

Proyecto Fin de Máster Máster Universitario en Ingeniería de Telecomunicaciones

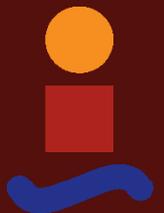
Estudio del método Common Spatial Patterns y sus variantes en interfaces cerebro-ordenador

Autor: F. Javier Olías Sánchez

Tutor: Sergio A. Cruces

**Dpto. Teoría de la Señal y Comunicaciones
Escuela Técnica Superior de Ingeniería
Universidad de Sevilla**

Sevilla, 2016



Proyecto Fin de Máster

Máster Universitario en Ingeniería de Telecomunicaciones

Estudio del método Common Spatial Patterns y sus variantes en interfaces cerebro-ordenador

Autor:

F. Javier Olías Sánchez

Tutor:

Sergio A. Cruces

Profesor Titular

Dpto. Teoría de la Señal y Comunicaciones

Escuela Técnica Superior de Ingeniería

Universidad de Sevilla

Sevilla, 2016

Proyecto Fin de Máster: Estudio del método Common Spatial Patterns y sus variantes en interfaces cerebro-ordenador

Autor: F. Javier Olías Sánchez

Tutor: Sergio A. Cruces

El tribunal nombrado para juzgar el trabajo arriba indicado, compuesto por los siguientes profesores:

Presidente:

Vocal/es:

Secretario:

acuerdan otorgarle la calificación de:

El Secretario del Tribunal

Fecha:

Agradecimientos

El desarrollo de este trabajo pone final a mis estudios de máster y a una etapa de mi vida. Es por eso que en esta página quiero agradecer a todas las personas que me han empujado a conseguir mis objetivos y a crearlos, porque debo admitir que no los he formado yo solo.

En primer lugar y aunque nunca lo lleguen a saber, quiero darles las gracias a mis músicos favoritos (*Kase-O, Morodo, Soldiers of Jah Army, Stick Figure, The Green, etc...*) por hacer las canciones que me han acompañado durante todos mis años de estudiante y en especial durante la realización de este trabajo porque sin su música todo hubiese sido más difícil.

Tampoco me puedo olvidar de mi primo Manolo, ni de mis amigos Ángel y Mayte porque siempre han estado a mi lado y con los que tantos buenos momentos he pasado. Ellos nunca han dudado en echarme una mano y me han enseñado muchas más cosas de las que he aprendido estudiando.

Quien no puede faltar en esta página es mi familia, porque sin su ayuda hoy yo no estaría aquí. Agradezco a mis padres la educación que me han dado. Ellos me han enseñado a no rendirme nunca, a ser tan cabezón como soy; me han educado para que sea libre y tome mis propias decisiones y me siento realmente afortunado por ello, ya que es algo que no todo el mundo ha podido disfrutar. Le doy las gracias a mis hermanas Isabel y Macarena por darme siempre su amor, por haber estado siempre de mi parte, porque junto a ellas formé mi primer equipo (aunque solo fuese para afrontar problemas de niños) y porque se que ellas nunca me van fallar.

Finalmente quiero agradecerle a Carmen todo su amor y cariño. Ella me ha enseñado a estudiar, a organizarme, a dedicarle el tiempo que cada cosa necesita. Quiero agradecerle todas las tardes, mañanas y noches que hemos pasado en las bibliotecas o en el cuarto de estudio. Quiero agradecerle todos los besos, todas las veces que me ha animado a seguir cuando ya no podía más, todas las fuerzas que me ha dado, porque ella ha estado conmigo durante la mayor parte de mis años de estudiante, con ella he aprobado todas las asignaturas de este máster y ha hecho que esta carrera sea mucho más fácil.

Resumen

Las comunicaciones Brain Computer Interface (BCI) consisten en una tecnología que permite que las personas puedan comunicarse con una máquina o un ordenador, usando para ello el cerebro y en la mayoría de los casos un casco de EEG.

Este campo supone un gran reto para la ingeniería (junto con muchas otras ramas de conocimiento) y en la actualidad se está investigando mucho y se están realizando grandes avances. La importancia de investigar y avanzar en la realización de estos sistemas se debe a que los sistemas BCI pueden ser de gran ayuda a personas que sufren de algunos trastornos de parálisis cerebral, o que padecen de otras enfermedades o discapacidades que impidan el uso normal de sus habilidades motoras. Se cree que estos sistemas pueden mejorar considerablemente la calidad de vida de estas personas, para las cuales pequeños avances y cambios implican grandes mejoras.

Common Spatial Patterns (CSP) es un algoritmo muy conocido y ampliamente usado que ha cobrado gran importancia durante los últimos años por sus aplicaciones en BCI para los sistemas basados en EEG multicanales. El algoritmo CSP consiste en encontrar un filtro espacial óptimo, que reduzca la dimensionalidad de las señales originales pudiendo tomar tantos canales como se desee.

El objetivo de este trabajo consiste en realizar un repaso sobre esta técnica y también sobre Linear Discriminant Analysis (LDA), que se trata de una técnica de clasificación lineal. Además, se ha implementado un algoritmo basado en CSP con el que se consiguen mejorar los resultados que se obtienen usando únicamente la técnica de CSP. El algoritmo desarrollado es capaz de distinguir entre dos clases y además, se ha realizado una extensión en la que se distingue entre cuatro clases usando un sistema de votaciones simple.

Para poder probar y comprobar el correcto funcionamiento de ambos algoritmos desarrollados, se han usado los datos procedentes de una competición pública de BCI, que ha sido usada como referencia en numerosos artículos. Esto nos ha permitido comparar los resultados obtenidos con nuestros algoritmos con aquellos obtenidos mediante otras técnicas y variantes de CSP, como sería el caso del algoritmo RSTFC, que también ha sido implementado y probado durante este trabajo.

