de la entropía conjunta se obtiene derivando la entropía $H(\hat{x}_i)$ con respecto a \mathbf{W} . Lo anterior se realiza calculando el gradiente de $H(\hat{x}_i)$.

$$\frac{\partial H(\hat{x}_i)}{\partial \mathbf{W}} = \frac{\partial}{\partial \mathbf{W}} (-E\{\log(|J|)\}) = \frac{\partial}{\partial \mathbf{W}} \log(|J|) \quad (4.21)$$

El término $|J|$ es el valor absoluta del Jacobiano. Donde la relación entre la densidad de salida $p(\hat{x})$ y la densidad de entrada $p(y)$ puede ser definida como $p(\hat{x}) = p(y)/|J(y)|$. Por lo tanto $|J|$ puede ser definido como una transformación de $p(y)$ a $p(\hat{x})$:

$$J(y) = \det \begin{bmatrix} \frac{\partial \hat{x}_1}{\partial y_1} & \cdots & \frac{\partial \hat{x}_1}{\partial y_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial \hat{x}_n}{\partial y_1} & \cdots & \frac{\partial \hat{x}_n}{\partial y_n} \end{bmatrix} \quad (4.22)$$

De donde cada elemento de la matriz tiene la forma:

$$\frac{\partial \hat{x}_i}{\partial y_j} = w_{ij} \frac{\partial \hat{x}_i}{\partial u_j} \quad (4.23)$$

Sustituyendo en la ecuación 4.24 se tiene que:

$$J(y) = \det \begin{bmatrix} w_{11} \frac{\partial \hat{x}_1}{\partial u_1} & \cdots & w_{1n} \frac{\partial \hat{x}_1}{\partial u_n} \\ \vdots & \ddots & \vdots \\ w_{1n} \frac{\partial \hat{x}_n}{\partial u_1} & \cdots & w_{nn} \frac{\partial \hat{x}_n}{\partial u_n} \end{bmatrix} \quad (4.24)$$

Como no existe una conexión entre las salidas después de las neuronas, las derivadas parciales de $\partial \hat{x}_i / \partial \hat{u}_j$, tomarán un valor diferente de cero solamente en los casos donde $i = j$. Por lo tanto, se puede definir la ecuación 4.24 como sigue:

$$J(y) = \det(\mathbf{W}) \prod_{i=1}^n \left| \frac{\partial \hat{x}_i}{\partial u_i} \right| \quad (4.25)$$

Sustituyendo la ecuación 4.25 en la ecuación 4.21:

$$\frac{\partial H(\hat{x}_i)}{\partial \mathbf{W}} = \frac{\partial}{\partial \mathbf{W}} \log \left(|\det(\mathbf{W})| \prod_{i=1}^n \left| \frac{\partial \hat{x}_i}{\partial u_i} \right| \right) = \frac{\partial}{\partial \mathbf{W}} \log |\det(\mathbf{W})| + \sum_{i=1}^n \frac{\partial}{\partial \mathbf{W}} \log \left| \frac{\partial \hat{x}_i}{\partial u_i} \right| \quad (4.26)$$

El primer término de la ecuación se expresa como:

$$\frac{\partial}{\partial \mathbf{W}} \log |\det(\mathbf{W})| = \frac{(\text{adj} \mathbf{W})^T}{\det \mathbf{W}} = (\mathbf{W}^T)^{-1} \quad (4.27)$$

Donde $\text{adj} \mathbf{W}$ es la matriz adjunta de \mathbf{W} . Debido a que el determinante de \mathbf{W} puede ser obtenido de:

$$\det(\mathbf{W}) = \sum_{j=1}^n w_{ij} \text{cof}(w_{ij}) \quad (4.28)$$

Donde $\text{cof}(w_{ij})$ es una matriz de cofactores de tamaño $i \times j$. El segundo término de la ecuación 4.26 puede ser obtenido como:

$$\frac{\partial}{\partial w_{ij}} \sum_{i=1}^n \log \left| \frac{\partial \hat{x}_i}{\partial u_i} \right| = \frac{1}{\frac{\partial \hat{x}_i}{\partial u_i}} \frac{\partial^2 \hat{x}_i}{\partial u_i^2} y_j^T \quad (4.29)$$

Definiendo la derivada de la no linealidad \hat{x}_i con respecto a u_i como una aproximación de la densidad de la fuente $p(u_i)$, como sigue:

$$p(u_i) = \frac{\partial \hat{x}_i}{\partial u_i} \quad (4.30)$$

De las ecuaciones 4.29 y 4.30, se obtiene que:

$$\frac{\partial \hat{x}_i}{\partial \mathbf{W}} \sum_{i=1}^n \log \left| \frac{\partial \hat{x}_i}{\partial u_i} \right| = \frac{\frac{\partial p(u)}{\partial u}}{p(u)} y^T \quad (4.31)$$

Sustituyendo los dos términos obtenidos, en la ecuación 4.32 se tiene:

$$\frac{\partial H(\hat{x}_i)}{\partial \mathbf{W}} = (\mathbf{W}^T)^{-1} + \left(\frac{\frac{\partial p(u)}{\partial u}}{p(u)} \right) y^T \quad (4.32)$$

Esta ecuación es el resultado del gradiente de la función de la entropía. Una forma de maximizar la entropía es por medio del gradiente natural, el cual se logra multiplicando por $\mathbf{W}^T \mathbf{W}$. Por lo tanto, la ecuación 4.32 se expresa como:

$$\Delta \mathbf{W} \propto \frac{\partial H(\hat{x}_i)}{\partial \mathbf{W}} \mathbf{W}^T \mathbf{W} = \left[I + \left(\frac{\frac{\partial p(u)}{\partial u}}{p(u)} \right) u^T \right] \mathbf{W} \quad (4.33)$$

Donde I es la matriz identidad y el segundo factor denota la no linealidad.

$$g(u) = -\frac{\frac{\partial p(u)}{\partial u}}{p(u)} \quad (4.34)$$

La ecuación 4.33 es:

$$\Delta \mathbf{W} \propto [I - g(u)u^T] \mathbf{W} \quad (4.35)$$

En donde el término $g(u)$ es la función no lineal y es determinante para la separación de fuentes. Algunas funciones no lineales fueron propuestas por Bell y Sejnowski. La ecuación 4.35 se conoce con el nombre de *regla de aprendizaje* y se emplea en un modelo neuronal recursivo para el ajuste de los pesos sinápticos en donde en cada iteración se realizará una actualización hasta encontrar su valor óptimo. Dicho modelo neuronal se conoce como la regla de Hebb y se expresa como:

$$\mathbf{W}(t+1) = \mathbf{W}(t) + \Delta \mathbf{W} \quad (4.36)$$

La función no lineal tiene un papel importante en la separación de las fuentes. Como se mencionó anteriormente, la función no lineal propuesta fue una función logística ($g_i(u_i) = \tanh(u_i)$), así la regla de aprendizaje se rescribe como:

$$\Delta \mathbf{W} \propto [I - \tanh(u)u^T] \mathbf{W} \quad (4.37)$$

4.3.1. Algoritmo INFOMAX

En la sección anterior se presentaron los fundamentos teóricos de INFOMAX, a continuación se presenta el algoritmo.

Algorithm 1 INFOMAX($eeg(t)$)

Require: Un registro EEG neonatal en formato ASCII

Ensure: Los componentes independientes u y la matriz de pesos W

- 1: Centrado de los datos, es decir, hacer su media cero.
 - 2: Elegir una matriz inicial $W = I$ (matriz identidad)
 - 3: Calcular $\hat{x}(t) = W(t)y(t)$; $i, j = 1, 2, \dots, n, i \neq j$
 - 4: Calcular $\Delta W(t) \propto [I - g(u(t))u^T(t)]$ donde $\varphi(u) = \tanh(u)$
 - 5: Calcular $W(t + \Delta t) = W(t) + \Delta W(t)$
 - 6: Actualizar $W(t) = W(t + \Delta t)$
-

Capítulo 5

Metodología

En esta sección se describe en términos generales la metodología propuesta para la detección y eliminación de artificios oculares en registros EEG neonatales reales. Cabe mencionar que se cuenta con una base de datos de registros EEG neonatales proporcionados por el Hospital Infantil de Tamaulipas, ubicado en Cd. Victoria, Tamps.

Para cumplir con los objetivos del proyecto la metodología propuesta está basada en análisis estadístico multivariado de los registros electroencefalográficos neonatales.

- a. **Preparación de los registros EEG neonatales.** Los archivos EEG neonatal generalmente vienen en el formato propio del equipo electroencefalografo. Para generalizar la aplicación de las técnicas se por convertirlos a formato ASCII. Una vez que se tiene el registro en formato ASCII, se procede a preprocesar los datos mediante un filtro pasa bajo y pasa alto con una frecuencia de corte especificada por un especialista en neurología pediátrica o la mejor reportada en la literatura.
- b. **Proceso Infomax.** Una vez preparados los registros, se procede a aplicar el algoritmo INFOMAX extendido, que implica la aplicación de un centrado, blanqueado y la aplicación del algoritmo INFOMAX, y como resultado se obtienen los componentes independientes.
- c. **Identificación de artificios oculares.** Se analizará cada uno de los componentes independientes, mediante el uso de mapas de actividad de cada compo-

nente en el cuero cabelludo, aquellos componentes que afecten la parte frontal, serán considerados como artificios oculares.

- d. **Eliminación de artificios oculares.** Una vez elegidos los componentes independientes como artificios, se procede a eliminarlos de la matriz de mezcla A .
- e. **Reconstrucción del registro EEG neonatal libre de artificios oculares.** Una vez eliminado los componentes de la matriz de mezcla A , se reconstruye el registro EEG neonatal, dicho registro corresponde al EEG neonatal libre de artificios oculares.
- f. **Evaluación de los resultados.** Debido a la naturaleza de los registros EEG, la evaluación de los resultados debe ser efectuada por un especialista en neurología pediátrica.

La Figura 5.1 bosqueja la metodología propuesta, en la cual se puede apreciar que tiene como dato de entrada un registro electroencefalográfico contaminado con diferentes artificios, en el presente trabajo de investigación nos referimos a los provocados por los movimientos oculares. La preparación de los registros da como resultado un EEG estandarizado en formato ASCII y filtrado.

Una vez preparados dichos registros se preprocesan utilizando un centrado y un blanqueado de datos, para facilitar la separación de las señales cerebrales de los artificios mediante la aplicación de INFOMAX. Una vez ejecutado el algoritmo ICA tiene como salida los componentes independientes, los cuales son utilizados para identificar y eliminar aquellos componentes que son considerados como artificios oculares. Finalmente se tiene un registro EEG neonatal libre de artificios oculares.

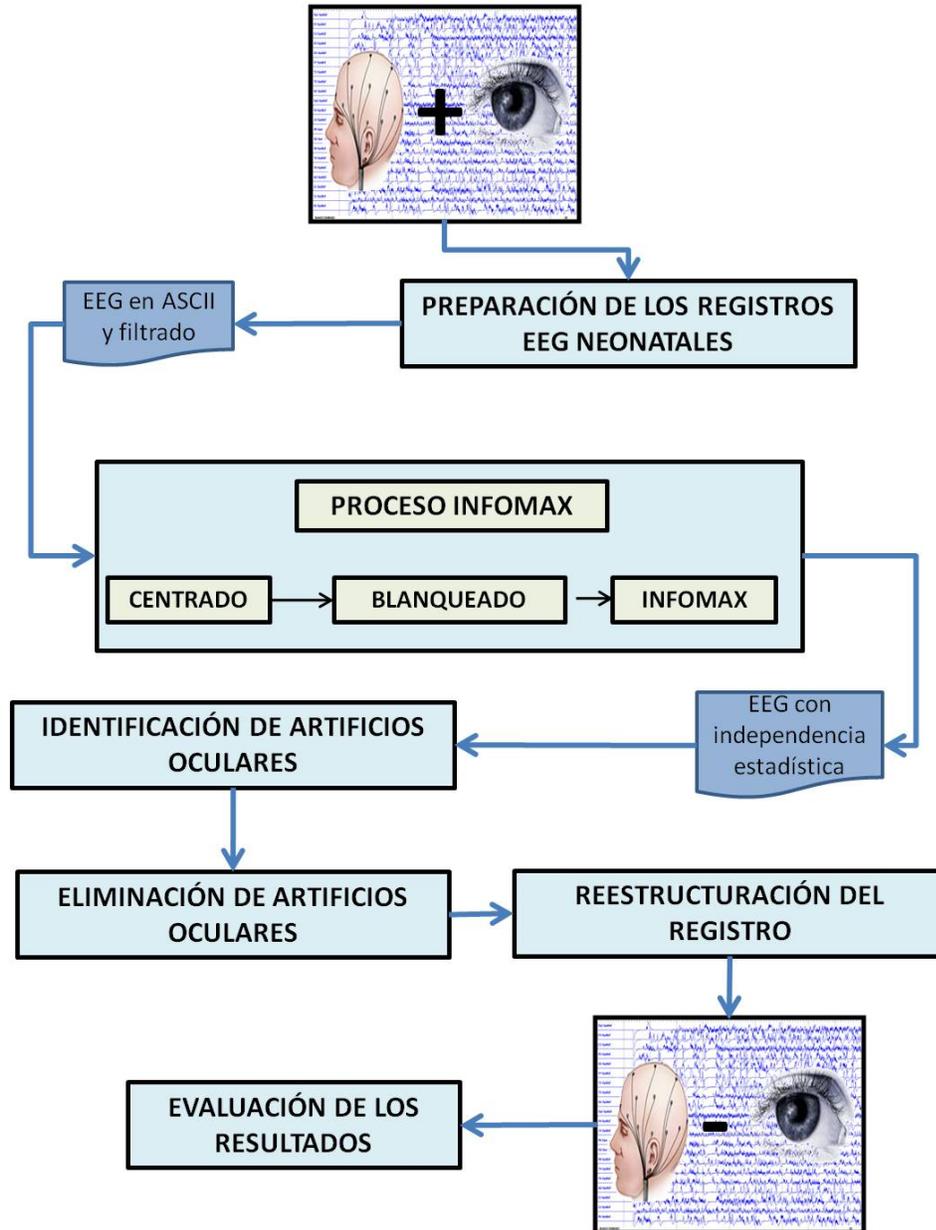


Figura 5.1: Metodología propuesta para la eliminación de artificios oculares en EEG neonatales

5.1. Preparación de los registros EEG neonatales

Para poder llevar a cabo la aplicación de las técnicas estadísticas multivariadas, es necesario que los datos cuenten con un formato estandar, para ello se proponen preparar los datos convirtiendo los registros EEG emitidos en un formato propio por el equipo electroencefalógrafo a un formato ASCII, esta conversión genera cuatro archivos ASCII(*Header, Annotation, SignalyData*). Finalmente una vez estandarizados los registros se procede a eliminar frecuencias no deseadas utilizando un filtro pasa bajo y un pasa alto, dando como resultado un registro EEG preparado para la aplicación de las técnicas utilizadas en el presente trabajo de investigación.

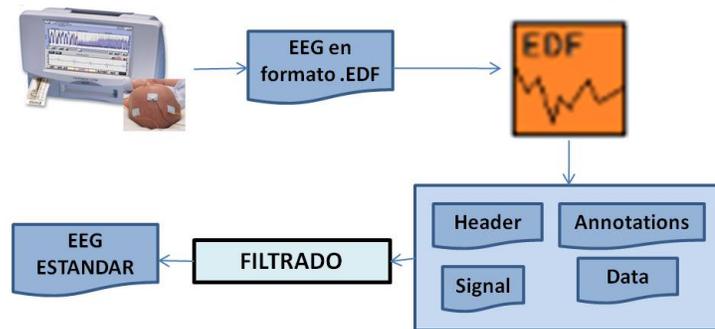


Figura 5.2: Estandarización de los registros EEG neonatales

5.1.1. Estandarización de los archivos EEG

Lo primero que se debe llevar a cabo es la conversión del formato .eeg, el cual es un formato propio emitido por el equipo electroencefalógrafo, en este caso el equipo con que cuenta el Hospital Infantil Tamaulipas es de la compañía Akonic, S.A., a un formato ASCII, esto con la finalidad de poder ser leídos y ser procesados mediante las herramientas y técnicas disponibles.

Para llevar a cabo la conversión primeramente los registros EEG neonatales se recuperaron en un formato EDF¹, para posteriormente ser convertidos a formato ASCII utilizando un software científico llamado EDFBrowser. El proceso para realizar la conversión se detalla a continuación:

¹El Formato de Datos Europeo (EDF), es un formato simple y flexible para intercambiar y almacenar señales físicas y biológicas multicanales[10]

Primeramente se tiene que abrir el software EDFBrowser, posteriormente se debe elegir **Open** del menú **File** (Ver Figura 5.3), para elegir el registro en formato EDF que queramos convertir a formato ASCII.



Figura 5.3: Menú File del software EDFBrowser

El software tiene la opción de determinar el número total de señales que se desea visualizar del conjunto total de señales que componen el registro EEG neonatal, como se muestra en la Figura 5.4. Una vez elegidas las señales a visualizar se debe dar clic en el botón **Add Signals**.

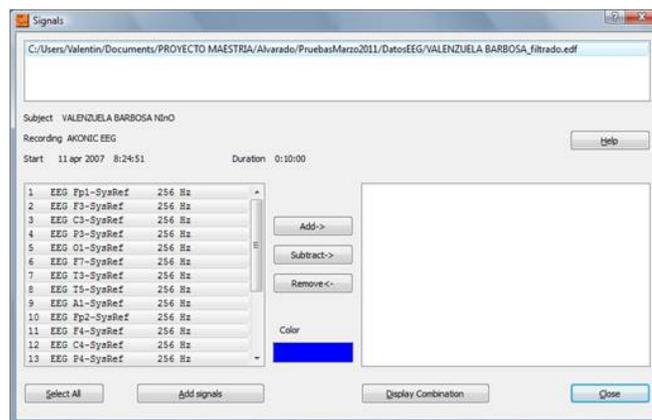


Figura 5.4: Señales a visualizar en la ventana de EDFBrowser

Ya que se tienen las señales visualizadas en el EDFBrowser, se procede a la conversión a formato ASCII, dicha conversión se lleva a cabo eligiendo la opción

Export EDF/BDF to ASCII del menú **Tools** (Ver Figura 5.5). Posteriormente se abre una ventana, donde se tiene que especificar la dirección donde se guardarán los archivos ASCII generados.