Por último, se han obtenido unas conclusiones de los sistemas BCI, así como de las distintas técnicas mencionadas anteriormente. Para ello nos hemos ayudado de gráficas y medidas obtenidas a partir de los resultados obtenidos. También hemos podido extraer conclusiones a partir de ilustraciones de los filtros espaciales calculados con CSP

abstract

The BCI communication is a technology which allows to people to communicate with a machine or a computer using their own brains and normally a EEG helmet.

This field of knowledge is a great challenge for engineering (and for other fields). Now days BCI communications are being very studied and developed, having great advances. The importance of developing better BCI system falls in the believing that it can help a lot of people who are suffering of some kind of paralysis or some kinds of motor inability. We believe that this technology can be very useful for those people and that their lifestyle could be increased significantly, to whom little improvement means a lot.

CSP is very a well-known algorithm and it has been widely used in many BCI system especially in those systems which use EEG helmets. CSP is used to reduce the dimensionality of the EEG signals allowing to choose how many channels the user needs.

The main of this work is to make a review in CSP algorithm and about LDA too, which is a classification technique. Furthermore, an algorithm based on these two techniques has been developed which improve the results obtained using the simple CSP. The algorithm developed is able to classify between two classes but an extension have been done making able to distingue between four imaginary movements. These extension is based in a simple vote system. Whit the purpose of comparing and checking the well behavior theses algorithms they have been test over a public set of data. This set of data is given from IV BCI competition. These data have been wildly used in multiple articles which has allow us to compare the algorithm whit others, like the RSTFC algorithm which have been also tried in this work.

Finally, some conclusions about BCI system and the algorithm have been done. To accomplish that we used some plots about the results and about the filters that CSP provides.

Índice Abreviado

Resumen	III
abstract	V
<i>Índice Abreviado</i>	VII
1 Introducción	1
1.1 Brain Computer Interface	2
1.2 Método de adquisición	4
1.3 El cerebro humano	5
1.4 Sistemas BCI	8
1.5 Conclusiones	11
2 Principales técnicas usadas	13
2.1 CSP	13
2.2 LDA	19
2.3 Conclusiones	22
3 Conjunto de datos	25
3.1 Datos experimentales	25
3.2 Grabación de los datos	26
3.3 Descripción de los datos	27

3.4	Conclusiones	29
4	Algoritmo Implementado	31
4.1	Visión general	31
4.2	Elección de parámetros	32
4.3	Partición de señales	33
4.4	Iniciación de CV	34
4.5	Banco de filtros	34
4.6	Cálculo de filtros espaciales	35
4.7	Extracción de características	36
4.8	Entrenamiento del clasificador	37
4.9	Pruebas de CV	37
4.10	Test	37
4.11	Extensión a cuatro clases	38
4.12	Conclusiones	40
5	RSTFC: A Novel Algorithm for Spatio-Temporal Filtering and Clasification of Single-Trial EEG	41
5.1	Metodología	41
5.2	Conclusiones	44
6	Resultados	45
6.1	Mano izquierda - Mano derecha	47
6.2	Mano izquierda - Ambos pies	52
6.3	Mano izquierda - Lengua	53
6.4	Mano derecha - Ambos pies	53
6.5	Mano derecha - Lengua	55
6.6	Ambos pies - Lengua	57
6.7	Multiclase	59

6.8	Regularized Spatio-Temporal Filtering and Classification	60
6.9	Conclusiones	63
7	Análisis de los resultados	65
7.1	Descomposición en bandas de frecuencia	65
7.2	Zonas de cerebro	66
7.3	Porcentajes de acierto	68
7.4	Multiclase	69
7.5	Conclusiones	72
8	Conclusiones	75
	<i>Índice de Figuras</i>	79
	<i>Índice de Tablas</i>	83
	<i>Bibliografía</i>	85

Índice

Resumen	III
abstract	V
<i>Índice Abreviado</i>	VII
1 Introducción	1
1.1 Brain Computer Interface	2
1.2 Método de adquisición	4
Electroencefalograma	4
1.3 El cerebro humano	5
1.3.1 Las ondas cerebrales	7
1.4 Sistemas BCI	8
1.4.1 Ondas μ	8
1.4.2 La onda P300	9
1.5 Conclusiones	11
2 Principales técnicas usadas	13
2.1 CSP	13
2.1.1 Extracción de características	17
2.2 LDA	19
2.2.1 Teorema de Bayes	19
2.2.2 LDA para dos clases	20
2.3 Conclusiones	22
3 Conjunto de datos	25
3.1 Datos experimentales	25
3.2 Grabación de los datos	26
3.3 Descripción de los datos	27
3.4 Conclusiones	29
4 Algoritmo Implementado	31
4.1 Visión general	31
4.2 Elección de parámetros	32
4.3 Partición de señales	33
4.4 Iniciación de CV	34

4.5	Banco de filtros	34
4.6	Cálculo de filtros espaciales	35
4.7	Extracción de características	36
4.8	Entrenamiento del clasificador	37
4.9	Pruebas de CV	37
4.10	Test	37
4.11	Extensión a cuatro clases	38
4.12	Conclusiones	40
5	RSTFC: A Novel Algorithm for Spatio-Temporal Filtering and Clasification of Single-Trial EEG	41
5.1	Metodología	41
5.2	Conclusiones	44
6	Resultados	45
6.1	Mano izquierda - Mano derecha	47
6.1.1	Usando una sola banda (Alfa y Beta a la vez)	47
6.1.2	Usando dos bandas (Alfa y Beta por separado)	48
6.1.3	Usando cuatro bandas	49
6.2	Mano izquierda - Ambos pies	52
6.3	Mano izquierda - Lengua	53
6.4	Mano derecha - Ambos pies	53
6.5	Mano derecha - Lengua	55
6.6	Ambos pies - Lengua	57
6.7	Multiclase	59
6.8	Regularized Spatio-Temporal Filtering and Classification	60
6.9	Conclusiones	63
7	Análisis de los resultados	65
7.1	Descomposición en bandas de frecuencia	65
7.2	Zonas de cerebro	66
7.3	Porcentajes de acierto	68
7.4	Multiclase	69
7.4.1	Regiones de decisión	69
7.4.2	Aciertos por clase	71
7.5	Conclusiones	72
8	Conclusiones	75
	<i>Índice de Figuras</i>	79
	<i>Índice de Tablas</i>	83
	<i>Bibliografía</i>	85

Notación

Tabla 1 Tabla de signos matemáticos usados en el trabajo.

Notación matemática usada	
T	Número de muestras temporales de un ensayo.
N	Número de filtros espaciales.
C	Número de canales (<i>número de sensores de la prueba EEG</i>).
\mathbf{X}	Señales de cada uno de los sensores (<i>filtradas con un filtro paso de banda</i>).
\mathbf{w}	Filtro espacial simple.
\mathbf{W}	Matriz de filtros espaciales.
\mathbf{Y}	Señales filtradas espacialmente.
\mathbf{y}	Señal filtrada espacialmente.
Σ	Matrices de estimación de covarianza de las señales \mathbf{X} .
σ^2	Varianzas de las señales filtradas.
\mathbf{u}	Vector que satisface que: $\mathbf{u}\mathbf{u}^\top = 1$
S_i	Variable auxiliar: $S_i = U^\top \Sigma_T^{-1/2} \Sigma_L \Sigma^{-1/2} U$
λ	Símbolo usado para representar autovalores.
ϕ	Símbolo usado para representar autovectores.
\mathbf{z}	vector de características obtenidas de las señales \mathbf{X}
C	Número de clases en una clasificación.
c	Variable que se refiere a las clases.
\mathbf{z}	Variable que se refiere al conjunto de las obsevaciones.
o	Variable que se refiere a una observación en concreto.
\mathbf{z}_n	Puntos del plano.
\mathbf{a}	Coefficientes que definen un plano.
\mathbf{b}	Término independiente en la definición de un plano.
ρ	Coefficiente de regularización
$X^{(\tau)}$	Matriz de señales, retrasada τ muestras temporales.
\tilde{X}	Matriz de señales aumentada.

1 Introducción

*Imagination is more important than knowledge.
Knowledge is limited. Imagination encircles the
world.*

ALBERT EINSTEIN

Los sistemas BCI proporcionan un canal de comunicaciones novedoso para que las personas puedan interactuar con su entorno. La principal idea de estos sistemas es decodificar el estado mental de un sujeto a partir de su actividad cerebral y usar esta información para controlar una aplicación en concreto o un dispositivo electrónico.

El objetivo de este trabajo es realizar un algoritmo de procesamiento de señales para poder decodificar en que parte del cuerpo está pensando una persona a través de un sistema BCI.

Como ya se ha dicho, los sistemas BCI ofrecen una forma de comunicación entre el cerebro y el entorno de una persona sin que intervengan otros órganos o músculos, de manera que podemos encontrar en esto una gran utilidad para aquellas personas que tengan algún tipo de discapacidad asociada al sistema motor. Esta tecnología les permitiría manejar ordenadores o máquinas que les proporcionen una mayor autonomía, aumentando así considerablemente su nivel de vida.

Además de esta motivación, se espera que el progreso de los sistemas BCI nos ayude a avanzar en el estudio del cerebro, del que aun a día de hoy se sabe muy poco a pesar de ser considerado uno de los órganos más importantes del cuerpo humano. Al trabajar con estos sistemas, cada vez vamos aprendiendo más acerca de como el cerebro codifica y decodifica la información, los estados mentales y los impulsos que genera, entre otras muchas cosas.

Por otro lado los sistemas BCI suponen un reto científico, abarcando diferentes ramas del conocimiento en ingeniería y medicina.

Para la realización y comprobación de los resultados obtenidos con el algoritmo desarrollado, se han usado los datos proporcionados por una competición BCI, de manera que podemos comparar nuestro algoritmo con el resto de algoritmos y técnicas desarrolladas, abstrayéndonos así de los problemas surgidos en la etapa de adquisición de los datos, como podrían ser el uso de diferentes sistemas de adquisición de datos, aparición de artefactos, ruido o entrenamiento de los pacientes.

1.1 Brain Computer Interface

La actividad cerebral produce señales que son detectables sobre el cuero cabelludo, en la superficie del cerebro o dentro del mismo. La Interfaz Cerebro Ordenador BCI, traduce estas señales en salidas que permiten al usuario comunicarse sin la participación de músculos debido a que las señales cerebrales no dependen del control neuromuscular [30]. Al hablar de una comunicación BCI nos referimos a una comunicación en un único sentido: desde el cerebro al ordenador. La comunicación desde el computador a la persona se realiza normalmente mediante estímulos visuales o auditivos, por lo que no supone un problema y no es un campo a abordar en el estudio de BCI.

Para conseguir una comunicación fiable se debe conseguir una tasa de acierto igual o superior al 70%. Este límite se ha establecido porque según el programa de soporte lingüístico del departamento de educación de Australia¹ la comunicación verbal es posible a ese nivel [26], y además este límite ha usado en numerosas publicaciones sobre BCI como [21] [5].

En un principio, cuando se comenzaron a desarrollar los sistemas BCI, conseguir este porcentaje de acierto era una tarea muy dura, ya que era necesario un entrenamiento muy intensivo del usuario para conseguir la capacidad de interactuar correctamente con un sistema BCI [44] [45]. Sin embargo en la última década, gracias a las aportaciones de las técnicas de aprendizaje máquina, se han conseguido numerosos avances que han facilitado y acortado el tiempo de entrenamiento de los usuarios de BCI, como son las técnicas de Support Vector Machine (SVM) y CSP.