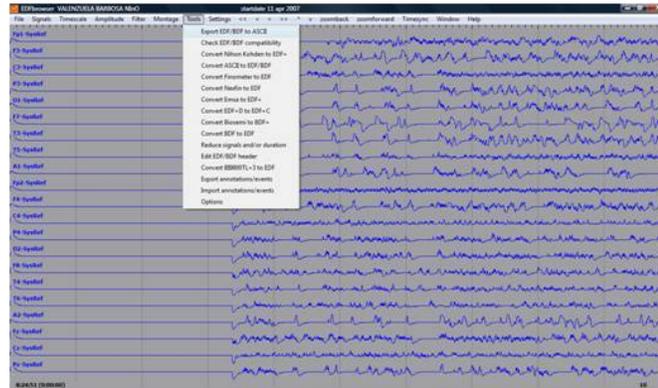


Figura 5.5: Convertir a ASCII desde el formato EDF

Los archivos generados por el proceso de estandarización de registro EEG son cuatro:

- * **Header:** El archivo *Header* contiene la versión del equipo (*Version*), datos del paciente (*Patient*), datos del registro (*Recording*), hora y fecha de registro (*Startdate* y *Starttime*) así como la duración del registro EEG (*Duration*) y número de señales (*NumSig*), (Ver Figura 5.6).

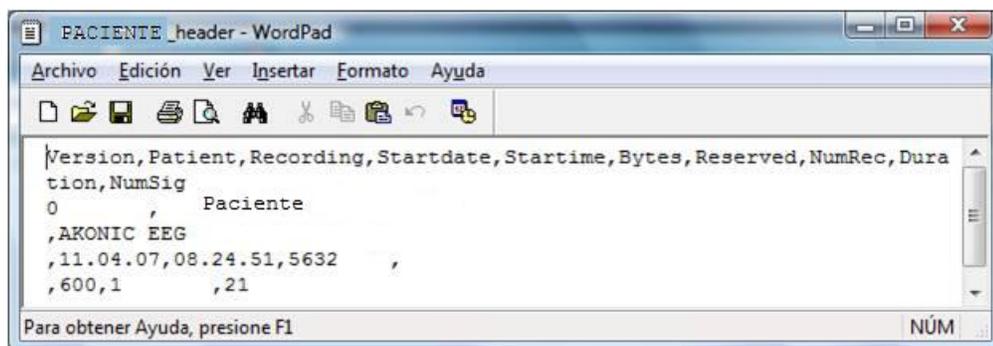
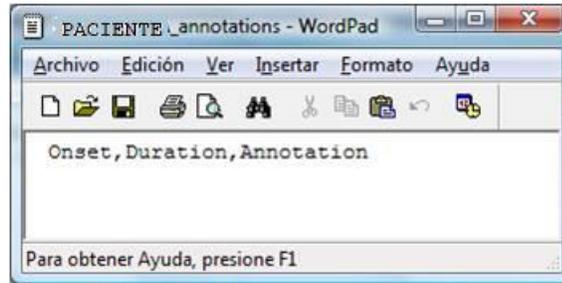
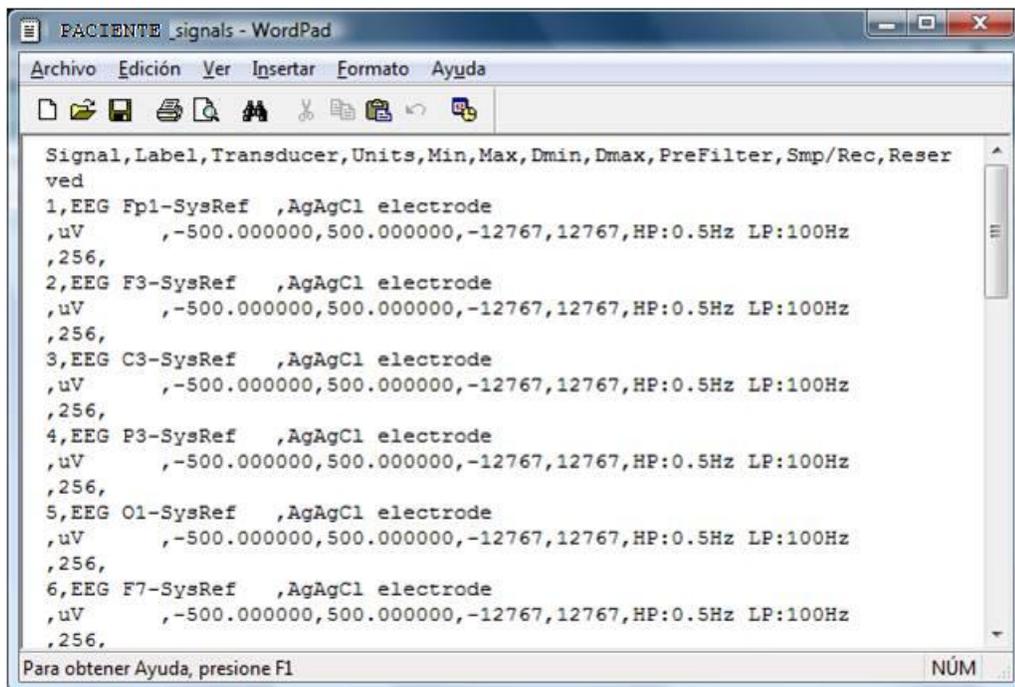


Figura 5.6: Contenido del archivo *header*

- * **Annotations:** El archivo *annotations* contiene la duración del registro EEG (*Duration*) y comentarios acerca del registro (*Annotation*), (Ver Figura 5.7).

Figura 5.7: Contenido del archivo *annotacions*

- * **Signals:** El archivo *signals* contiene la descripción de cada una de las señales involucradas en el registro EEG como lo son: la frecuencia de muestreo (*Smp/Rec*), el filtrado de equipo (*PreFilter*), las unidades de medida en μV (*Units*), (Ver Figura 5.8).

Figura 5.8: Contenido del archivo *signals*

- * **Data:** El archivo *Data* contiene los valores discretos del registro, (Ver Figura 5.9).

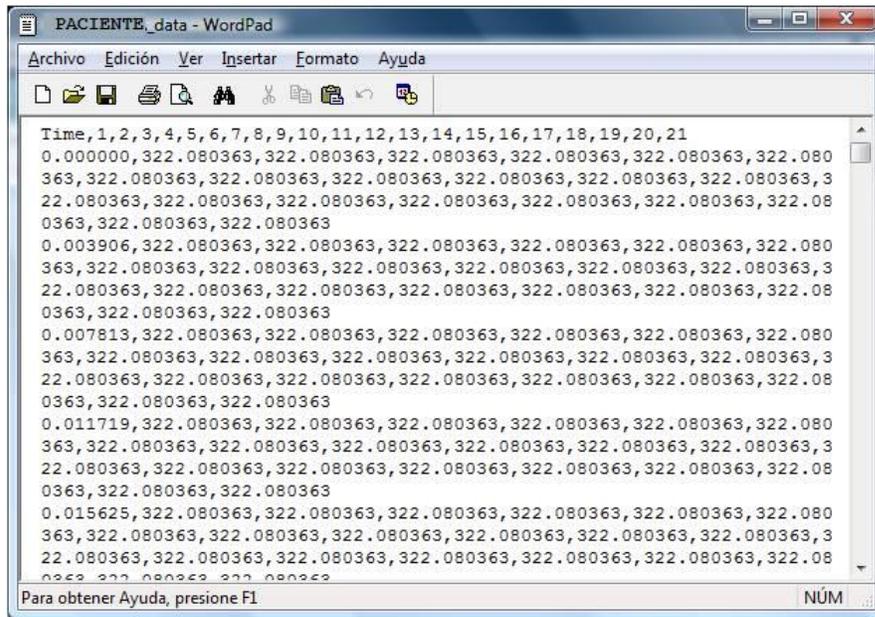


Figura 5.9: Contenido del archivo *data*

La magnitud de la matriz de datos del archivo *Data*, se determina de la siguiente manera:

$$\begin{aligned} \text{ renglones} &= (\text{Smp}/\text{Rec} \times \text{Duration}) + 1 \\ \text{filas} &= \text{NumSig} + 1 \end{aligned}$$

En el caso particular de los datos del Hospital Infantil, estos fueron muestreados a 256Hz, y son aproximadamente 10 minutos de duración del registro (600 segundos), la cantidad de electrodos o señales son 21. Por lo tanto, el formato de salida es una matriz de aproximadamente 153601 filas X 22 columnas. La matriz de datos consta de un renglon de encabezado que representan las señales involucradas y la columna uno representa las series de tiempo y el resto representa cada una de las señales del registro.

5.1.2. Filtrado de los registros EEG neonatales

Una vez recuperado el estudio EEG, el siguiente paso es llevar a cabo el filtrado de los datos con la finalidad de eliminar frecuencias no deseadas en los registros

EEG. El primer filtro que se aplica es el pasa bajo con un frecuencia de corte de 3 Hz, posteriormente se aplica un filtro pasa alta con un corte de frecuencia de 15 Hz. Dichas frecuencias son recomendadas por el Dr. Luis Carlos Ortega Tamez, especialista en neurología pediátrica en el Hospital Infantil de Tamaulipas.

Para llevar a cabo el filtrado de las señales en el software EDFBrowser, se debe ir al menú **Filter**, posteriormente seleccionar aplicar nuevo filtro (*New*). Se abrirá una ventana en la que se debe especificar que tipo de filtro llevar a cabo y a que señales se les debe aplicar dicho filtro. Cabe mencionar que los filtros se pueden aplicar a una o mas señales dentro de un mismo registro. Para fines de este trabajo de investigación se le aplicarán los filtros a todas las señales del registro, como se muestra en la Figura 5.10.

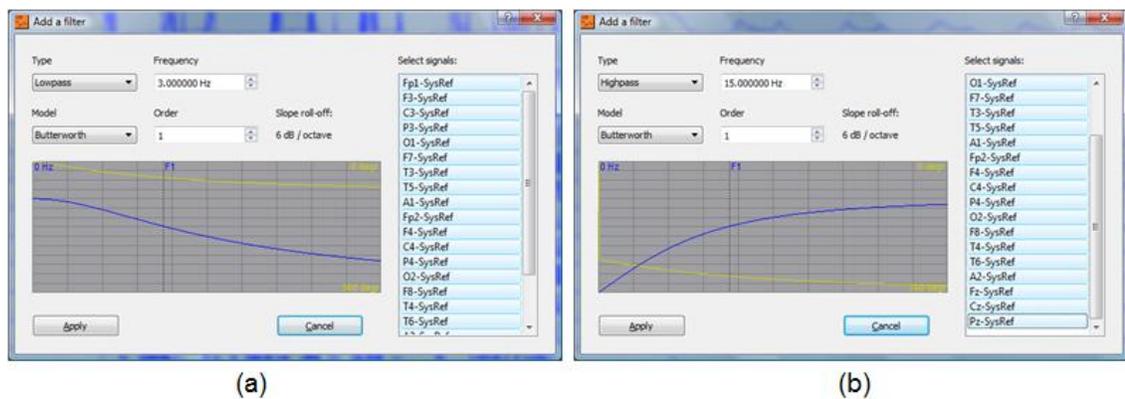


Figura 5.10: La Figura (a) muestra la configuración del filtro pasa bajo y el (b) del filtro pasa alto

La Figura 5.11 en la parte superior muestra el EEG en crudo y en la figura inferior se muestra el EEG después de aplicar los filtros anteriores.

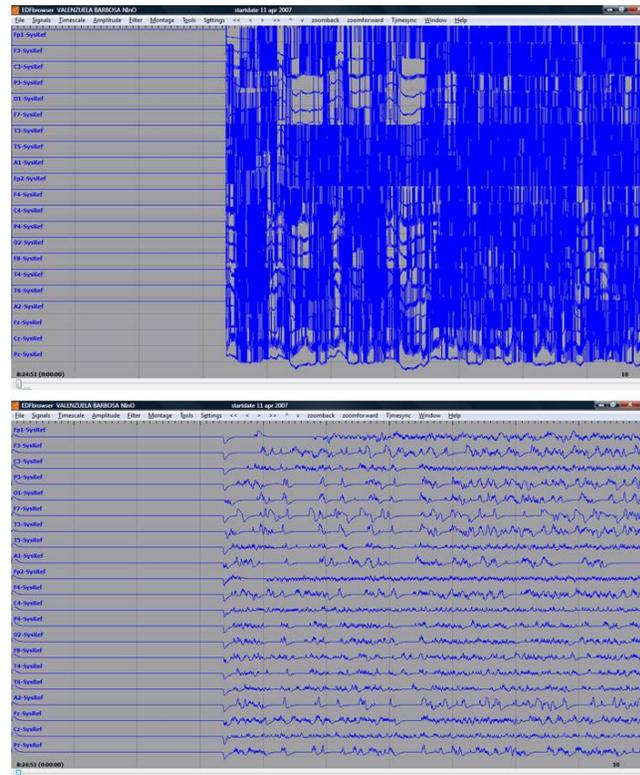


Figura 5.11: Figura superior: EEG en crudo; Figura inferior: EEG filtrado

5.2. Pre-procesamiento INFOMAX extendido

Es más sencillo estimar la matriz de mezcla W si las señales medidas tienen una media de cero, una varianza de uno y una correlación cero. Los datos deben ser pre-procesados para alcanzar la separación de las fuentes y la obtención de la matriz W .

Para que las señales tengan una media de cero, es necesario restarles a cada una de las señales la media de dicha señal (*Centrado*). Una matriz de covarianzas puede ser formada tomando la covarianza entre pares de señales y así formar una matriz. La matriz de covarianza deberá ser cuadrada y simétrica, este paso se puede llevar a cabo mediante la aplicación del PCA (*Blanqueado*).

5.2.1. Centrado de los datos

El primer paso para la aplicación del Análisis de Componentes Independientes es centrar la matriz de mezcla observada. Este paso se hace solamente para simplificar el proceso ICA y no afecta su resultado. El centrado es llevado a cabo simplemente eliminando la media de cada mezcla observada e implica que las fuentes son también centradas. La media puede ser regresada fácilmente a los datos después de aplicar el algoritmo ICA.

La Figura 5.12 muestra un ejemplo de la gráfica de la relación de las señales $Fp1$ y $Fp2$. En la figura del lado izquierdo puede observar que los datos están totalmente esparcidos por todo el plano, mientras que en la figura del lado derecho se puede observar que los datos fueron centrados, en este paso aun existe dependencia estadística de las señales, la cual será sustraída posteriormente por el algoritmo INFOMAX.

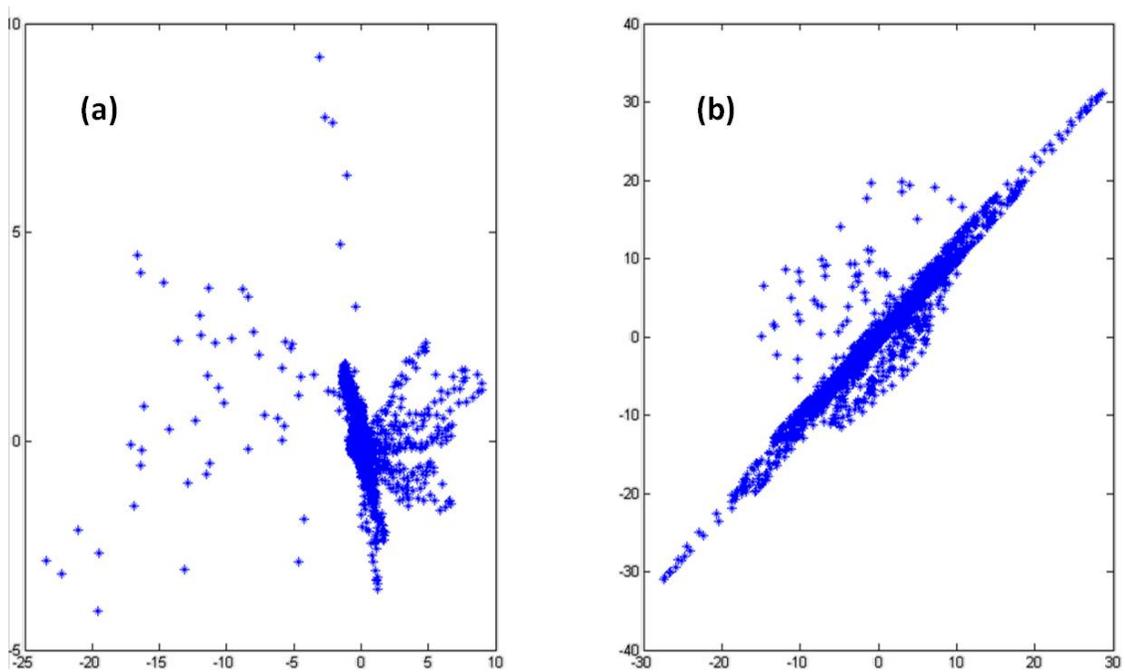


Figura 5.12: Señales $Fp1$ y $Fp2$ antes y después del centrado

5.2.2. Blanqueado de los datos

El siguiente paso es el blanqueado de los datos y es aplicado después del proceso de centrado y antes de aplicar el algoritmo ICA. El procedimiento de blanqueado es una transformación lineal de los datos a un nuevo vector de datos compuesto de elementos los cuales son no correlacionados y con varianza la unidad.