¹ Education and Training Department: <http://www.education.vic.gov.au/Pages/default.aspx>

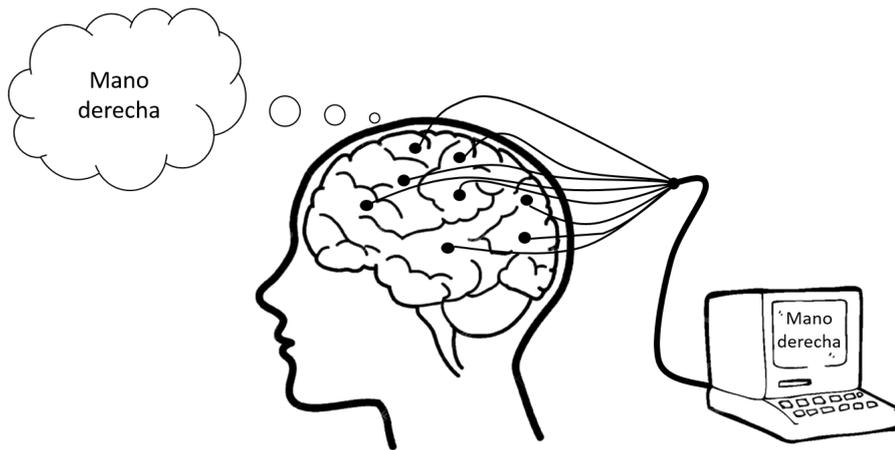


Figura 1.1 Esquema BCI. *Fuente:* Imagen modificada [2] .

Como cualquier sistema de comunicaciones, un sistema BCI se compone de:

- Una entrada, que en este caso sería la intención ² del usuario.
- Una salida que serían los comandos que interpreta la máquina.
- Un canal de comunicaciones que posibilite la transmisión de señales desde el cerebro al ordenador. Para este propósito se usarán los cascos de Electroencefalograma (del inglés Electroencephalography) (EEG).
- Un codificador de fuente. En BCI el propio cerebro humano es el encargado de realizar la tarea de codificador de fuente por lo que es necesario disponer de un lenguaje que el cerebro sea capaz de codificar directamente en señales que puedan ser transmitidas a través del canal. Es decir, disponer de un conjunto de símbolos que los cascos de EEG sean capaces de captar.
- Un decodificador de fuente, para que el ordenador sea capaz de interpretar esa información.

Teniendo en cuenta los distintos elementos que componen un sistema BCI, podemos ver que este tipo de comunicaciones abarcan un gran número de ramas. Entre las que más importancia tienen podemos encontrar la ingeniería electrónica que tiene gran importancia en el desarrollo del casco que capta las señales cerebrales. Por otro lado, tiene especial importancia el procesado de señales, que será la ciencia que usemos para interpretar las señales cerebrales y en la que se centra este trabajo y por último, aunque no menos importante, la neurología, que es la ciencia que estudia el cerebro desde un punto de vista fisiológico y biológico y nos permite entender el funcionamiento del cerebro.

² Desplazar un selector en una dirección, detener un contador...etc

1.2 Método de adquisición

Existen varias alternativas para lograr la adquisición de las señales propias de la actividad cerebral y con ella los símbolos en los que el usuario va a codificar la información. Entre ellas podemos destacar todas las técnicas basadas en campos magnéticos: Magnetoencefalografía (MEG), funcional Magnetic Resonance Imaging (fMRI) y funcional Near-Infrared Spectroscopy (fNRI). Sin embargo, para la adquisición de señales mediante estas técnicas son necesarios grandes dispositivos que requieren de mucha energía por lo que son pruebas que se realizan en un periodo de tiempo acotado, debido al consumo de energía y a la comodidad a la hora de trabajar con ellos (tamaño, precio,...), pero a pesar de ello existen varios estudios que han explorado la posibilidad de estas técnicas de adquisición para su uso en BCI [13][14][27][46]. El resultado de la prueba suele ser representado con una imagen o con un volumen, pero no proporcionan una medida por cada instante de tiempo durante el transcurso de la prueba, es decir, tienen una baja resolución temporal. Este es el motivo por el que en BCI se suelen usar únicamente dos técnicas de adquisición de señales: Electroencefalografía (del inglés Electroencephalography) (EEG) y ECoG. Ambas técnicas permiten ser usadas en tiempo real y permiten la movilidad de la persona que se somete a la prueba. No hay que olvidar que el objetivo de BCI es proporcionar a las personas una comunicación lo más cómoda posible.

La electroencefalografía (EEG) consiste en realizar un implante de una formación de electrodos en la corteza cerebral para poder medir la actividad de las distintas zonas del cerebro. De esta manera se consigue tener una relación señal a ruido mucho mayor que al usar EEG, consiguiendo además un ancho de banda mucho mayor (de 0 a 500HZ). Se ha demostrado que en estas bandas superiores a la banda gamma existe información importante relacionada con las funciones motoras y del habla [37] [4]. El principal problema de usar esta técnica es que, como ya se ha comentado, es necesario realizar un implante, lo que implica una operación quirúrgica en la cabeza del paciente para colocar una formación de sensores. Lógicamente los riesgos de esta operación son muy elevados, por lo que la gran mayoría de los estudios se realizan usando EEG. A continuación estudiamos este método de adquisición.

Electroencefalograma

El electroencefalograma se trata de un método de adquisición de señales para monitorizar la actividad cerebral. Se trata de un método no invasivo en el que se colocan unos electrodos sobre la corteza craneal que miden la variación de voltaje resultado de las corrientes iónicas que se dan en las neuronas cerebrales (12).

En un principio era un método muy usado para detectar enfermedades del cerebro como tumores, epilepsia o alzheimer; pero otras técnicas de adquisición de imágenes más adecuadas para la detección de estas patologías le ganaron el terreno. Hoy en día se sigue usando sobre todo para detectar epilepsia y para BCI. Además, el EEG se usa para evaluar problemas relacionados con el sueño, ya que existen algoritmos para detectar automáticamente si un paciente está dormido o despierto o las distintas fases de sueño [16]. En la Figura 1.2 se muestra un EEG de un paciente que ha empezado a sufrir una crisis

tónico-clónica que es el caso más común de epilepsia.

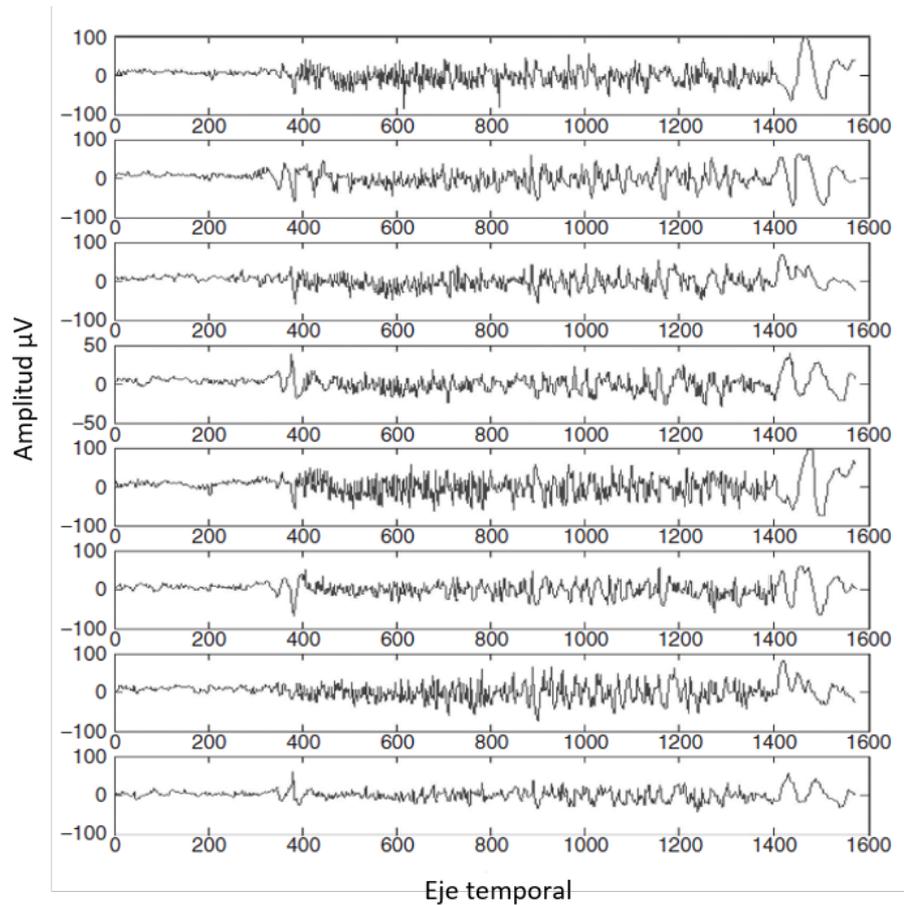


Figura 1.2 EEG de una crisis tónico-clónica. *Fuente:* Imagen extraída de [40] .

1.3 El cerebro humano

El cerebro humano es el órgano más complejo que poseemos y un gran desconocido en la actualidad, a pesar de que cada vez se dispone de más información del funcionamiento del mismo. Es imposible abordar un trabajo de BCI sin tener unos conocimientos básicos del cerebro, de manera que en este apartado se va a hacer un repaso de las principales características que nos pueden ser de utilidad en el estudio de BCI.

En primer lugar es necesario describir el cerebro físicamente: dispone de dos hemisferios y cuatro lóbulos. Los hemisferios son casi simétricos y distinguimos dos (dependiendo del lado del cuerpo humano en el que se encuentren): el hemisferio izquierdo y el hemisferio derecho. El hemisferio derecho controla la parte izquierda del cuerpo y el hemisferio izquierdo la parte derecha, es decir, cada uno controla la parte opuesta del cuerpo. Además el hemisferio izquierdo se encarga de la lógica, de las matemáticas y del lenguaje articulado, mientras que el hemisferio derecho se encarga de las habilidades artísticas y musicales.

En cuanto a los lóbulos, son las zonas del cerebro que se encargan de distintas funciones del cuerpo y podemos distinguir:

- **Lóbulo frontal:** es el más grande de todos los lóbulos. Entre sus principales funciones se encuentran el control de la conducta, la planificación y la regulación de las emociones [8].
- **Lóbulo parietal:** esta es la zona del cerebro que se encarga de controlar las extremidades del cuerpo, como las piernas y los brazos, así como otras actividades motoras. Como ya se ha adelantado, el lóbulo parietal del hemisferio izquierdo controla las extremidades derechas y viceversa. Recalcamos este hecho ya que nos aprovecharemos de él en nuestro sistema BCI como veremos. Además, dentro de este lóbulo, se pueden distinguir dos regiones: la región que esta más próxima al lóbulo frontal es la encargada de pensar los movimientos y se llama corteza motora. Por otro lado tenemos la corteza somatosensorial, que es la parte más próxima al lóbulo occipital, esta parte interpreta el sentido del tacto, siendo capaz de distinguir si un objeto es suave o áspero, si esta húmedo o seco, etc... [31].
- **Lóbulo occipital:** es el encargado de la visión, es decir, interpreta los estímulos visuales y los transforma en información útil. Lesiones en esta zona del cerebro pueden causar alucinaciones y otros trastornos relacionados con el sentido de la visión. [23].
- **Lóbulo temporal:** Sus funciones tienen que ver con la memoria y el reconocimiento de patrones. Este lóbulo recibe información del resto de lóbulos, además de recibir la información proveniente de los oídos, por lo que tiene gran relación con el sentido del equilibrio.