Esta característica de los datos se puede obtener llevando a cabo el proceso de PCA, en el cual construye una transformación lineal que escoge un nuevo sistema de coordenadas para el conjunto original de datos en el cual la varianza de mayor tamaño del conjunto de datos es capturada en el primer eje (llamado el Primer Componente Principal), la segunda varianza más grande es el segundo eje, y así sucesivamente. Para construir esta transformación lineal debe construirse primero la matriz de covarianza o matriz de coeficientes de correlación.

En otras palabras, la función de PCA es transformar un conjunto de variables correlacionadas a un conjunto menor o igual de variables no correlacionadas, gráficamente se puede ver en la Figura 5.13. En la Figura (a) se muestra un conjunto de datos que están correlacionados y en la figura (b) se muestra el mismo conjunto de datos no correlacionados.

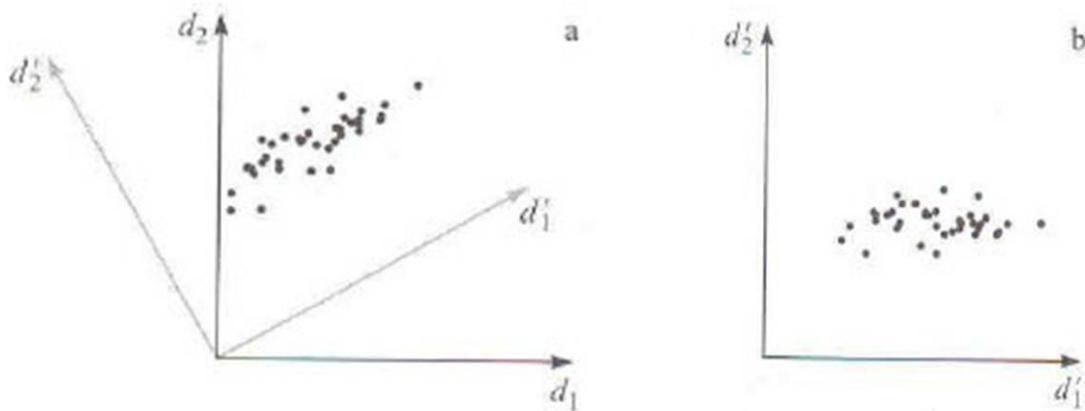


Figura 5.13: Función de PCA

5.3. Extensión propuesta del algoritmo INFOMAX para la detección de artificios oculares

Una limitación que se presenta con el algoritmo de INFOMAX es que solo logra la separación de las señales cuando tienen distribuciones sub-gaussianas y no separa cuando son distribuciones super-gaussianas. Para resolver dicha limitación se presenta el algoritmo INFOMAX extendido, que es una modificación del algoritmo INFOMAX original.

5.3.1. Aplicación de INFOMAX extendido

El objetivo de INFOMAX extendido es proveer una regla de aprendizaje sencilla que pueda separar señales con una variedad de distribuciones. Una forma generalizada de la regla de aprendizaje es si las fuentes con distribución super o sub gaussianas es aproximadamente la función de densidad de probabilidad estimada.

La regla de aprendizaje para distribuciones sub-gaussianas es:

$$\Delta W \propto [I + \tanh uu^T - uu^T] W \quad (5.1)$$

La regla de aprendizaje para distribuciones super-gaussianas es:

$$\Delta W \propto [I - \tanh uu^T - uu^T] W \quad (5.2)$$

La diferencia entre la regla de aprendizaje entre las distribuciones super y sub gaussianas es el signo antes de la función \tanh , y puede ser determinado por una regla de identificación de la distribución. Sin embargo como no hay una definición general para las distribuciones sub y super gaussianas una forma válida para el criterio de selección de signo esta basado en el criterio de estabilidad.

El criterio entre la regla de aprendizaje para sub y super gaussianas es:

$$\Delta W \propto [I - K \tanh uu^T - uu^T] W \left\{ \begin{array}{l} k_i = 1 : \text{super - gaussianas} \\ k_i = -1 : \text{sub - gaussianas} \end{array} \right\} \quad (5.3)$$

Donde k_i son los elementos de la matriz diagonal K de $N \times N$ dimensiones. La

selección del signo puede ser derivado del análisis de las soluciones separadas. En dicho análisis el promedio es aproximadamente una perturbación de primer orden en la matriz de separación.

Por lo tanto la regla para seleccionar el signo queda de la siguiente manera:

$$k_i = \text{sig} (E \{ \text{sech}^2(u_i) \} E \{ u_i^2 \} - E \{ [\tanh u_i] u_i \}) \quad (5.4)$$

5.3.2. Obtención de los componentes independientes

Para obtener los componentes independientes, es necesario conocer primeramente el algoritmo de INFOMAX extendido, que es mostrado a continuación. En el proceso de INFOMAX extendido se puede obtener la matriz de pesos W así como también los componentes independientes u , de cada registro EEG neonatal al cual se le haya aplicada dicho algoritmo.

Algorithm 2 INFOMAX($eeg(t)$)

Require: Un registro EEG neonatal en formato ASCII

Ensure: Los componentes independientes u y la matriz de pesos W

- 1: Centrado de los datos, es decir, hacer su media cero.
 - 2: Blanqueado de los datos aplicando PCA.
 - 3: Elegir una matriz inicial $W = I$ (Matriz Identidad)
 - 4: Calcular $\hat{x}(t) = W(t)y(t)$; $i, j = 1, 2, \dots, n$, $i \neq j$
 - 5: Calcular el signo $k_i = \text{sig} (E \{ \text{sech}^2(u_i) \} E \{ u_i^2 \} - E \{ [\tanh u_i] u_i \})$
 - 6: Calcular $\Delta W(t) \propto [I - \varphi(u(t))u^T(t)]$ donde $\varphi(u) = \tanh(u)$
 - 7: Calcular $W(t + \Delta t) = W(t) + \Delta W(t)$
 - 8: Actualizar $W(t) = W(t + \Delta t)$
 - 9: Calcular $u(t) = W(t) * eeg(t)$
-

5.3.3. Identificación de componentes independientes asociados a artificios oculares

Posterior a la obtención de los componentes independientes del EEG, se procede a identificar aquellos componentes que son indicativos de posibles artificios oculares. La identificación de los componentes se lleva a cabo utilizando mapas de actividad

cerebral. La intensidad se representa por una escala de colores. En la Figura 5.14, para cada componente se muestra un mapa de como se ven afectados cada uno de los electrodos por el componente. Como los artificios oculares afectan a la zona frontal del cerebro, se deben seleccionar los componentes de mayor intensidad relativa en esa zona.

Dichos mapas son generados como sigue:

- a. La columna de la matriz inversa W indican la intensidad con que un componente contribuye a cada electrodo.
- b. La intensidad de un electrodo es mapeado en un círculo.
- c. Un mapa de colores es asignado para mostrar la magnitud de los componentes en cada electrodo.
- d. Correspondiente a cada posición del electrodo, un mapa en dos dimensiones puede ser generado.

En el ejemplo mostrado en la Figura 5.14 se puede ver que los componentes 4, 6, 9, 19 y 20, se ven afectados en las zonas frontales, es decir, se tomarían como posibles artificios oculares.

5.4. Eliminación de artificios oculares

Los registros EEG son buenos indicadores de las anormalidades en el sistema nervioso central. La presencia de artificios producidos por la contaminación de los movimientos del glóbulo ocular y el párpado pueden enmascarar la presencia de crisis en el registro EEG neonatal. Este es un problema muy conocido y que aparece en el registro EEG neonatal como una interferencia, causando serios problemas a la hora del análisis e interpretación de los registros [40]. En base a ello, la importancia de la eliminación de los artificios oculares, para ello, una vez identificado aquellos componentes como posibles artificios oculares, se procede a eliminarlos.

La forma en la cual se llevará a cabo la eliminación es mediante el uso de la matriz de mezcla A , la cual es la matriz inversa de W . La matriz W es una matriz cuadrada de tamaño del número de canales registrados en el EEG neonatal, en este caso la

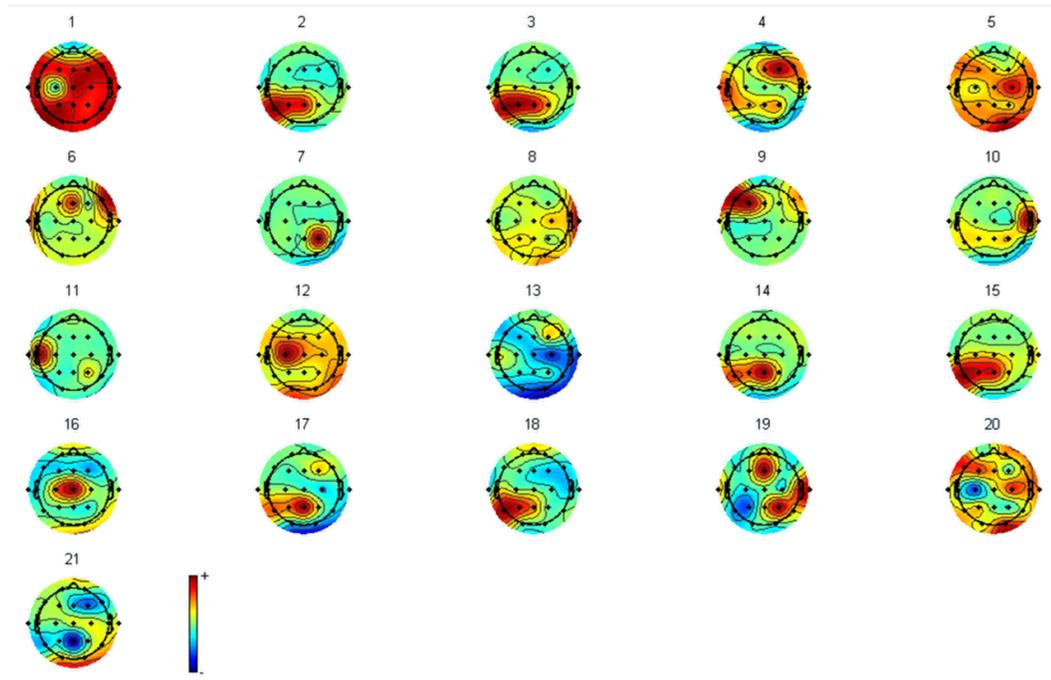


Figura 5.14: Identificación de artificios oculares

dimensión de la matriz sería de 21×21 . En la Figura 5.15 se muestra un ejemplo del contenido de una matriz W , en la cual se puede observar que las dimensiones son las antes mencionadas.

Para llevar a cabo la eliminación de los posibles artificios oculares se procede a eliminar los componentes seleccionados de la matriz A , esto se realiza haciendo ceros la columna correspondiente a dicho componente. De acuerdo al ejemplo de la Figura 5.14, se determinaron que los componentes 4, 6, 9, 19 y 20 se van a eliminar de la matriz A , como se muestra en la Figura 5.16.

5.5. Reconstrucción de el registro EEG neonatal libre de artificios oculares

Una vez eliminado los componentes asociados a artificios oculares, se debe reconstruir el registro EEG neonatal libre de artificios oculares. De acuerdo a la formula $y(t) = W \times x(t)$, se tiene que $x'(t) = A \times y(t)$, donde x' es la aproximación de las

-0.001552	-0.003485	-0.186277	-0.187059	-0.228808	-0.140085	-0.250124	-0.24976	-0.152748	-0.044204	-0.158238	-0.184104	-0.251897	-0.238117	-0.292529	-0.285946	-0.274803	-0.206241	-0.150562	-0.154394	-0.183212
0.002581	0.007571	0.172499	0.17143	0.235912	0.120673	0.310556	0.304838	-0.018081	-0.041895	0.009859	0.082673	0.20895	0.035713	-0.426971	-0.392164	-0.378015	-0.242644	0.056225	0.057107	0.072223
-0.000175	0.001779	0.118498	0.118447	-0.062645	-0.17684	-0.145515	-0.149303	-0.325323	-0.000906	0.214603	0.221556	-0.168609	0.13065	0.098532	0.091077	0.082581	-0.050989	0.227945	0.233415	0.220677
-0.000084	0.000678	-0.005191	-0.005354	0.159005	0.140433	0.381499	0.360942	-0.004157	-0.065916	-0.167039	-0.167478	-0.554405	-0.152262	0.173059	0.159444	0.155135	-0.318389	-0.151272	-0.152543	-0.16375
-0.001774	-0.005919	-0.166783	-0.166209	-0.013567	-0.050409	0.046028	0.056286	-0.099449	-0.014518	-0.173029	-0.168032	0.689152	0.040959	0.150377	0.134305	0.13406	-0.445185	-0.200076	-0.202957	-0.162238
-0.0037	-0.006677	-0.01744	-0.017642	0.033204	-0.498872	0.214276	0.232944	-0.715658	0.05493	-0.040555	-0.014558	0.007861	0.036224	0.000087	0.002303	0.00431	0.34757	-0.026574	-0.027566	-0.014565
0.002012	0.016716	0.485052	0.4845	0.230505	0.080738	-0.252524	-0.253355	-0.179838	-0.262198	-0.398433	-0.080369	0.025261	0.03755	0.04451	0.037406	0.041212	0.054431	-0.163097	-0.161812	-0.07892
-0.00571	-0.014334	-0.133483	-0.129919	0.329876	-0.380884	-0.113795	0.005991	0.285882	0.092888	-0.199551	0.101756	-0.186166	0.661113	-0.058124	-0.051407	-0.047563	-0.058526	-0.176563	-0.179014	0.103062
0.002939	0.002451	-0.255593	-0.254432	0.502482	0.568977	-0.185674	-0.135726	-0.425955	0.089729	0.158023	0.020961	-0.04561	0.129845	-0.041229	-0.005716	0.015863	0.058144	-0.031916	-0.039463	0.007818
0.00357	0.015889	-0.077928	-0.077985	-0.555963	0.346145	0.159461	0.055547	-0.168671	-0.446178	-0.180244	0.113562	-0.056087	0.485543	-0.031083	-0.016416	-0.016523	0.064829	-0.036057	-0.030728	-0.114108
-0.006073	-0.021686	0.210747	0.204881	-0.333034	0.204466	0.069118	-0.032323	-0.091774	0.75792	0.090013	-0.00797	-0.042281	0.189613	-0.043182	-0.010005	-0.00874	-0.032029	-0.251308	-0.251171	-0.029124
0.007096	0.010983	-0.113848	-0.117002	0.000873	-0.089478	-0.00816	-0.007074	0.030448	-0.275655	0.697774	-0.396299	-0.023416	0.285527	-0.005796	-0.03743	-0.021078	0.003504	-0.089932	-0.108235	-0.385602
-0.004241	-0.008643	-0.009166	0.003458	-0.011697	0.06382	-0.008555	-0.001142	-0.019005	0.234843	-0.335738	-0.429494	-0.015882	0.279117	0.036672	-0.061518	-0.014594	-0.007254	0.471041	0.453484	-0.375558
-0.00047	-0.020754	-0.028187	-0.037172	0.10935	-0.060965	0.68022	-0.726159	0.018498	-0.008295	-0.005374	0.00358	0.002876	0.016193	0.017932	0.008663	-0.034782	0.003803	-0.002668	-0.014686	-0.003749
0.014188	0.015482	0.00836	-0.015793	0.009557	0.02891	-0.027671	0.028329	-0.018075	0.013049	0.008099	-0.01678	0.000924	0.007551	0.550432	0.190362	-0.811195	0.006947	-0.079911	0.047119	0.041539
0.127255	0.269843	0.048487	-0.052011	-0.001526	-0.01891	-0.004573	-0.000792	0.008714	0.005289	-0.018317	0.201069	-0.000579	0.01573	-0.499632	0.712198	-0.179164	-0.004741	0.002232	0.064794	-0.288745
0.295084	0.905459	-0.020995	0.002652	0.000321	-0.010624	0.016089	-0.013402	0.001171	0.0351	-0.002346	0.007383	-0.000692	-0.000636	0.170259	-0.247404	0.072174	-0.000151	0.038252	-0.061832	0.016951
0.121666	-0.122982	0.143865	-0.147727	-0.000135	-0.000131	-0.007809	0.008217	0.00086	-0.006101	-0.005242	0.59505	-0.00023	0.00544	0.20193	-0.272339	0.0677	0.001309	-0.209285	0.223172	-0.61025
-0.150811	0.135236	0.094574	-0.097779	0.003094	-0.001188	-0.004047	0.00256	0.000398	0.004961	0.006741	-0.224313	-0.000236	0.003072	-0.080211	0.013218	0.069115	0.00017	-0.651768	0.65075	0.206658
-0.925769	0.287627	-0.075841	0.07272	0.000528	0.000787	0.003783	-0.003438	0.001434	-0.000332	0.002425	0.155717	0.000551	0.000369	0.033148	-0.016773	-0.017125	0.000039	0.072433	-0.069128	-0.159853
-0.109781	0.034141	0.682777	-0.684368	0.001368	0.000345	-0.001637	0.001596	0.000453	-0.000622	-0.000429	-0.09737	-0.000033	0.00145	0.002563	-0.005838	0.003241	0.000015	0.150941	-0.148858	0.100595

Figura 5.15: Matriz W

fuentes originales.

De esta forma obtendremos un registro EEG libre de artificios en función de la matriz A . La formula $x'(t) = A \times y(t)$ es la que será utilizada para reconstruir los registros EEG neonatales libres de artificios.

En la Figura 5.17 se muestra un ejemplo de un registro EEG neonatal libre de artificios oculares.

5.6. Evaluación de los resultados

La evaluación de proyectos es un proceso por el cual se determina la calidad del método propuesto. Es decir, se intenta conocer qué tanto un proyecto ha logrado cumplir sus objetivos o bien qué tanta capacidad poseería para cumplirlos.