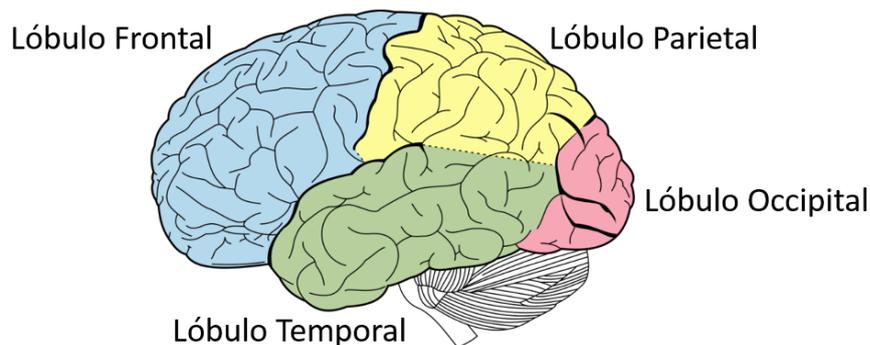


Figura 1.3 Lóbulos del cerebro. *Fuente:* Imagen extraída de [19] .

Entrando más en detalle, el cerebro está compuesto principalmente por neuronas que son células que trabajan conjuntamente. Dentro de las neuronas encontramos el cuerpo de la célula, las dendritas y los axones. El cuerpo celular tiene un único núcleo que contiene la mayoría de los metabolismos de las neuronas, especialmente los relacionados con la síntesis de las proteínas. Los axones son largos cilindros (que en los humanos pueden tener una longitud de entre una porción de milímetro hasta algo más de un metro) por los que se

transmiten señales eléctricas. Las dendritas se conectan a los extremos de los axones o a otras dendritas y entregan y reciben los impulsos nerviosos de otras neuronas.

Los nervios son conjuntos de axones en forma de cordones que comunican al cerebro con todos los órganos del cuerpo. La información transmitida por los nervios se llama potencial de acción [40]. Los potenciales de acción son un cambio temporal del potencial en la membrana que es transmitido por los axones y son generados por un intercambio de iones entre las membranas de las células. Los potenciales de acción pueden ser provocados por muchos tipos de estímulos, como por ejemplo estímulos químicos, visuales, eléctricos, de presión, de tacto o de fuerza.

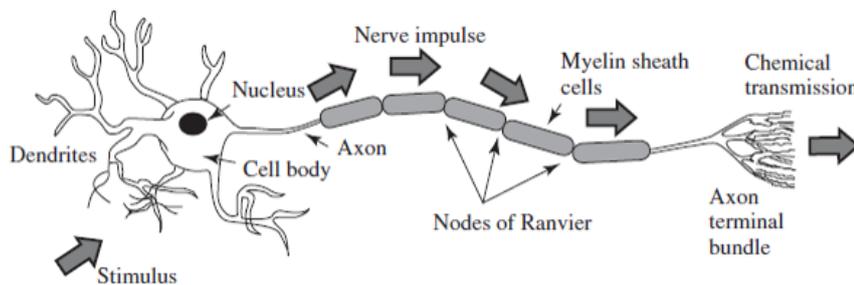


Figura 1.4 Esquema de una neurona. *Fuente:* Imagen extraída de [40].

1.3.1 Las ondas cerebrales

A medida que se fueron realizando estudios de EEG, se descubrieron distintos ritmos cerebrales que se asignaron a distintos estados mentales de los pacientes o usuarios que se sometían a la prueba. El estudio de estos ritmos ha supuesto un gran avance en las comunicaciones BCI. Vamos a ver que existen principalmente cinco tipos de ondas distinguidas por su rango de frecuencias: las ondas Delta, Theta, Alfa, Beta y Gamma.

Las ondas Delta (δ) son aquellas que tienen una frecuencia en el rango de 0.5-4Hz. Estas ondas se asocian principalmente a un sueño profundo y al ser tan lentas pueden ser confundidas fácilmente por artefactos debidos a movimiento, aunque analizando las señales es fácil ver si se trata de artefactos o de ondas Delta.

Las ondas Theta (θ) están en el rango de frecuencias de 4-7.5Hz. Estas ondas se asocian con la meditación profunda y con la inspiración creativa y aparecen cuando los pacientes están concentrados en realizar distintos tipos de tareas, por este motivo también se asocia a procesos en los que interviene la memoria [28].

Las ondas Alfa (α) aparecen en la zona posterior de la cabeza, normalmente sobre el lóbulo occipital con una frecuencia que oscila en el rango de 8-13 Hz. Suelen aparecer cuando los pacientes están relajados y pueden desaparecer si los pacientes están concentrados, tienen ansiedad o al escuchar sonidos no familiares. A día de hoy el origen fisiológico de las ondas Alfa es desconocido, pero se considera que son las ondas más conocidas e

investigadas de los ritmos cerebrales. Las ondas Alfa también son conocidas como ondas Berger, en honor al fundador del EEG que fue quien las descubrió.

Las ondas Beta (β) son aquellas ondas cuyas frecuencias están en el rango 14-26Hz. Este es el ritmo asociado a un pensamiento normal de una persona despierta que está prestando atención al mundo exterior o resolviendo algunos problemas. Están muy relacionada con las ondas μ y pueden ser bloqueadas por actividades motoras o por estímulos sobre el sentido del tacto.

Por último, a las ondas con frecuencias que están por encima de 30 Hz se las llama ondas Gamma (γ). Estas ondas se usan para detectar algunos tipos de enfermedades en las pruebas de EEG [43] [12] [9] [34]

1.4 Sistemas BCI

Como se ha adelantado, uno de los requisitos para la comunicación BCI es disponer de un lenguaje cuyos símbolos puedan ser codificados directamente por el cerebro humano y que el ordenador sea capaz de interpretar. Esta es una de las razones por la cual el lenguaje escogido no puede ser arbitrario, ya que si por ejemplo definimos como lenguaje la amplitud de un pulso cuadrado, el cerebro será incapaz de generar esta señal, porque físicamente le es imposible, y aunque tomemos como lenguaje símbolos que el cerebro si es capaz de codificar, deben ser símbolos que el cerebro pueda codificar voluntariamente y por sí mismo. Por ejemplo se ha demostrado que se pueden inducir ondas cerebrales a 10 Hz con distinta amplitud cuando se expone a la persona a ciertos estímulos visuales [32], pero estas ondas no pueden ser usadas como lenguaje porque son inducidas externamente y el cerebro no es capaz de generarlas por sí mismo. Con esto queremos decir que aunque el cerebro sea físicamente capaz de crear ciertos tipos de “símbolos” además debe de ser capaz de hacerlo bajo la voluntad de la persona y de manera intencionada.

Por estas razones, vemos que no tenemos demasiada libertad a la hora de elegir un lenguaje, además de ser difícil de optimizar o decidir cuál es el mejor.

Actualmente existen dos variantes de sistemas BCI que se basan en diferentes aspectos fisiológicos del cerebro: aquellos basados en ondas μ y aquellos basados en la onda P300. A continuación haremos un breve repaso de ellos.

1.4.1 Ondas μ

Las ondas μ o sensoriomotoras son patrones síncronos de actividad eléctrica que involucran grandes cantidades de neuronas del lóbulo parietal. Estas ondas suelen aparecer a frecuencias que están en el rango 7.5-12.5 Hz y aparecen con mayor amplitud cuando el cuerpo está en reposo. Cuando una persona realiza un movimiento de alguna parte del cuerpo, las ondas μ desaparecen en determinadas partes del lóbulo parietal, concretamente en la corteza motora, y dependiendo de la parte del cuerpo que se esté moviendo desaparece en una parte concreta de la corteza motora. A este fenómeno se le llama evento relacionado

con la desincronización (ERD), mientras que cuando la onda μ aparece normalmente se conoce como evento relacionado con la sincronización (ERS), esto ocurre cuando el sujeto se relaja después de la realización del movimiento imaginario. Con práctica los usuarios pueden llegar a imaginar estos movimientos sin nunca llegar a realizarlos [3].

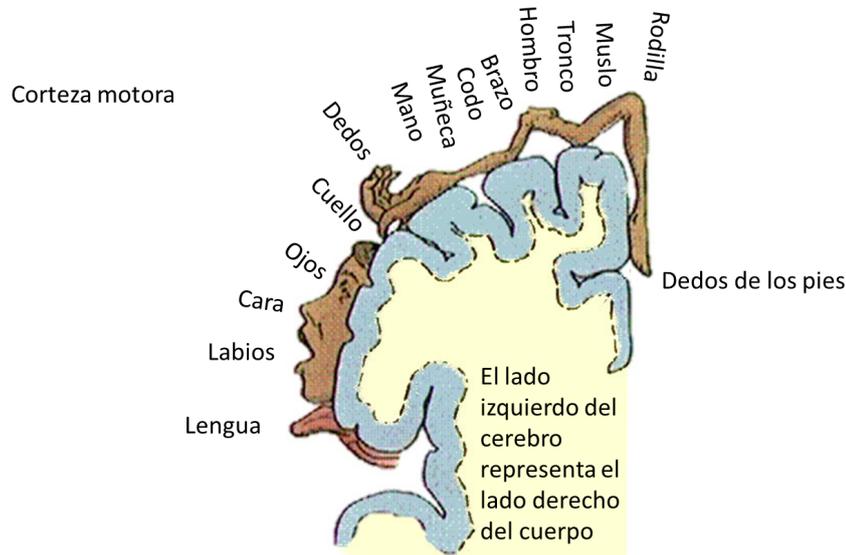


Figura 1.5 Correspondencia entre las zonas del lóbulo parietal y las zonas del cuerpo que controlan. *Fuente:* Imagen modificada [1].

Los ERD pueden ser captados por los cascos de EEG y esto lleva a un sistema BCI en el que los usuarios imaginan movimientos con distintas partes del cuerpo siendo el computador capaz de distinguir que parte del cuerpo el usuario tenía intención de mover, posibilitando la creación de un lenguaje.

Actualmente también se está trabajando con en el Potencial de Predisposición (Readiness Potential: RP), el cual se encuentra sincronizado y es previo (600 a 800 ms) a la realización o imaginación de un movimiento [15].

El algoritmo desarrollado en este trabajo se apoya en el paradigma de los ERD porque se puede encontrar más bibliografía y además nos permite diferenciar tipos de movimiento mientras que el potencial de predisposición está mas relacionado con los sistemas BCI que buscan la aparición o no de un evento y no la clasificación de tipos de movimiento como tal.

1.4.2 La onda P300

La onda P300 es un potencial evocado, es decir, es una reacción a un estímulo que no debe de ser periódico en el tiempo. La onda P300 es un potencial de acción que indica a la persona que un evento de interés ha ocurrido. Por ejemplo, la onda P300 aparece en el cerebro de los deportistas que van a realizar cualquier tipo de carrera cuando el árbitro da

la señal de salida.

El experimento típico para el estudio de la P300 es poner a personas frente a una pantalla en la que irán apareciendo patrones (distintas formas por ejemplo) e indicarles que pulsen un botón cuando aparezca un patrón en concreto. Numerosos estudios han demostrado que alrededor de unos 300ms después de la aparición del estímulo aparece un potencial positivo que puede ser capturado por los sistemas de EEG, especialmente en la zona parietal del cerebro[24].

El potencial aparece normalmente si se dan estas cuatro condiciones. En primer lugar se le debe presentar una secuencia de eventos aleatorios. En segundo lugar se deben de separar los eventos en dos categorías. La tercera condición es que se debe exigir que el usuario realice alguna acción, o tarea a partir de la regla anterior. La última condición es que el evento asociado a la tarea no sea muy frecuente [40].

De manera que podemos ver que la onda P300 no es exactamente una acción voluntaria del cerebro pero a pesar de ello está muy relacionada con la intención de la persona, por lo que puede ser útil para la comunicación BCI ya que indica la aparición de un evento que la persona está esperando. Una de las principales aplicaciones de este potencial en BCI es el “P300-based spelling paradigm” desarrollado por Dounchin [17].