En una evaluación de proyectos siempre se produce información para la toma de decisiones, por lo cual también se le puede considerar como una actividad orientada a mejorar la eficacia de los proyectos en relación con sus fines, además de promover mayor eficiencia en la asignación de recursos.

En este proyecto participan expertos en el área de señales y reconocimiento de patrones el cual liderea el Dr. Claudio Castellanos Sánchez, la Dra. Laura Cruz

Capítulo 6

Pruebas y resultados

En esta sección se presentan algunas pruebas realizadas con ayuda de la herramienta MATLAB. Los datos con los que se realizaron las experimentaciones fueron proporcionados por el Hospital Infantil Tamaulipas ubicado en Cd. Victoria, los cuales constan de 21 canales y fueron muestreados a 256 Hz.

En estas pruebas se estuvo trabajando con el EEG con características propias de artificios oculares, es decir, un especialista del Hospital Infantil de Tamaulipas hizo las observaciones sobre dicho EEG, en el cual se indicó en que momento el neonato tenía los ojos abierto y en que momento los cerró. A partir de estas especificaciones se aplicó INFOMAX a todo el registro para ver el comportamiento de los componentes independientes.

6.1. Ambiente experimental

6.1.1. Infraestructura

Las experimentaciones se llevaron a cabo en una computadora portátil. Dichas experimentaciones se realizaron bajo las mismas circunstancias y mismo equipo de cómputo. Las características del equipo de cómputo se enlistan a continuación:

- * Computadora portátil marca Dell Studio 1537
- * Sistema operativo Windows Vista Home Premium

- * Procesador Intel Core 2 Duo a 2.26 GHz y 2.27 GHz
- * Memoria RAM de 3Gb
- * Sistema operativo de 32 bits

6.1.2. Base de datos

La base de datos consta de 4 registros EEG de 4 neonatos, con una duración promedio de registro de 10 minutos (600 segundos) aproximadamente. El registro consta de 21 canales de acuerdo al Sistema Internacional 10-20 ($Fp1$, $F3$, $C3$, $P3$, $O1$, $F7$, $T3$, $T5$, $A1$, $Fp2$, $F4$, $C4$, $P4$, $O2$, $F8$, $T4$, $T6$, $A2$, Fpz , Fz , Cz).

Los registros fueron recolectados en el Hospital Infantil de Tamaulipas, ubicado en Cd. Victoria, Tamps., bajo la dirección del Dr. Luis Carlos Ortega Tamez. Estos registros fueron muestreados a 256 Hz, es decir, 256 datos por segundo.

6.2. Determinación de la tasa de aprendizaje

6.2.1. Objetivo

Determinar el valor del parámetro alfa, en otras palabras, el valor de la tasa de aprendizaje para registros EEG neonatales

6.2.2. Método

La experimentación se realizó utilizando el código implementado en MATLAB, con datos de registros EEG neonatales como datos de entrada y con un valor alfa por defecto que es de 0.0004. Este es el valor de la tasa de aprendizaje que utilizó en la separación de señales de audio. Con este valor de alfa no se obtuvo resultados en la aplicación con registros EEG debido a que en un punto ya no se tenían valores numéricos sino que había pérdida de datos y como resultado teníamos un valor denominado NaN , es decir, un valor indefinido.

Para resolver este problema se empezó a probar con diferentes valores para alfa para poder encontrar no el valor óptimo pero sí un valor de alfa que nos pueda completar la matriz de salida con valores definidos numéricamente.

0.1 - 0.9
0.001 - 0.009
0.0001 - 0.0009
0.00015 - 0.00095
0.000015 - 0.000095
0.0000001 - 0.0000009

Cuadro 6.1: Valores para el parámetro alfa en la regla de aprendizaje

La tabla 6.1 muestra algunos de los valores con los que se probó la implementación de infomax.

Los valores para alfa con los cuales la matriz de salida del algoritmo es completamente numérica son los valores de la tabla anterior que se encuentran en el rango (0.0000001 - 0.000009). A partir de estos valores ya se tiene una matriz numérica, el siguiente paso es encontrar los valores óptimos de alfa para los datos EEG neonatales que tenemos como entrada.

La figura 6.1 presenta una de las ventanas de cómo se muestra los datos de salida con valores no definidos. En este caso se utilizó el valor por defecto del algoritmo con $\alpha=0.0004$, con dicho valor a partir del segundo 4 empezaron los valores indefinidos. En otras palabras, de acuerdo a la naturaleza de los datos, el valor de la tasa de aprendizaje que logra separar señales de audio, no logra separar las fuentes de los registros EEG neonatales.

6.2.3. Resultado

Mediante la aplicación de esta prueba se pudo determinar el valor del parámetro alfa del algoritmo INFOMAX para registros EEG neonatales. El valor de la tasa de aprendizaje para registros EEG neonatales se encuentra en el rango de (0.0000001 - 0.000009). Como se puede observar, el valor de la tasa de aprendizaje son valores muy pequeños, pero esto es debido a la cantidad de datos con los que trabaja el algoritmo infomax, en este caso, el número de datos promedio es de 153000 datos por cada uno de los registros.

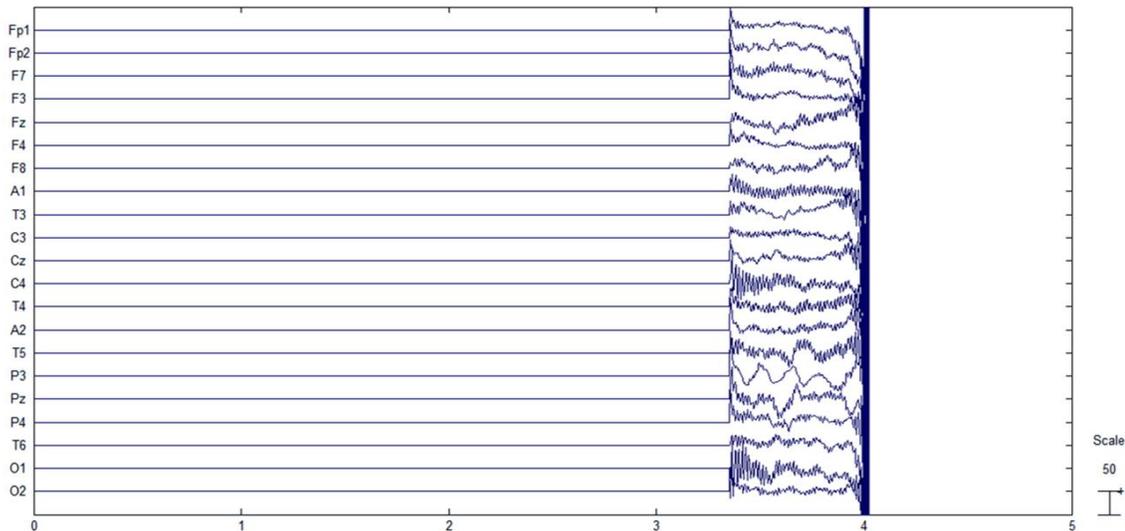


Figura 6.1: ICA en el segmento de 15 segundos de un EEG

6.3. Aplicación de INFOMAX

6.3.1. Objetivo

Aplicar el algoritmo INFOMAX para separar las señales cerebrales de las señales artificiales.

6.3.2. Método

ICA fue aplicado a todo el registro para calcular los componentes independientes con la finalidad de posteriormente eliminar el artificio ocular. El número de componentes independientes es igual al número de canales presentes en el registro, para facilitar los cálculos en la técnica ICA. La aplicación de ICA fue realizado sobre los datos filtrados y eliminando los segundos que se utilizan para calibrar el equipo, normalmente son los primeros cuatro segundos.

La Figura 6.2 muestra el EEG neonatal de un paciente filtrado, al cual se aplica el algoritmo INFOMAX, para obtener los componentes independientes y así tener la separación de las señales del cerebro de las señales artificiales.

Los componentes independientes obtenidos de dicho segmento del registro son los que se muestran en la figura 6.3

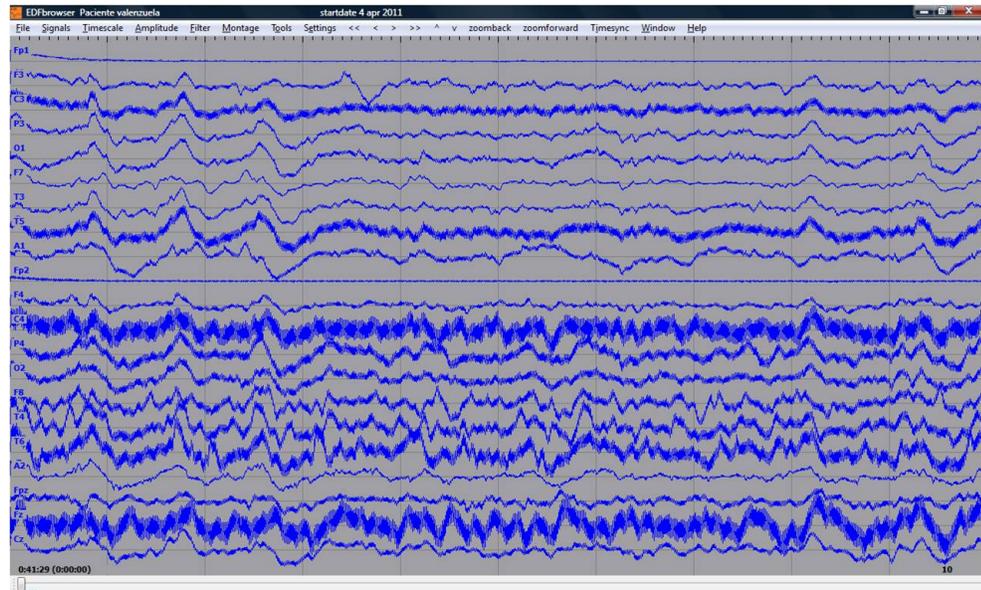


Figura 6.2: EEG filtrado

6.3.3. Resultado

Como se puede observar en las figuras, el algoritmo INFOMAX no logra separar las señales cerebrales de las señales artificiales, esto es debido a la naturaleza de los datos. Para ello, es necesario la aplicación del algoritmo INFOMAX extendido, ya que este último algoritmo logra la separación de las fuentes con distribuciones sub gaussianas y super gaussianas, mientras que en el INFOMAX original solo logra la separación con una sola distribución.

6.4. Aplicación de INFOMAX extendido

6.4.1. Objetivo

Aplicar el algoritmo INFOMAX extendido para separar las señales cerebrales de las señales artificiales.

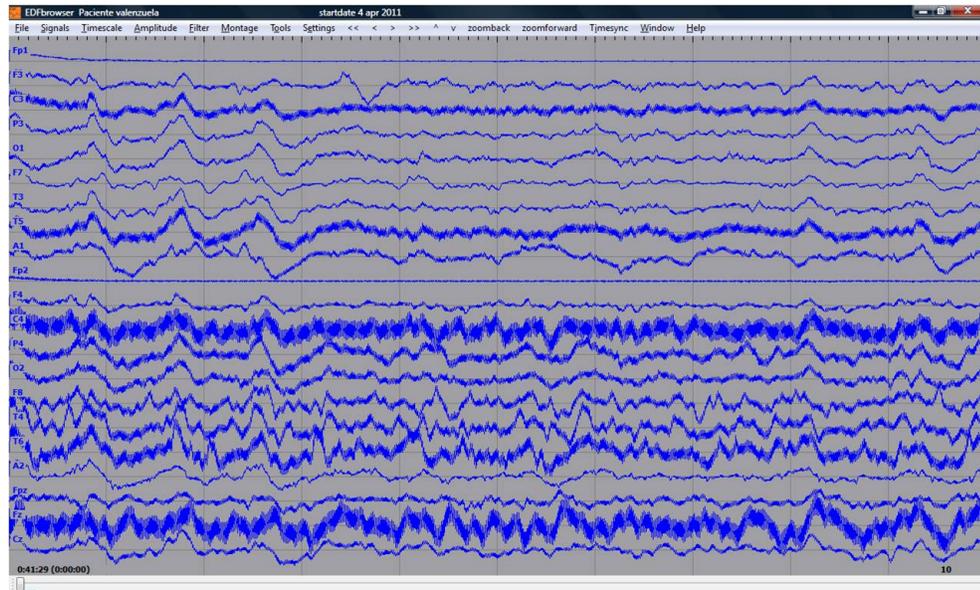


Figura 6.3: Componentes independientes después de aplicar INFOMAX

6.4.2. Método

Debido a que el algoritmo INFOMAX no logra separar las señales del registro EEG neonatal, se procede a hacerle la extensión al algoritmo para que separe señales con distribuciones sub y super gaussianas. El algoritmo que se aplica es el INFOMAX extendido.

La Figura 6.4 muestra el EEG neonatal de un paciente filtrado, al cual se aplica el algoritmo INFOMAX, para obtener los componentes independientes y así tener la separación de las señales del cerebro de las señales artificiales.

Los componentes independientes obtenidos de dicho segmento del registro son los que se muestran en la figura 6.5

La Figura 6.7 muestra los mapas de actividad para cada uno de los componentes independientes. Los componentes 13, 14, 15 y 19, son los identificados como artificios oculares, debido a que son los componentes que afectan las zonas frontales del cerebro.

6.4.3. Resultado

La modificación del algoritmo INFOMAX al INFOMAX extendido logra separar satisfactoriamente las señales cerebrales, para posteriormente la eliminación de los

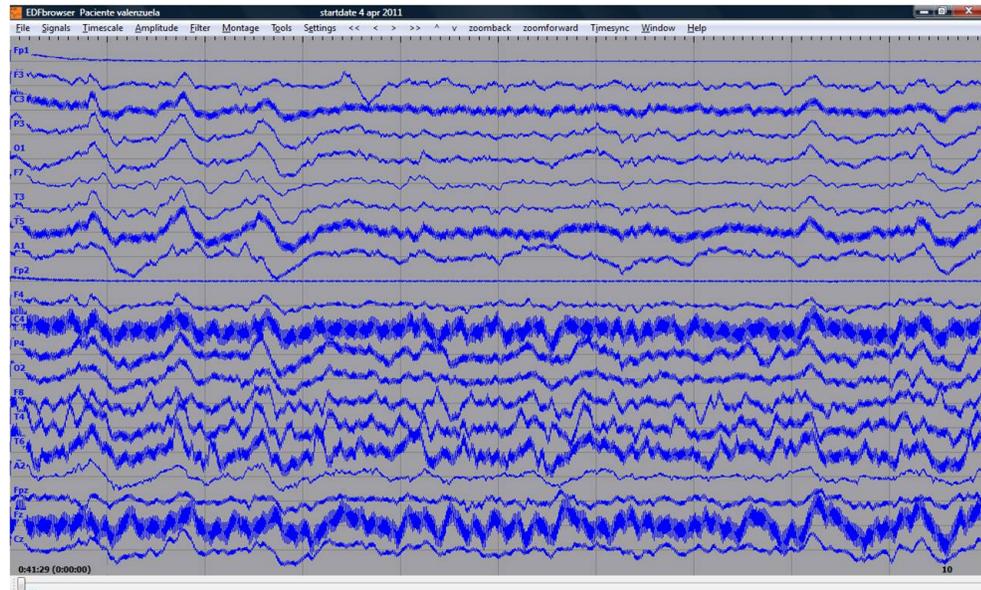


Figura 6.4: EEG filtrado

artificios oculares. Una vez separada la señal en sus componentes independientes, se procede a graficar los mapas de actividad del cerebro y así poder verificar cuál de ellos afectan directamente las partes frontales del cerebro.

Finalmente una vez eliminados los componentes considerados como artificios oculares, se debe proceder a la etapa de reconstrucción de la señal libre de artificios oculares. La Figura 6.8 muestra un segmento de 10 segundos libre de artificios oculares.

6.5. Comentarios

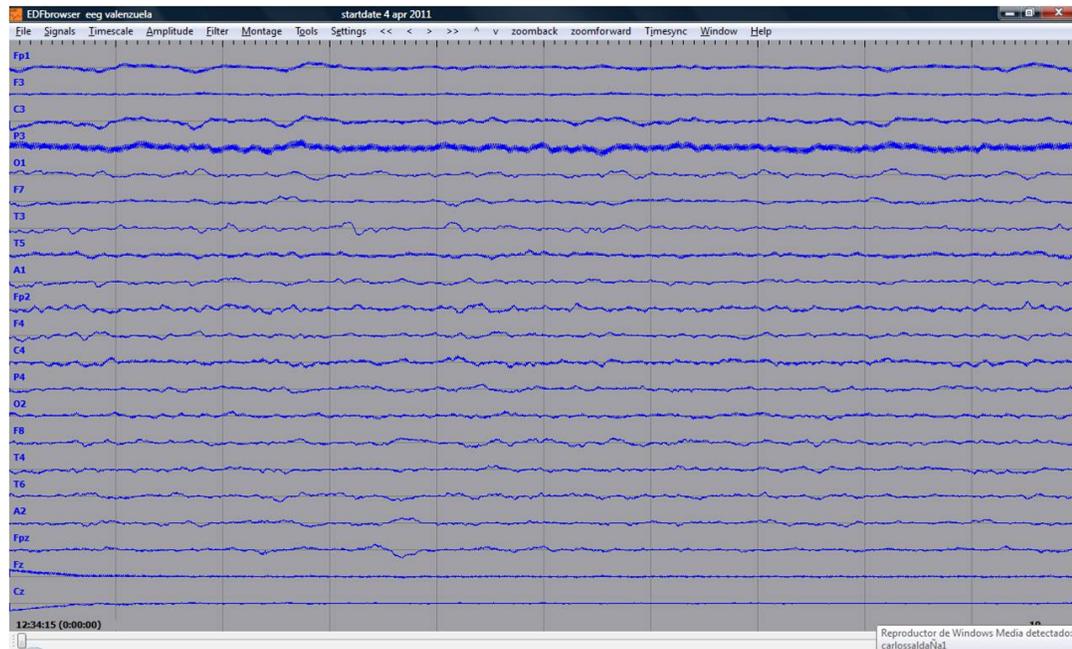


Figura 6.5: Componentes independientes después de aplicar INFOMAX extendido

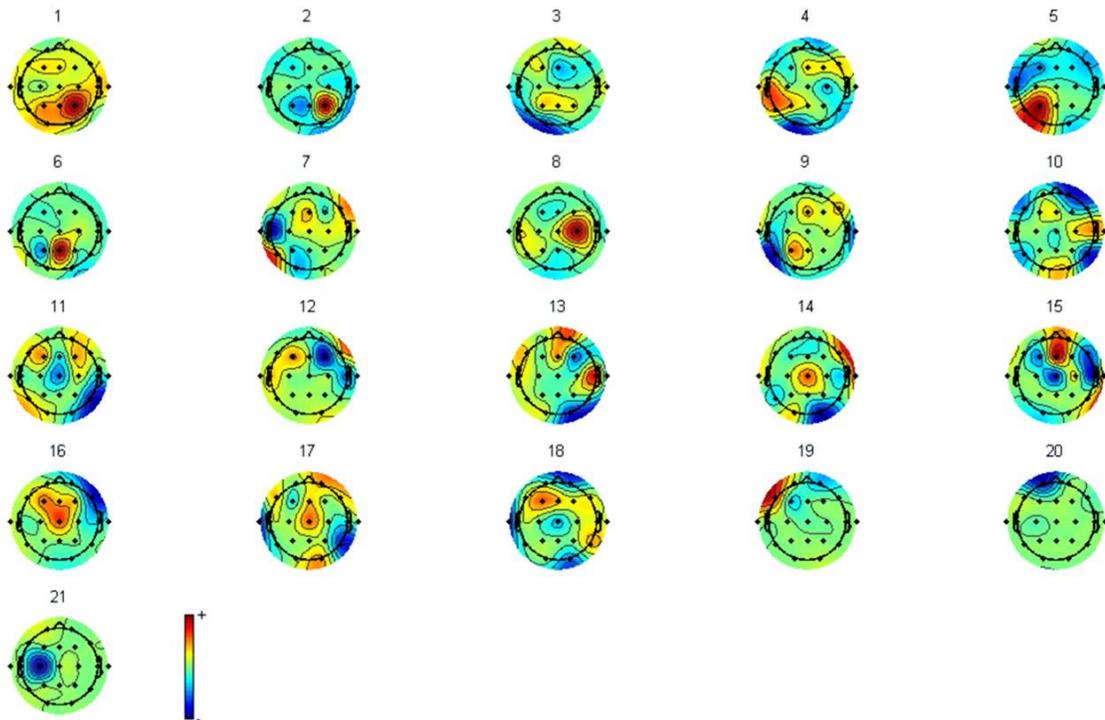


Figura 6.6: Mapas de actividad de los componentes independientes

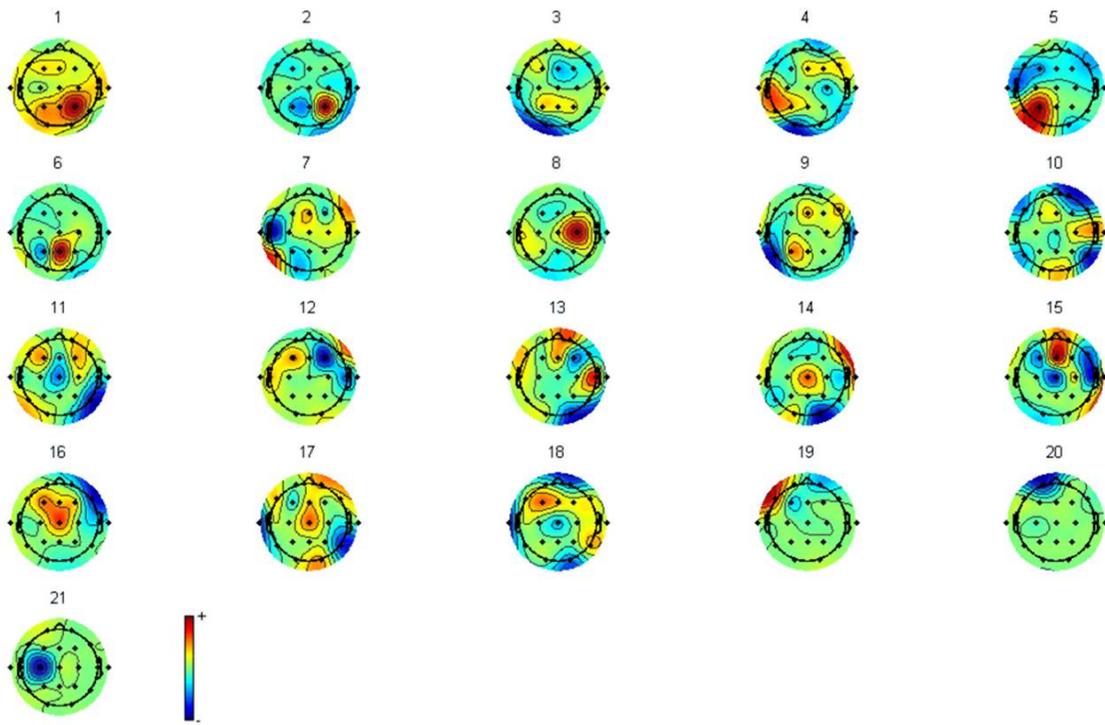


Figura 6.7: Mapas de actividad de los componentes independientes

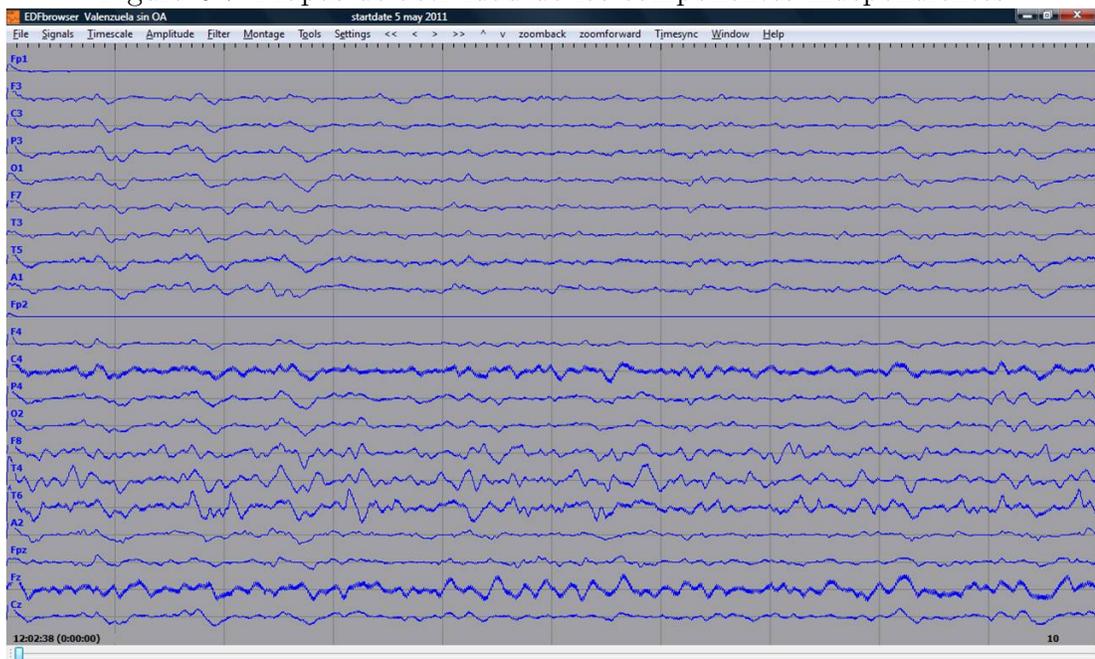


Figura 6.8: EEG libre de artificios oculares

Capítulo 7

Conclusiones y perspectivas

La estructura de un EEG neonatal es muy diferente a la de un EEG de un niño mayor y a la de un adulto. Por esta razón, durante el período de maduración, el EEG debe ser interpretado cuidadosamente en base a normas estadísticas. De acuerdo con lo anterior, trabajar con el tratamiento de señales EEG neonatales no es una tarea sencilla.

La pregunta planteada en esta tesis fue que si era posible la automatización de la detección y eliminación de artificios oculares en registros EEG neonatales, para lo cual se tiene una metodología semi-automática para dicho problema abordado, ya que se realizó la investigación de las técnicas de procesamiento de señales basadas en BSS mediante modelos neuronales artificiales, con las cuales, fue posible encontrar los componentes, los cuales son considerados como artificios oculares.

En este trabajo se presenta una metodología para la detección y eliminación de artificios oculares presentes en registros EEG neonatales. La metodología propuesta consta de las siguientes etapas: a) preparación de datos; b) pre-procesamiento ICA, c) detección de artificios oculares, d) eliminación de componentes identificados como artificios oculares y; e) reestructuración de la señal.

Del análisis de los resultados experimentales se puede concluir que la etapa de preparación de datos de entrada, utilizando un centrado y un blanqueo de datos, facilita la separación de las señales de entrada y por consiguiente la eliminación de artificios.

De acuerdo con los resultados obtenidos en experimentaciones realizadas, ICA es

una técnica aplicable y efectiva para la eliminación una amplia variedad de artificios en los registros EEG neonatales.

De acuerdo a la experimentación realizada se puede concluir lo siguiente:

- a. Se determinó la tasa de aprendizaje para el algoritmo INFOMAX extendido, la cual es 0,000004.
- b. La etapa de pre-procesamiento de los datos antes de la aplicación de INFOMAX extendido, es de vital importancia para obtener mejores resultados.
- c. Es importante el manejo del tamaño de la ventana analizado, debido a que con ello se puede determinar el tamaño idoneo para la detección y eliminación de artificios oculares. De acuerdo con los resultados de INFOMAX extendido se puede concluir que es suficiente el registro de por lo menos 60 segundos. Es decir, con al menos 60 segundos de registro EEG neonatal sería suficiente para la detección y eliminación de los artificios oculares.

7.1. Aportaciones

El presente trabajo de investigación dió como resultados los siguientes puntos:

- a. Una metodología semi-automática para la detección y eliminación de artificios oculares en registros EEG neonatales.
- b. El valor de la tasa de aprendizaje para el algoritmo INFOMAX extendido para problemas en EEG neonatales.
- c. Un proceso para la preparación de los datos, el cual incluye la estandarización de los formatos EEG a formato ASCII, además de un proceso de filtrado.
- d. Documentación de la técnica de ICA, esto es debido a que en la literatura se presenta la técnica dando por hecho que el lector tiene conocimientos matemáticos y probabilísticos previos. En este caso, se presenta los fundamentos necesarios para un mejor entendimiento de ICA.
- e. Implementación de INFOMAX extendido, basado en la implementación realizada en la tesis doctoral del Dr. Luz Noé Oliva Moreno [38].

7.2. Trabajos futuros

Como trabajos futuros de esta tesis, se mencionan algunos de ellos:

- a. La evaluación de los resultados por parte de un especialista en neurología pediátrica.
- b. Detección y eliminación de otros tipos de artificios fisiológicos como los musculares, cardíacos, etc.
- c. Implementación de otras técnicas para la detección y eliminación de artificios oculares en EEG neonatales.
- d. Evaluación del método propuesto con registros EEG neonatales diferentes a los registros proporcionados por el Hospital Infantil de Tamaulipas.
- e. Documentación y evaluación del modelo de filtrado y realizar una comparación, esto con la finalidad de determinar con cual de los modelos se tienen mejores resultados.
- f. Aplicación y análisis de INFOMAX extendido sobre registros EEG en algun montaje como neonatal, referencial, entre otros.
- g. Desarrollo de un software no comercial que realice la detección y eliminación de artificios fisiológicos.

Apéndice A

Software EDFBrowser

Es una herramienta y visualizador multiplataforma y de código fuente libre hecho para el tratamiento de series de tiempo como EEG, EMG, ECG, Bioimpedancia, entre otros [52].

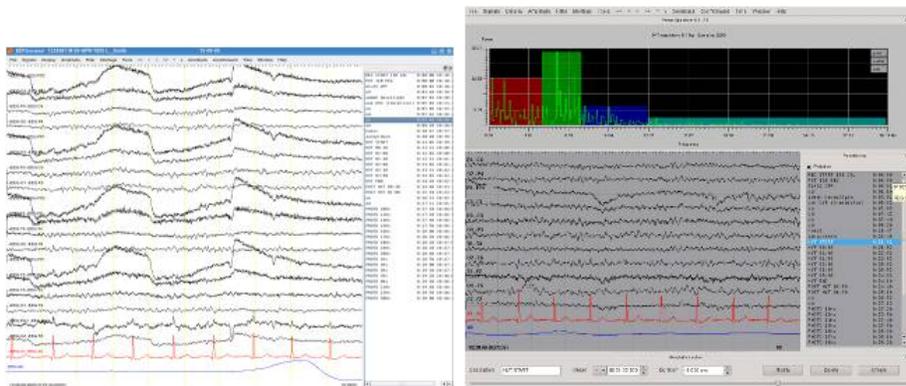


Figura A.1: Ejemplos de áreas de trabajo del software EDFBrowser

A.1. Características

A continuación se enlistan algunas de las características que posee el software:

- * Soporta diferentes formatos de archivos: EDF, EDF+, BDF, BDF+.
- * Convierte del formato Nihon Kohden a EDF+ (incluyendo las anotaciones).

- * Convierte del formato Emsa (*.PLG) a EDF+ (incluyendo las anotaciones).
- * Convierte del formato ASCII a EDF/BDF.
- * Convierte del formato Finometer (Beatscope) a EDF.
- * Convierte del formato Bmeye Nexfin (FrameInspector) a EDF.
- * Convierte del formato WAV a EDF.
- * Editor de anotaciones.
- * Editor de los encabezados.
- * Diferentes modelos de filtrado como son: Butterworth, Chebyshev y Bessel, de orden uno hasta el orden ocho.
- * Poder espectral.
- * Soporte para diferentes montajes.
- * Visualiza señales de diferentes archivos al mismo tiempo.
- * Convierte del formato EDF/EDF+/BDF/BDF+ a ASCII.
- * Convierte del formato EDF+D a EDF+C.
- * Convierte del formato BDF(+) a EDF(+).
- * Disponible para Linux y para Windows.

A.2. Comentarios

El software EDFBrowser es una herramienta de suma importancia para la realización de la presente tesis. Fue por medio de este software se pudo visualizar los archivos EEG de neonatos, así como la aplicación del filtrado y la conversión de los archivos EEG a un formato ASCII para poder ser procesados y analizados con las herramientas disponibles.

Apéndice B

Conceptos de estadística

En este apéndice se muestran los conceptos básicos de teoría de probabilidad, estadística y procesos aleatorios. Dichos conceptos son utilizados en los métodos basados en BSS. Para mayor información se pueden consultar la siguiente literatura [13, 21, 28, 43].

La teoría de probabilidad trata con el promedio del fenómeno de las masas ocurriendo secuencialmente o simultáneamente. Se ha observado que ciertos promedios aproximan a un valor constante tanto como el número de observaciones incrementa. El objetivo de este promedio es describir y predecir los eventos sucesivos en términos de probabilidades.

En el análisis de una señal aleatoria, es de interés el estudio de su Función de Densidad de probabilidad y su función de distribución de probabilidad, ya que proporcionan información sobre la ocurrencia de algún valor determinado.

Se dice que una variable aleatoria que puede asumir cualquier valor dentro de algún intervalo especificado, es una *variable continua*. Sin embargo, una *variable aleatoria discreta* que solo pueda asumir ciertos valores, puede ser tratada con el mismo método.

B.1. Función de Distribución de Probabilidad o Función de Distribución

Sea X una variable aleatoria y x cualquier valor permitido (o evento) de esta variable. La función de distribución de probabilidad es definida como la probabilidad acumulada de que el evento x de la variable aleatoria observada X ocurra, por lo tanto la probabilidad es menor o igual que el valor permitido x , esto es:

$$P_X(x) = P_r(X \leq x) \tag{B.1}$$

Puesto que la función de distribución de probabilidad es una probabilidad, debe satisfacer los siguientes axiomas básicos:

1. $0 \leq P_X(x) \leq 1 \quad -\infty < x < \infty$
2. $P_X(-\infty) = 0 \quad P_X(\infty) = 1$
3. $P_X(x)$ es no decreciente tanto como x incrementa
4. $P_r(x_1 < X < x_2) = P_X(x_2) - P_X(x_1)$

Algunas funciones de distribución mostradas en la Figura B.1. Indicando una variable aleatoria continua con posibles valores en el intervalo de $-\infty$ a $+\infty$.

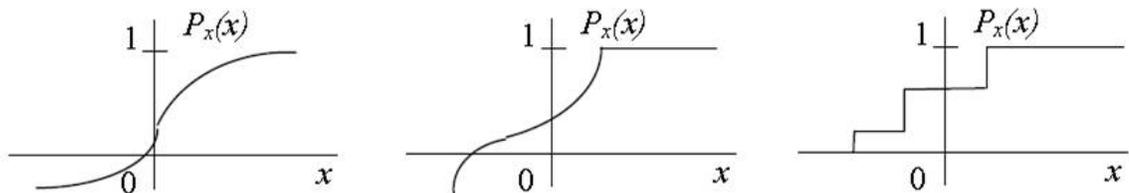


Figura B.1: Ejemplos de posibles funciones de distribución de probabilidad

B.2. Función de Densidad de probabilidad

También conocida como la distribución de probabilidad. La cual como su nombre lo indica, define como se distribuye la probabilidad a través de la densidad de repeti-

ción de una variable aleatoria. Esta forma es más conveniente para muchos cálculos de interés.

Esta función de densidad se puede obtener a través de tomar la derivada de la función de distribución de probabilidad ($P_x(x)$).

$$px(x) = \lim_{\epsilon \rightarrow 0} \frac{P_x(x + \epsilon) - P_x(x)}{\epsilon} = \frac{dP_x(x)}{dx} \tag{B.2}$$

Rescribiendo la ecuación B.2, la función de densidad de probabilidad puede ser interpretada como:

$$px(x)dx = P_r(x < X \leq x + dx) \tag{B.3}$$

La ecuación B.3 establece que la probabilidad del elemento, $px(x)dx$, es la probabilidad de que el evento de la variable aleatoria X se encuentre en el intervalo de posibles valores de x y $x + dx$.

Puesto que $px(x)$ es una función de densidad de probabilidad y no una probabilidad, no es necesario que sus valores sean menor que 1; pero no negativos. Sus propiedades se pueden resumir en los siguientes axiomas:

1. $px(x) \geq 0 \quad -\infty < x < +\infty$
2. $\int_{-\infty}^{\infty} px(x) dx = 1$
3. $P_x(x) = \int_{-\infty}^x px(u) du$
4. $\int_{x_1}^{x_2} px(x) dx = P_r(x_1 < X \leq x_2)$

Algunas funciones de densidad de probabilidad son mostradas en la figura B.2:

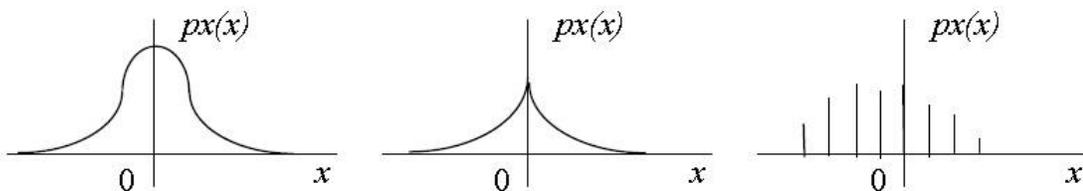


Figura B.2: Ejemplos de posibles funciones de densidad de probabilidad

Particularmente la función de densidad para una variable aleatoria discreta consiste de un conjunto de funciones delta, cada una teniendo un área igual a su magnitud.

Existen muchas posibles funciones de densidad pero algunas cuantas de estas, tienen un significado en el análisis de procesamiento de señales.

B.3. Valor medio y valor cuadrático medio

Uno de los conceptos más importantes asociados en estadística es el valor medio o promedio de una variable o variables aleatorias. El valor medio o promedio de una variable que esta en función del tiempo, se obtiene integrando sobre todo el intervalo del tiempo que se toma la variable; otra forma de tomar el promedio de una variable, es considerar el número de muestras y promediarlas sobre el número de éstas.

Diferentes notaciones son utilizadas para encontrar el valor medio. En ingeniería el símbolo más empleado es $E\{X\}$ que usualmente se lee el valor esperado de X o la esperanza matemática de X .

$$\bar{X} = E\{X\} = \int_{-\infty}^{\infty} xp(x) dx \quad (\text{B.4})$$

El valor esperado de cualquier función puede obtenerse de forma similar:

$$E\{F(X)\} = \int_{-\infty}^{\infty} f(x)p(x) dx \quad (\text{B.5})$$

Una función de particular importancia es $f(x) = x^n$, puesto que esta función nos permite generar los momentos de una variable aleatoria.

$$\bar{X}^n = E\{X^n\} = \int_{-\infty}^{\infty} x^n p(x) dx \quad (\text{B.6})$$

Los momentos más importantes son para $n = 1$ y $n = 2$ el cual genera el valor medio (B.4) y el valor cuadrático medio (B.7), respectivamente.

$$\bar{X}^2 = E\{X^2\} = \int_{-\infty}^{\infty} x^2 p(x) dx \quad (\text{B.7})$$

B.4. Varianza y Desviación estándar

Otro de los conceptos importantes en estadística es el momento central, el cual se define como la diferencia entre una variable aleatoria y su valor medio. Por lo tanto el n -ésimo momento central es:

$$\overline{X - \bar{X}}^n = E\{\{X - \bar{X}\}^n\} = \int_{-\infty}^{\infty} (x - \bar{X})^n p(x) dx \quad (\text{B.8})$$

El momento central para $n = 1$ es cero, mientras que para $n = 2$ se le denomina la varianza, usualmente simbolizada por σ^2 .

$$\sigma^2 = \overline{X - \bar{X}}^2 = \overline{X^2} - \bar{X}^2 \quad (\text{B.9})$$

Por esta razón se define la varianza, como la diferencia entre el valor cuadrático medio y el cuadrado del valor medio. La raíz cuadrada de la varianza se conoce como la desviación estándar σ .

Algunas de las desventajas de los momentos, es que existen distribuciones para el cual los momentos no son finitos y que el conocimiento de estos no necesariamente especifica la función de densidad de probabilidad. Sin embargo, para muchas de las distribuciones los momentos no cumplen con estas condiciones.

B.5. Cumulantes

Los cumulantes son otras constantes que describen y especifican las propiedades y las características de una distribución y son más útiles, desde el punto de vista teórico, que los momentos. Los cumulantes de orden k pueden calcularse a partir de los momentos de igual orden o inferior a k . Los cumulantes del orden 1 al 4, respecto al origen y en función de los momentos son:

$$K_1 = \bar{X} (\text{Media}) \quad (\text{B.10})$$

$$K_2 = \overline{X^2} - \bar{X}^2 (\text{Varianza}) \quad (\text{B.11})$$

$$K_3 = \overline{X^3} - 3\overline{X^2}\overline{X} + 2\overline{X}^3 \quad (\text{B.12})$$

$$K_4 = \overline{X^4} - 4\overline{X^3}\overline{X} - 3\overline{X^2}^2 + 12\overline{X^2}\overline{X}^2 - 6\overline{X}^4 \quad (\text{B.13})$$

Si las variables aleatorias están centradas en su media, es decir $\overline{X} = 0$, las expresiones de los cumulantes se simplifican en:

$$K_1 = 0 \quad (\text{B.14})$$

$$K_2 = \overline{X^2} \quad (\text{B.15})$$

$$K_3 = \overline{X^3} (\textit{skewness}) \quad (\text{B.16})$$

$$K_4 = \overline{X^4} - 3\overline{X^2}^2 (\textit{kurtosis}) \quad (\text{B.17})$$

En donde el cumulante normalizado de tercer orden determina la asimetría (*skewness*) de una distribución y el de cuarto orden determina la agudeza o curtosis (*kurtosis*). Para una variable con distribución normal (gaussiana) los cumulantes de orden superior a 2 son cero; es decir $K_K = 0 (\forall K > 2)$. Se verá más adelante la utilidad de cada uno de los cumulantes de orden superior para una distribución normal.

B.6. Variable aleatoria Gaussiana o Normal

Una de las funciones de densidad más estudiadas es la Gaussiana o normal. Alguna de las razones para su importancia, son:

1. Esta proporciona un buen modelo matemático para muchos fenómenos aleatorios observados físicamente.
2. La combinación lineal de variables aleatorias Gaussianas da como resultado una nueva variable aleatoria Gaussiana.

3. Los procesos aleatorios del cual una variable aleatoria Gaussiana es derivado, puede ser completamente especificado, en el sentido estadístico, esto es, el conocimiento de su primero y segundo momento.
4. En el análisis del sistema, el proceso Gaussiano es frecuentemente el único para el cual un análisis estadístico completo puede ser realizado en una situación lineal y no lineal.

La representación matemática de la función de densidad Gaussiana es:

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left[\frac{-(x - \bar{X})^2}{2\sigma^2} \right] \quad -\infty < x < \infty \quad (\text{B.18})$$

Donde \bar{X} y σ^2 son la media y la varianza, respectivamente. La figura B.3 muestra la función de densidad y su correspondiente función de distribución de una variable aleatoria con media cero ($\bar{X} = 0$) y varianza unitaria ($\sigma = \sigma^2 = 1$). Algunas características de este tipo de distribución, son las siguientes:

1. Tienen un unico máximo y este ocurre en el valor de la media.
2. La funcion de densidad de probabilidad es simétrica alrededor de su media.
3. El ancho de la distribución, es directamente proporcional a la desviación estandar σ .
4. El valor máximo de la función de densidad es inversamente proporcional a la desviación estandar σ .

Usualmente la función de densidad de probabilidad para este tipo de distribución se considera normalizada, es decir, con media cero y varianza unitaria ($\bar{X} = 0, \sigma^2 = 1$) y se define como:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp \left[-\frac{x^2}{2} \right] dx \quad (\text{B.19})$$

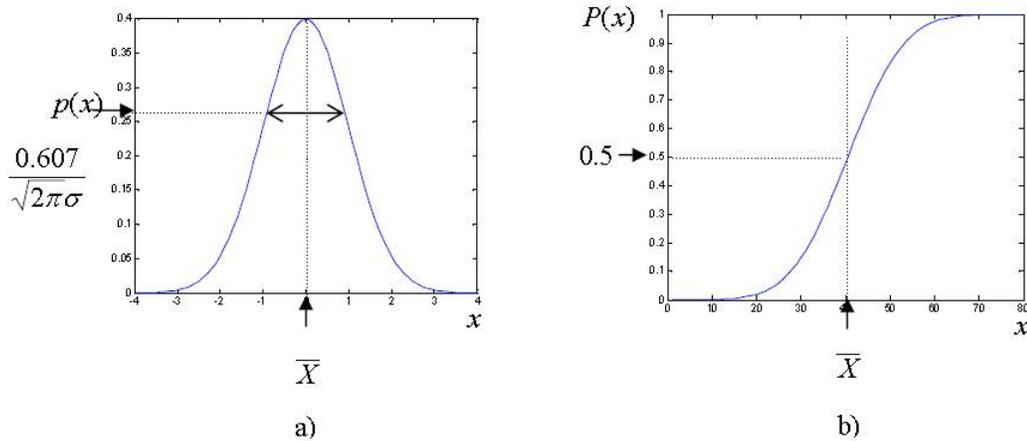


Figura B.3: Variable aleatoria Gaussiana: a)función de densidad y b)función de probabilidad

Donde $\Phi(x)$ representa esta función de distribución normalizada. Efectuando un cambio de variable, de la ecuación B.18, se obtiene que:

$$\Phi(x) = \Phi \left[-\frac{x - \bar{X}}{\sigma} \right] \tag{B.20}$$

El n -ésimo momento central de una variable aleatoria gaussiana con media cero y varianza unitaria, puede expresarse como:

$$\begin{aligned} \overline{(X - \bar{X})^n} &= 0 \text{ } n \text{ impar} \\ &= (\sigma^2 = 1), (3\sigma^4 = 3), (5\sigma^6 = 5), \dots, (n - 1)\sigma^n \text{ } n \text{ par} \end{aligned} \tag{B.21}$$

Bajo ciertas condiciones, que suelen cumplir la mayoría de las funciones de distribución, si dos variables aleatorias, tienen sus respectivos momentos iguales (con que tengan los 3 o 4 primeros momentos iguales), se puede considerar que sus funciones son idénticas.

Considerando el centrado de la media en cero y normalizando la varianza de la señal, los cumulantes, efectuando ($\bar{X} = 0, \sigma^2 = 1$), es decir, efectuando el siguiente cambio de variable.

$$s = \frac{x - \bar{X}}{\sigma} \quad (\text{B.22})$$

Donde s es la señal normalizada; es decir que tiene media cero y varianza unitaria. Los cumulantes normalizados K_K , son los cumulantes de las señales estandarizadas. El cumulante normalizado de orden 3 (K_3), mide la asimetría de la función de distribución ya que es cero si la función de densidad de probabilidad de x posee un eje de simetría. Si x esta centrada, el factor de asimetría (γ_1) resulta ser:

$$\gamma_1 = \tilde{\chi}_3 = \frac{E\{s^3\}}{E\{s^2\}^{3/2}} = \frac{\mu_3}{\mu_2^{3/2}} = \frac{k_3}{k_2^{3/2}} \quad (\text{B.23})$$

Si la distribución es simétrica: $\gamma_1 = 0$ (ya que toda distribución de probabilidad simétrica tiene los momentos impares nulos, y en particular $\mu_3 = 0$).

El cumulante estandarizado de orden 4, \tilde{K}_4 , es el factor de agudeza o curtosis (γ_2). Si la señal está centrada en la media:

$$\gamma_2 = \tilde{K}_4 = \frac{E\{s^4\}}{E\{s^2\}^2} - 3 = \frac{\mu_4}{\mu_2^2} - 3 = \frac{k_4}{k_2^2} \quad (\text{B.24})$$

Si $\gamma_2 = 0$ se dice que la distribución es mesoaguda (gaussiana). Si la agudeza es negativa ($\gamma_2 < 0$) se dice que la distribución es *sub-gaussiana* (ó *platicúrtica*), ya que tiende a cero en el infinito más rápidamente que la distribución gaussiana. Por el contrario, si la agudeza es positiva ($\gamma_2 > 0$), la función de distribución tiende a cero en el infinito más lentamente que la distribución gaussiana, se dice que es *super-gaussiana* (ó *leptocúrtica*).

B.7. Múltiples variables aleatorias

El término múltiples, considera que se trabaja con dos o mas variables aleatorias, pero por simplicidad se consideran solo dos variables (extendiendo el mismo estudio a un grupo mayor). Definiendo su función de distribución de x y y como:

$$P(X) = P[X \leq x] \quad (\text{B.25})$$

$$P(Y) = P[Y \leq y] \quad (\text{B.26})$$

Estas distribuciones de probabilidad, también se le conoce como probabilidades marginales. Expresando la densidad de probabilidad conjunta (de juntura) de las dos variables se define como:

$$P(X, Y) = P[X \leq x, Y \leq y] \quad (\text{B.27})$$

Algunas de las propiedades análogas de la probabilidad conjunta con respecto a la distribución de una sola variable se muestra a continuación:

1. $0 \leq P_X(x, y) \leq 1 \quad -\infty < x < \infty$
2. $P(-\infty, y) = P(x, -\infty) = P(-\infty, -\infty) = 0$
3. $P(\infty, \infty) = 1$
4. $P(x, y)$ es no decreciente tanto como x ó y incrementa
5. $P(\infty, y) = P(y) \quad P(x, \infty) = P(x)$

El valor esperado de la función de densidad de probabilidad conjunta (o de juntura) de dos variables aleatorias puede ser determinado por las siguiente expresión:

$$E[f(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)p(x, y) dx dy \quad (\text{B.28})$$

Considerando el caso general de $f(X, Y) = XY$. La ecuación B.28 se puede representar como:

$$E[XY] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xyp(x, y) dx dy \quad (\text{B.29})$$

Esta expresión determina la correlación de dos variables aleatorias.

Las propiedades de la función de distribución conjunta son análogas a aquellas de una sola variable, y se mencionan a continuación:

1. $p(x, y) \geq 0 \quad -\infty < x < +\infty \quad -\infty < y < +\infty$
2. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x, y) dx dy = 1$
3. $P(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(u, v) du dv$

$$4. p(x) = \int_{-\infty}^{\infty} p(x, y) dy \quad p(y) = \int_{-\infty}^{\infty} p(x, y) dx$$

$$5. P[x_1 < X \leq x_2, y_1 < Y \leq y_2] = \int_{x_1}^{x_2} \int_{y_1}^{y_2} p(x, y) dy dx$$

También es posible determinar la función de densidad conjunta derivando la función de distribución. Puesto que son dos variables independientes la derivada obtendrá un resultado parcial.

Otro concepto es la probabilidad condicional. Esta es la probabilidad de un evento X , dado que otro evento Y ha ocurrido y se expresa como $P(X|Y)$. Entonces la probabilidad de un evento X , dado un evento Y , se define por la ecuación B.30.

$$P(X|Y) = \frac{P(X, Y)}{P(Y)} \quad (\text{B.30})$$

Análogamente se define $P(Y|X)$

$$P(Y|X) = \frac{P(Y, X)}{P(X)} \quad (\text{B.31})$$

donde $P(X, Y) = P(Y, X)$ es la probabilidad conjunta del evento X y Y . Y $P(X)$ ó $P(Y)$ es la probabilidad marginal de los eventos particulares de X y Y , respectivamente. Si el evento X esta contenido en Y ($X \subset Y$) entonces $(X, Y) = X$ y la probabilidad $P(X|Y)$ esta dado por la ecuación B.32.

$$P(X|Y) = \frac{P(X)}{P(Y)} \geq P(X) \quad (\text{B.32})$$

Si ($Y \subset X$), entonces $(X, Y) = Y$ y

$$P(X|Y) = \frac{P(Y)}{P(Y)} = 1 \quad (\text{B.33})$$

La Figura B.4 muestra la distribución de densidad de probabilidad de x y y , en la cual se denota la densidad marginal, de densidad condicional ($y|x$) y conjunta (de juntura) (y, x).

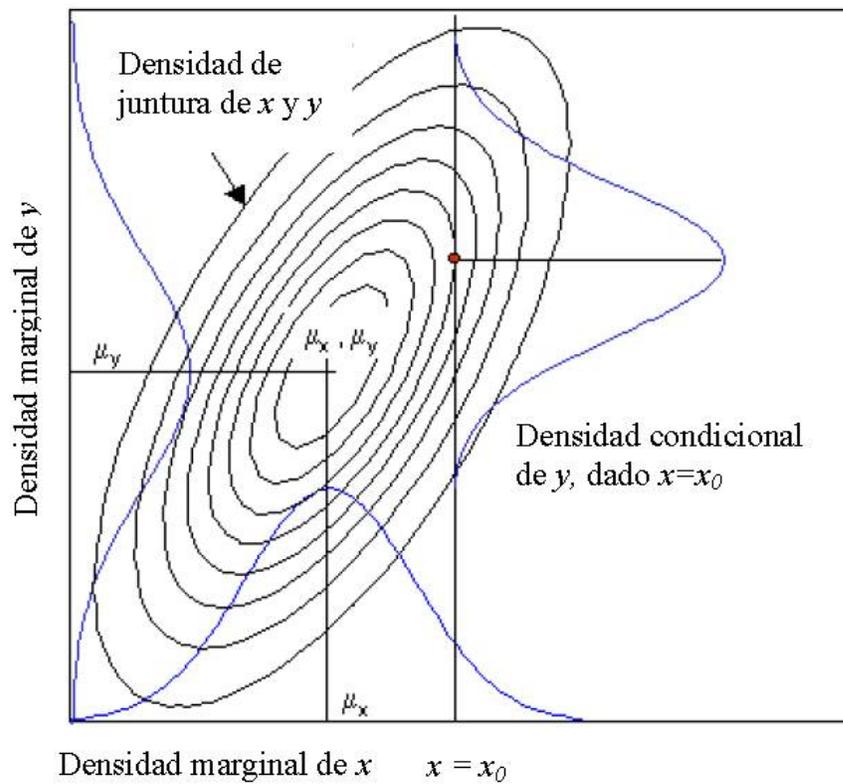


Figura B.4: Distribución de densidades de probabilidad, marginales, condicionales y conjunta (de juntura)

B.8. Independencia estadística

Considerando el caso de dos variables aleatorias X y Y . Cuando dos variables aleatorias son estadísticamente independientes, el conocimiento de una variable aleatoria no da información acerca del valor de la otra y matemáticamente en términos de probabilidad, la independencia estadística se define como:

$$P(X, Y) = P(X)P(Y) \tag{B.34}$$

En general si existe n eventos, la expresión B.35 se representa como:

$$P(X, Y, \dots, n) = p(X)p(Y), \dots, p(n) \tag{B.35}$$

En el caso que se tome cualquier función, se satisface la misma propiedad básica.

$$E\{g(X)h(Y)\} = E\{g(X)\}E\{h(Y)\} \quad (\text{B.36})$$

Donde $g(X)$ y $h(Y)$ son cualquier función integrable.

$$\begin{aligned} E\{g(X)\}E\{h(Y)\} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x)h(y)p_{XY}(x, y) dy dx \\ &= \int_{-\infty}^{\infty} g(x)p_X(x) dx \int_{-\infty}^{\infty} h(y)p_Y(y) dy = E\{g(X)\}E\{h(Y)\} \end{aligned} \quad (\text{B.37})$$

La misma condición de independencia estadística se cumple para el caso multi-dimensional

$$E\{g(X)g(Y)g(Z)\dots\} = E\{g(X)\}E\{g(Y)\}E\{g(Z)\}\dots \quad (\text{B.38})$$

Donde $g()$ son funciones arbitrarias.

Como se mencionó anteriormente la probabilidad condicional se representa como:

$$P(Y, X) = P(X, Y) = P(X|Y)P(Y) = P(Y|X)P(X) \quad (\text{B.39})$$

Despejando de los dos últimos términos de la ecuación B.39, el término $P(X|Y)$:

$$P(X, Y) = \frac{P(Y|X)P(X)}{P(Y)} \quad P(Y) \neq 0 \quad (\text{B.40})$$

El término $P(Y)$ se puede representar como:

$$P(Y) = P(Y|X_1)P(X_1) + P(Y|X_2)P(X_2) + \dots + P(Y|X_n)P(X_n) \quad (\text{B.41})$$

Sustituyendo la ecuación B.41 en B.40, se obtiene:

$$P(X|Y) = \frac{P(Y|X)P(X)}{P(Y) = P(Y|X_1)P(X_1) + P(Y|X_2)P(X_2) + \dots + P(Y|X_n)P(X_n)} \quad (\text{B.42})$$

La ecuación B.40 o la B.42 son referidas como el *Teorema de Bayes*. Por lo tanto, la expresión de $P(Y|X)$, se deduce como:

$$P(Y|X) = \frac{P(X|Y)P(Y)}{P(X)} \quad P(X) \neq 0 \quad (\text{B.43})$$

Si las dos variables (X y Y) son independientes, se obtiene su probabilidad marginal, esto es:

$$P(Y|X) = \frac{P(X|Y)P(Y)}{P(X)} = \frac{P(X, Y)}{P(X)} = \frac{P(X)P(Y)}{P(X)} = P(Y) \quad (\text{B.44})$$

Equivalentemente:

$$P(X|Y) = \frac{P(Y|X)P(X)}{P(Y)} = P(X) \quad (\text{B.45})$$

B.9. Teorema de límite central

El teorema del límite central o teorema central del límite concierne a la suma de un gran número de variables aleatorias independientes con mismas funciones de densidad de probabilidad. Del cual el resultado se aproximará a una función de densidad Gaussiana, tanto como el número de variables independientes se incremente. Esta suma queda definida como:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_n \quad (\text{B.46})$$

En donde X_n , son las variables aleatorias independientes. El teorema del límite central tiene una importante consecuencia en el análisis de componentes independientes.

B.10. Correlación

La correlación indica la fuerza y la dirección de una relación lineal entre dos variables aleatorias. Se considera que dos variables aleatorias cuantitativas están correlacionadas cuando los valores de una de ellas varían sistemáticamente con respecto

a los valores homónimos de la otra.

Si dos variables aleatorias X y Y tienen los valores posibles x y y , entonces el valor esperado de su producto es conocido como la correlación, definido como:

$$E\{X, Y\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xyp(x, y) dx dy = \overline{XY} \quad (\text{B.47})$$

La correlación de dos variables aleatorias independientes continuas es igual a sus probabilidades marginales.

$$\begin{aligned} E\{X, Y\} &= \int_{-\infty}^{\infty} xp(x) dx \int_{-\infty}^{\infty} yp(y) dy \\ &= E\{X\}E\{Y\} = \overline{XY} \end{aligned} \quad (\text{B.48})$$

Si dos variables aleatorias tienen su media diferente de cero, es conveniente restarle este valor.

$$E\{(X - \bar{X})(Y - \bar{Y})\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \bar{X})(y - \bar{Y}) dx dy \quad (\text{B.49})$$

Esta ecuación es conocida como la *covarianza*, análogamente a la varianza de una sola variable.

El coeficiente de correlación o covarianza normalizada es un índice estadístico (entre -1 y +1) que indica la relación lineal entre dos variables aleatorias. El cálculo del coeficiente de correlación lineal se realiza dividiendo la covarianza por el producto de las desviaciones estándar de ambas variables:

$$r = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} \quad (\text{B.50})$$

El coeficiente de correlación es una forma simple de la covarianza normalizada. Note que si dos variables aleatorias de media cero no están correlacionadas, entonces su covarianza es igual a 0. Sí $r = 1$ se tiene una total dependencia de las dos variables. Sí $r = -1$ existe una dependencia negativa, es decir, que cuando una de ellas aumenta, la otra disminuye en idéntica proporción.

Apéndice C

Conceptos de álgebra lineal

El álgebra lineal, es la rama de la matemática que concierne al estudio de vectores, espacios vectoriales, transformaciones lineales y sistemas de ecuaciones lineales. Los espacios vectoriales son un tema central en la matemática moderna; por lo que el álgebra lineal es usada ampliamente en álgebra abstracta y análisis funcional. El álgebra lineal considera espacios de dimensión arbitraria o incluso de dimensión infinita. Un espacio vectorial de dimensión n se dice que es n -dimensional. La mayoría de los resultados encontrados en 2 y 3 dimensiones pueden extenderse al caso n -dimensional.

En este apéndice, se darán algunos conceptos básicos del álgebra lineal. Estos conceptos son utilizados en algunos métodos de ICA. Para mayor información puede consultarse en [23, 36, 50].

C.1. Sistema de ecuaciones lineales

Un sistema lineal, es un sistema que cumple las propiedades de escalado y de superposición y se describe con la siguiente ecuación:

$$a_1x_1 + a_2x_2 + \dots + a_nx_n = b \tag{C.1}$$

Donde a_1, a_2, \dots, a_n y b_n son números y x_1, x_2, \dots, x_n son variables. Un sistema lineal de m ecuaciones y n incógnitas es un sistema de la forma:

$$\begin{aligned}
 a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\
 a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\
 &\vdots \\
 a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= b_m
 \end{aligned}
 \tag{C.2}$$

A este sistema se le conoce como sistema de ecuaciones lineales. Una forma de representar a este conjunto de ecuaciones es mediante la notación de Matrices y Vectores.

C.2. Matrices y vectores

Denotando el sistema de ecuaciones (C.2) C.3 en forma matricial y vectorial, se tiene de la forma:

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}
 \tag{C.3}$$

Simplificando la notación de C.3, se tiene que:

$$\mathbf{Ax}=\mathbf{b}
 \tag{C.4}$$

Donde \mathbf{A} es una matriz de dimensión $m \times n$, donde n son las columnas y m las filas, \mathbf{x} es un vector fila de dimensión n y \mathbf{b} es el vector columna de dimensión m . Cuando una matriz es de dimensión $n \times n$, se dice que es una matriz cuadrada de dimensión n . Cabe hacer notar que las matrices son denotadas con mayúsculas y en negrita, a diferencia de los vectores que son denotados en minúscula.

Un vector fila de dimensión m es una matriz de dimensión $1 \times m$, cuyos elementos se representan como una sucesión de datos en forma horizontal. Como por ejemplo, el vector fila (1×8) .

$$\mathbf{b} = [a_1 \ a_2 \ a_3 \ a_4 \ a_5 \ a_6 \ a_7 \ a_8]$$

Un vector columna de dimensión n es una matriz de dimensión $(n \times 1)$, cuyos elementos se representan como una sucesión de datos en forma vertical. Como por ejemplo, el vector columna (4×1) .

$$\mathbf{b} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} \quad (\text{C.5})$$

C.3. Matriz Identidad

La matriz de identidad \mathbf{I} , algunas veces denominada matriz unidad, es una matriz cuadrada en la cual los elementos situados sobre la diagonal principal son iguales a 1 y el resto de los elementos son iguales a 0.

$$\begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} \quad (\text{C.6})$$

La matriz identidad es una matriz que cumple la propiedad de ser el elemento neutro del producto de matrices. Esto quiere decir que el producto de cualquier matriz por la matriz identidad no tiene ningún efecto.

C.4. Producto punto de dos vectores

En matemáticas el producto escalar, también conocido como producto interno o producto punto, es una función definida sobre un espacio vectorial cuyo resultado es una magnitud escalar. El nombre espacio escalar se utiliza para denominar un espacio vectorial real sobre el que se ha definido una operación de producto interior que tiene como resultado un número real. El producto punto de \mathbf{a} y \mathbf{b} se define como:

$$\langle \mathbf{a}, \mathbf{b} \rangle = \sum_{i=1}^n a_i b_i \quad (\text{C.7})$$

De manera similar se define el producto punto de dos vectores columnas o de un vector fila y un vector columna. En todos los casos, los vectores deben tener el mismo número de componentes. El producto escalar en el caso particular de dos vectores en el plano, o en un espacio euclideo N -dimensional, se define como el producto de sus módulos multiplicado por el coseno del ángulo θ que forman.

C.5. Producto entre matrices

Dada una matriz \mathbf{A} y una matriz \mathbf{B} , el producto de estas dos matrices se puede realizar sólo si el número de columnas de la matriz \mathbf{A} es el mismo que el número de filas de la matriz \mathbf{B} , es decir sí \mathbf{A} es una matriz $m \times n$ y \mathbf{B} es una matriz $n \times p$, entonces su producto matricial \mathbf{AB} es la matriz $m \times p$ (m filas, p columnas) dada por:

$$c_{ij} = \sum_{k=1}^m \sum_{s=1}^m a_{ik} a_{sj} \quad (\text{C.8})$$

Esta expresión es el producto punto del i -ésima fila de \mathbf{A} y la j -ésima columna de \mathbf{B} .

C.6. Matriz invertible

Una matriz \mathbf{A} de dimensiones $n \times n$ se dice que es invertible, inversible, inversa o no singular si y solo si existe una matriz \mathbf{B} de dimensiones $n \times n$ (cuadrada) tal que $\mathbf{AB}=\mathbf{BA}=\mathbf{I}$. En tal caso, \mathbf{B} suele denominarse con \mathbf{A}^{-1} y se denomina como la inversa de \mathbf{A} .

Algunas propiedades de la matriz inversa, son:

$$(\mathbf{BA})^{-1} = \mathbf{A}^{-1}\mathbf{B}^{-1} \quad (\text{C.9})$$

Si la matriz es invertible, también lo es su transpuesta, y el inverso de su transpuesta es la transpuesta de su inversa, es decir:

$$(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T \quad (\text{C.10})$$

Y, evidentemente:

$$(\mathbf{A}^{-1})^{-1} = \mathbf{A} \quad (\text{C.11})$$

C.7. Determinante

El determinante de una matriz \mathbf{A} , es un escalar o polinomio, que resulta de obtener todos los productos posibles de una matriz de acuerdo a una serie de restricciones, siendo denotado como $|\mathbf{A}|$ o $\det(\mathbf{A})$. El valor numérico es conocido también como módulo de la matriz. A continuación se muestran unos ejemplos.

Ejemplo 1:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

El $\det(\mathbf{A}) = a_{11}a_{22} - a_{12}a_{21}$

Ejemplo 2:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

El $\det(\mathbf{A}) = a_{11} [a_{22}a_{33} - a_{23}a_{32}] - a_{12} [a_{21}a_{33} - a_{23}a_{31}] + a_{13} [a_{21}a_{32} - a_{22}a_{31}]$

Teorema 1 Sea \mathbf{A} una matriz cuadrada. \mathbf{A} es no singular si y solo si el $\det(\mathbf{A})$ es distinto de cero.

Teorema 2 Sea \mathbf{A} una matriz cuadrada de dimensión n y sea \mathbf{b} un vector columna de dimensión n . El sistema $\mathbf{Ax}=\mathbf{b}$ tiene un único vector solución \mathbf{x} si y solo si \mathbf{A} es no singular.

Corolario 1 *El sistema $\mathbf{Ax}=\mathbf{0}$ tiene una solución distinta de cero si y solo si el determinante de \mathbf{A} es igual a cero.*

C.8. Independencia lineal

Un conjunto de vectores es linealmente independiente si ninguno de ellos puede ser escrito con una combinación lineal de los restantes. Sea $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \dots, \mathbf{x}_n$ vectores en el espacio vectorial V , son linealmente independientes si

$$a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + a_3\mathbf{x}_3 + \dots + a_n\mathbf{x}_n = \mathbf{0} \quad (\text{C.12})$$

Se dice que son linealmente independientes si existen números $a_1, a_2, a_3, \dots, a_n$, no todos igual a cero.

Entre las propiedades de los vectores linealmente dependientes e independientes encontramos:

1. Un conjunto de vectores es linealmente dependiente si y solamente si alguno de los vectores es combinación lineal de los demás.
2. Si un conjunto de vectores es linealmente independiente cualquier subconjunto suyo también lo es.

Sean $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \dots, \mathbf{x}_n$ vectores en \mathfrak{R}^n se dice que tales vectores generan a \mathfrak{R}^n si y solo si para cada vector en \mathfrak{R}^n existen $a_1, a_2, a_3, \dots, a_n$ tales que:

$$a = a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + a_3\mathbf{x}_3 + \dots + a_n\mathbf{x}_n \quad (\text{C.13})$$

Una base de \mathfrak{R}^n es un conjunto en \mathfrak{R}^n que son linealmente independientes y que generan a \mathfrak{R}^n .

Los vectores $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \dots, \mathbf{x}_n$ son una base de \mathfrak{R}^n si y solo si el determinante de la matriz cuyas columnas son $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \dots, \mathbf{x}_n$ es distinto de cero.

C.9. Valores y Vectores propios (eigenvalores, eigenvectores)

Sea \mathbf{A} una matriz cuadrada de dimensión n . Un vector \mathbf{x} se dice que es un eigenvector o valor característico de la matriz \mathbf{A} si y solo si existe una constante λ tal que $\mathbf{Ax} = \lambda \mathbf{x}$. La constante λ se denomina el eigenvalor o valor característico de \mathbf{A} asociado a un vector \mathbf{x} .

Sea \mathbf{A} una matriz cuadrada de dimension n . Los eigenvalores de λ de \mathbf{A} se obtienen de resolver la ecuacion polinomial.

$$|\mathbf{A} - \lambda\mathbf{I}| = 0 \tag{C.14}$$

La cual se denomina la ecuación característica asociada a la matriz cuadrada \mathbf{A} . Los vectores propios se obtienen de resolver el sistema $\mathbf{Ax} = \lambda \mathbf{x}$, donde λ es una de las raíces de la ecuación característica.

Apéndice D

Conceptos de teoría de la información

La Teoría de la Información es una teoría matemática creada por Claude Shannon en el año 1948 y que forma la piedra angular sobre la que se ha desarrollado toda la teoría actual de la comunicación y la codificación. Esta teoría establece los límites de cuánto se puede comprimir la información y de cuál es la máxima velocidad a la que se puede transmitir información. La Teoría de la Información es, por tanto una teoría de límites alcanzables: máxima compresión de datos y máxima tasa de transmisión de información transmitida sin errores. Las aplicaciones de esta teoría son enormes y abarcan desde las ciencias de la computación (criptografía, aprendizaje), la ingeniería eléctrica (Teoría de la comunicación y teoría de la codificación), la estadística o la biología (secuencias de ADN, código genético).

Una gran parte de los métodos de ICA se basan en conceptos de la Teoría de la información que se presentan en este apéndice. Pueden encontrarse más detalles en las referencias [14, 14, 35].

D.1. Entropía

La información se encuentra muy relacionada a la aleatoriedad o «sorpresa» de un resultado. Por ejemplo considere el caso de lanzar una moneda al aire. La probabilidad que caiga cara o sol es del 50 %, se dice que este experimento es equiprobable.

Y conocer previamente el resultado, no resulta posible si no se cuenta con mas información.

Shannon propuso la entropía como una medida apropiada de la incertidumbre. La entropía es definida como:

$$H(X) = \sum_{x \in A_x} P(x) \log \frac{1}{P(x)} \quad (\text{D.1})$$

Donde X es el conjunto de la variable aleatoria x , donde éste puede tomar posibles valores. Cuando se usa el logaritmo natural o neperiano, la unidad de información es el *nat*, y cuando se usa el logaritmo en base 2 la unidad es el *bit*. La Figura D.1 muestra la gráfica de la entropía en donde se observa, que la máxima entropía se alcanza cuando la probabilidad es del 50% y de cero cuando se conoce certeramente su valor. Es decir, que cuando en la situación no hay «sorpresas» el evento no lleva «información» ya que se sabe a priori cual será el mensaje. Mientras que, al aumentar el nivel de «incertidumbre» entonces habría mas «sorpresa» y por lo tanto mas información. En consecuencia, las palabras «incertidumbre», «sorpresa» e «información» están relacionadas.

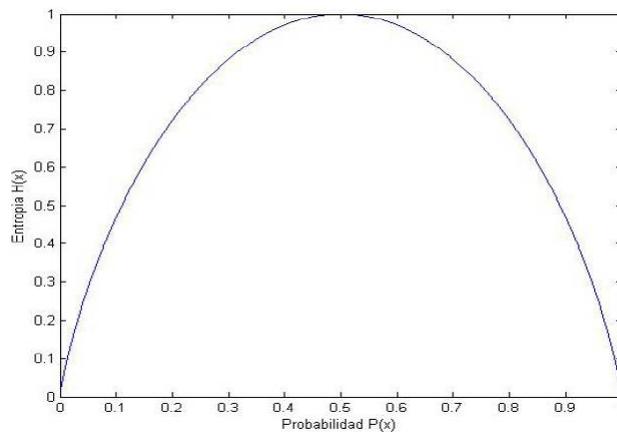


Figura D.1: Entropía para el lanzamiento de una moneda

Note que la Figura es una curva simétrica y alcanza su valor máximo cuando se obtiene la máxima incertidumbre.

D.2. Entropía conjunta, marginal y condicional

La *entropía conjunta* (o de juntura) es una medida de la entropía y mide la información contenida en un sistema de dos o mas variables. La entropía conjunta se define como:

$$H(X, Y) = \sum_{x, y \in A_x A_y} P(x, y) \log \frac{1}{P(x, y)} \quad (D.2)$$

Donde $P(x, y)$ es la función de probabilidad conjunta de las variables x y y . Si estas dos variables aleatorias fueran independientes, se cumpliría la siguiente condición $P(x, y) = P(x)P(y)$. Entonces se obtendría que la entropía resultaría en:

$$\begin{aligned} H(X, Y) &= \sum_{x, y \in A_x A_y} P(x)P(y) \log \frac{1}{P(x)P(y)} & (D.3) \\ &= \sum_{x, y \in A_x A_y} P(x)P(y) \log \frac{1}{P(x)} + \sum_{x, y \in A_x A_y} P(x)P(y) \log \frac{1}{P(y)} \\ &= \sum_{x \in A_x} P(x) \left(\sum_{y \in A_y} P(y) \right) \log \frac{1}{P(x)} + \sum_{y \in A_y} P(y) \left(\sum_{x \in A_x} P(x) \right) \log \frac{1}{P(y)} \\ &= \sum_{x \in A_x} P(x) \log \frac{1}{P(x)} + \sum_{y \in A_y} P(y) \log \frac{1}{P(y)} \\ H(X, Y) &= H(X) + H(Y) \end{aligned}$$

Esto demuestra que la entropía conjunta para dos variables independientes, es la suma de sus entropías marginales o individuales.

Para medir la incertidumbre de una variable X dada otra variable Y , se define la entropía condicional de X dado Y , escribiéndose como $H(X|Y)$. La entropía condicional $H(X|Y)$ representa la cantidad de incertidumbre que existe en X después de haber observado a Y , y se define como:

$$\begin{aligned} H(X|Y) &= H(X, Y) - H(Y) & (D.4) \\ \text{con } 0 &\leq H(X|Y) \leq H(X) \end{aligned}$$

La diferencia $H(X) - H(X|Y)$ representa la incertidumbre de la entrada que se va a obtener observando la salida. Esta cantidad se denomina información mutua entre las variables aleatorias X e Y . Denotándola como $I(X;Y)$.

D.3. Información mutua

La *información mutua* es la medida de información que tiene una variable con respecto a otra. Para dos variables aleatorias, se tiene que:

$$I(X;Y) = H(X) - H(X|Y) \quad (\text{D.5})$$

La información mutua $I(X;Y)$ es simétrica y no negativa. Donde $H(Y|X)$ es la entropía condicional. El primer miembro de la expresión D.5 es la media de la información dada por la salida del sistema X , menos la media de la información dada por Y suponiendo que conocemos la entrada X .

La Figura D.2 muestra una interpretación visual de la entropía de la entrada X , representada con el círculo de la izquierda, la entropía de la salida Y se representa con el círculo de la derecha y la información mutua entre X e Y se representa con la intersección de ambos círculos.

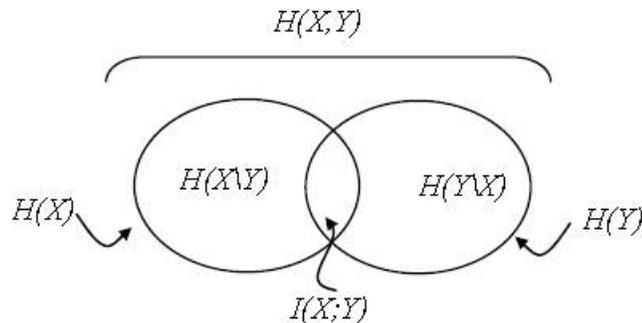


Figura D.2: Relación entre la información mutua $I(X;Y)$ y las entropías $H(X)$ y $H(Y)$.

D.4. Divergencia de Kullback-Leibler

La divergencia Kullback-Leibler o entropía relativa, determina la distancia entre dos densidades de probabilidad. Tomando valores positivos, e iguales que cero cuando las dos distribuciones son semejantes. Esta es una consecuencia directa de la estricta convexidad de el logaritmo negativo, y la aplicación de la clásica desigualdad de Jensen.

$$E \{f(y)\} \geq f(E\{y\}) \quad (\text{D.6})$$

Asignando a $f(y) = -\log(y)$ y asumiendo que $y = p^2(x)/p^1(x)$, se tendrá que:

$$E \{f(y)\} = E \left\{ -\log \frac{p^2(x)}{p^1(x)} \right\} = \int p^1(x) \left\{ -\log \frac{p^2(x)}{p^1(x)} \right\} dx \quad (\text{D.7})$$

Obteniendo la igualdad si las dos distribuciones son semejantes. La divergencia de Kullback-Leibler, no es propiamente una medida de la distancia.

La divergencia de Kullback-Leibler tiene las siguientes propiedades importantes: La información Mutua es positiva y cero si y solo si las variables son independientes. Esto es una consecuencia directa de la divergencia de Kullback-Leibler.

$$D(P\|Q) = \sum_x P(x) \log \frac{P(x)}{Q(x)} \quad (\text{D.8})$$

La información mutua es una forma especial de la entropía relativa.

$$I(X; Y) = D(P(x, y)\|P(x)P(y)) \quad (\text{D.9})$$

Bibliografía

- [1] James Rowan A. y Tolunsky Eugene. Conceptos básicos sobre eeg con mini-atlas. *Elsevier España*, 2004.
- [2] Lee Edward A. y Varaiya Pravin. *Structure and Interpretation of signals and systems*. Addison-Wesley, United States, 2000. ISBN 978-0-201-74551-1.
- [3] Vivaldi Ennio A. y Maldonado Pedro. Computers in biomedics research: I. analysis of bioelectrical signals. *Revista médica de Chile*, 2001.
- [4] Hyvärinen Aapo, Karhunen Juha, y Oja Erkki. *Independent Component Analysis*. John Wiley and Sons, New York, 2001. ISBN 0-471-22131-7.
- [5] Hyvärinen Aapo, Ramkumar Pavan, Parkkonen Lauri, y Hari Riitta. Independent component analysis of short-time fourier transforms for spontaneous eeg/meg analysis. *NeuroImage*, 2010.
- [6] Husain AM. Review of neonatal eeg. american journal of electroneurodiagnostic technology. *American Journal of Electroneurodiagnostic Technology*, 2005.
- [7] Ron Angevin y Ricardo. *Retroalimentación en el entrenamiento de una interfaz cerebro computadora usando técnicas basadas en realidad virtual*. Tesis Doctoral, Universidad de Málaga, España, 2005.
- [8] Antony J. Bell y Terrence J. Sejnowski. An information-maximization approach to blind separation and blind deconvolution.
- [9] B. Boashash, Mesbah M., y P Colditz. Newborn eeg seizure pattern characterisation using time-frequency analysis. ICASSP. IEEE Computer Society, Washington, DC, 2001.

-
- [10] Kemp Bob. Edf: a simple format for graphical analysis results from polygraphic siesta recordings. *J Sleep*, 1998.
- [11] Díaz Martínez Carina, Díaz Martínez Annia, y Céspedes Fernández Pedro L. Valor del electroencefalograma en neonatología. *Revista Mexicana Neurociencia*, 2006.
- [12] Guerrero Mosquera Carlos, Malanda Trigueros Armando, y Iriarte Franco Jorge. Detección de crisis epilépticas a partir de señales eeg mediante el método de extracción de crestas. XXIII Congreso Anual de la Sociedad Española de Ingeniería Biomédica, España, 2005.
- [13] G. R. Cooper. *Probabilistic Methods of Signal and System Analysis*. Rinehart and Winston Inc, 1971.
- [14] Thomas M. Cover. *Elements of Information Theory*. A Wiley Interscience Publication, 1991.
- [15] Binnie C. D. Eeg, paediatric neurophysiology, special techniques and applications. *Elsevier Health Sciences*, 2003.
- [16] Johnson Dallas E. *Métodos Multivariados Aplicados al Análisis de Datos*. Internacional Thompson, 2000.
- [17] Zhu Dan-hua, Tong Ji-jun, y Chen Yu-quan. An ica-based method for automatic eye blink artifact correction in multi-channel eeg. ICITAB, IEEE, Shenzhen, China, 2008.
- [18] Delamonica. *Electroencefalografía*. El Ateneo Pedro García S.A, segunda ed^{ón}., 1984. ISBN 950-02-0176-3.
- [19] s.a. de c.v. Departamento técnico Llog. ¿sabe usted para que sirve el ajuste de filtros en sus equipos de pnd? 2009.
- [20] Stephen Faul, G. Gregorcic, G. Boylan, W. Marnane, G. Lightbody, y S. Connolly. Gaussian process modeling of eeg for the detection of neonatal seizures. *IEEE Transactions on Biomedical Engineering*, 2007.

-
- [21] William Feller. *An introduction to probability theory and its applications*. Wiley International Edition, 1968.
- [22] Piryatinska A. and Terdik G., Woyczynski W. A., Loparo K. A., y A. Scher M. S. and Zlotnik. Automated detection of neonate eeg sleep stages. *Comput. Methods Prog. Biomed.*, 2009.
- [23] Charles F. Van Loan Gene H. Golub. *Matrix Computations*. The Johns Hopkins University Press, 1993.
- [24] Vaclav Gerla, Lenka Lhotska, Vladimir Krajca, y Karel Paul. Multichannel analysis of the newborn eeg data. *Congress Center Du Lac*, 2006.
- [25] Guerrero Graciela, Mayor Luis C., y Quijano Carlos A. Electroencefalograma neonatal. *Guía 7 Neurológica*, 2005.
- [26] B.R. Greene, W.P. Marnane, G. Lightbody, R.B. Reilly, y G.B. Boylan. Classifier models and architectures for eeg-based neonatal seizure detection. *IEEE Transactions on Signal Processing*, 2008.
- [27] Hagmann, Cornelia Franziska, Robertson, Nicola Jayne, Azzopardi, y Denis. Artifacts on electroencefalograms may influence the amplitude-integrated eeg classification: A qualitative analysis in neonatal. *Encephalopathy Pediatrics*, 2006.
- [28] Richar W. Hamming. *The Art of Probability for scientists and engineers*. Addison-Wesley Publishing Company, 1991.
- [29] Brotchie J., Rankine L., Mesbah M., Colditz P., y Boashash B. Robust time-frequency analysis of newborn eeg seizure corrupted by impulsive artefacts. 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Lyon, France, 2007.
- [30] Martínez J. y Tejeiro. Eeg normal.

-
- [31] Muñoz Juliana, Rivera Jorge, y Duque Edison. Análisis de componentes principales e independientes aplicados a reducción de ruido en señales electrocardiográficas. *Scientia et Technica*, 2008.
- [32] Amar Kachenoura, Laurent Albera, Lotfi Senhadji, y Pierre Comon. Ica: A potential tool for bci system, an in-depth look, brain computer interfaces. *IEEE signal processing magazine*, 2008.
- [33] Poh KoK y Marziliano P. Analysis of neonatal eeg signals using stockwell transform. Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society., 2007.
- [34] Poh KoK y Marziliano Pina. Compression of neonatal eeg seizure signals with finite rate of innovation. Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference, 2008.
- [35] Tee Won Lee. *Independent Component Analysis: Theory and applications*. Kluwer Academic Publisher, California, 2000. ISBN 0-7923-8261-7.
- [36] Steven J. Leon. *Linear Algebra With Applications*. Prentice Hall, 1990.
- [37] Wong Lisa y Abdulla W. Time-frequency evaluation of segmentation methods for neonatal eeg signals. Proceedings of the 28th IEEE, Annual International Conference, New York, 2006.
- [38] Oliva Moreno Luz Noé. *Sistemas de Procesamiento de Señales para el Análisis de Información Multidimensional*. Tesis Doctoral, Cinvestav, México, D.F., 2008.
- [39] Rangaraj M. y Rangayyan. *Biomedical Signal Analysis: A Case-Study Approach*. Wiley-IEEE Press, United Stated, first ed^{ón}., 2001. ISBN 0471208116.
- [40] Carlos Guerrero Mosquera y Angel Navia Vazquez. Automatic removal of ocular artifacts from eeg data using adaptive filtering and independent component analysis. 17 th European Signal Processing Conference (EUSIPCO), Glasgow, Scotland, 2009.

-
- [41] Mukherjee, A. Karayiannis, N.B. Glover, J.R. Hrachovy, R.A. Frost, J.D., Jr. Mizrahi, y E.M. Evaluation of cosine radial basis function neural networks in detection of artifacts in neonatal eeg. Proceedings of the 25th Annual International Conference of the IEEE, México, 2003.
- [42] Leonardo Palacios. Breve historia de la electroencefalografía. *Acta Neurológica Colombiana*, 2002.
- [43] A. Papoulis. *Probability, Random Variables, and Stochastic Processes*. McGraw Hill, 3rd edition ed^{ón}., 1991.
- [44] Zarjam Pega, Mesbah Mostefa, y Boashash B. Comparing two time-scale and time-frequency based methods in newborns eeg seizure detection. *IEEE, International Conference on Signal Processing and Communications*, 2007.
- [45] L. Rankine, N. Stevenson, M. Mesbah, y B. Boashash. A nonstationary model of newborn eeg. *IEEE Transactions on Biomedical Engineering*, 2007.
- [46] Romero S, Mañanas Ma, Riba J, Morte A, y Barbanoj Mj. Nuevo método automático de filtrado ocular en señales eeg espontáneas basado en análisis de componentes independientes. *XXV Jornadas de Automática*, 2004.
- [47] Romero Lafuente Sergio. *Reducción de artefactos en señales electroencefalográficas mediante nuevas técnicas de filtrado automático basadas en separación ciega de fuentes*. Tesis Doctoral, Universidad Politécnica de Cataluña, Cataluña, España, 2010.
- [48] Faul Stephen. *Automated Neonatal Seizure Detection*. Tesis Doctoral, National University, Ireland, Cork, 2007.
- [49] N. Stevenson, L. Rankine, M. Mesbah, y B. Boashash. Newborn eeg seizure simulation. *WDIC*, 2005.
- [50] Gilbert Strang. *Algebra lineal y sus aplicaciones*. Massachusetts Institute of Technology, Fondo Educativo Interamericano, 1982.
- [51] Steven T. y Karris. *Signals and System: with MATLAB Applications*. Orchard Publications, California, segunda ed^{ón}., 2003. ISBN 0-97095 11-8-3.

-
- [52] Teunis van Beelen. Edfbrowser. 2010. URL <http://www.teuniz.net/edfbrowser/>.
- [53] Deburchgraeve W, Cherian PJ, De Vos M, Swarte RM, Blok JH, y Visser GH. Automated eeg artefact reduction in eeg as a preprocessing step for neonatal seizure detection. *ESAT-SISTA*, 2008.
- [54] Deburchgraeve W, Cherian PJ, De Vos M, Swarte RM, Blok JH, y Visser GH. Automated neonatal seizure detection mimicking a human observer reading eeg. *Clinical Neurophysiology*, 2008.
- [55] J. Xi, J. F. Chicharo, A. C. Tsoi, y W. C. Siu. On the infomax algorithm for blind signal separation.
- [56] Stein Y. y Jonathan. *Digital Signal Processing: A Computer Science Perspective*. John Wiley and Sons, United States, 2000. ISBN 0-471-29546-9.