Este sistema de deletreo consiste en presentar una matriz de caracteres, las columnas y filas se van iluminando aleatoriamente a una alta frecuencia, por ejemplo 8 veces por segundo. El usuario selecciona un carácter concentrándose en él y contando cuantas veces este se ilumina. La columna o fila que contiene este carácter evoca la onda P300 que es detectada por el ordenador que a partir de esta información es capaz de estimar que fila y columna contiene el carácter. Para hacerlo calcula la media de varias pruebas y detecta que fila o columna ha tenido mayor amplitud de la onda P300 o en cuales se ha dado más veces. La tabla que se usa se puede ver en la Figura 1.6.

Usando esta técnica se conseguía deletrear hasta 7.8 caracteres por minuto.



Figura 1.6 Tabla para deletrear usada en [17]. *Fuente:* Imagen extraída de [17].

1.5 Conclusiones

En este capítulo hemos visto una introducción a los sistemas BCI y se han comentado las diferentes modalidades que existen en el estado del arte.

También se han explicado los principios fisiológicos en los que se apoyan los sistemas BCI. En esta línea cabe destacar el papel que juegan las distintas ondas cerebrales, y como resulta mucho más interesante estudiar lo que ocurre en el cerebro a un alto nivel al recibir distintos tipos de ordenes o estímulos, que estudiar lo que ocurre en cada una de las neuronas particularmente.

Por ultimo también hay que destacar la importancia de estudiar detenidamente las distintas zonas del cerebro, donde se producen con mas intensidad cada una de las ondas estudiadas y las zonas del cerebro que activan cada una de las zonas del cerebro porque estos conocimientos nos permiten aprovechar la diversidad que nos aporta el colocar sensores al rededor de toda la cabeza y tomar decisiones mas certeras.

2 Principales técnicas usadas

Solving problems is a practical art, like swimming, or skiing, or playing the piano; you can learn it only by imitation and practice.

GEORGE POLYA

En este capítulo vamos a describir las principales técnicas y algoritmos que se han empleado para el desarrollo de este trabajo. En concreto vamos a ahondar en CSP, una técnica relacionada con el procesamiento de señal y en LDA, una técnica usada en algoritmos de Machine Learning para hacer clasificaciones, entre otras cosas. El algoritmo implementado en este trabajo está basado en estas dos técnicas que a continuación describiremos en detalle.

2.1 CSP

El acrónimo CSP hace referencia a Common Spatial Patterns que es una técnica matemática usada en el procesamiento de señales para separar señales multivariantes en subcomponentes con distintas varianzas.

La técnica CSP fue propuesta por primera vez con el nombre de *Fukunaga-Koontz Transform* en [18] como extensión de Principal Component Analysis (PCA) y desde entonces ha sido ampliamente usada en BCI para maximizar la distancia entre dos clases de movimientos. Un filtro CSP maximiza la varianza de las señales filtradas de EEG de una clase de movimientos mientras que la minimiza para las señales de la otra clase. El desarrollo de esta técnica surge de manera natural cuando intentamos maximizar la diferencia de varianzas entre las dos señales filtrándolas espacialmente.

En lo que sigue, vamos a suponer que estamos tratando únicamente con dos clases de movimientos, a los que nos vamos a referir con los subíndices L y R (de izquierda y derecha

en inglés). Cuando hablemos de señales filtradas nos estaremos refiriendo a señales filtradas espacialmente y no frecuentemente y además usaremos la siguiente notación:

- T : Número de muestras temporales de un ensayo.
- N : Número de filtros espaciales.
- C : Número de canales (*Número de sensores de la prueba EEG*).
- $\mathbf{X} \in \mathbb{R}^{C \times T}$: Señales de cada uno de los sensores (*filtradas con un filtro paso de banda*).
- $\mathbf{w} \in \mathbb{R}^{C \times 1}$: Filtro espacial simple.
- $\mathbf{y} \in \mathbb{R}^{N \times T}$: Señal filtrada espacialmente. Se definen como:

$$\mathbf{y} = \mathbf{w}^\top \mathbf{X} \quad (2.1)$$

- $\Sigma \in \mathbb{R}^{C \times C}$: Matrices de estimación de covarianza de las señales \mathbf{X} , cada elemento de la matriz se define como:

$$\Sigma(i,j) = \frac{1}{T-1} \sum_{l=1}^T (\mathbf{X}_i(l) - \mu_i)(\mathbf{X}_j(l) - \mu_j) \quad (2.2)$$

Donde $\mathbf{X}_i(l)$ hace referencia al elemento l de la fila i y μ_i hace referencia a la media de la fila i .

- σ^2 : Varianzas de las señales filtradas. Se definen como:

$$\sigma^2 = \mathbb{E}[\mathbf{y}\mathbf{y}^\top] = \mathbf{w}^\top \Sigma \mathbf{w} \quad (2.3)$$

Como se decía, se pretende encontrar un filtro espacial que maximice la diferencia entre las varianzas de las señales de cada clase de movimiento, es decir:

$$\hat{\mathbf{w}} = \arg \max_{\mathbf{w}} \frac{\sigma_L^2 - \sigma_R^2}{\sigma_L^2 + \sigma_R^2} = \arg \max_{\mathbf{w}} \frac{\mathbf{w}^\top (\Sigma_L - \Sigma_R) \mathbf{w}}{\mathbf{w}^\top (\Sigma_L + \Sigma_R) \mathbf{w}} \quad (2.4)$$

El termino del denominador de la ecuación (2.4) se añade para normalizar, porque sin él la diferencia crecería haciendo crecer el módulo del filtro y eso no es lo que se pretende.

La ecuación (2.4) se trata de un cociente de Rayleigh que se resolverá a continuación. Para ello, definimos:

$$\Sigma_T = \Sigma_L + \Sigma_R = \Sigma_T^{1/2} \Sigma_T^{1/2} \quad (2.5)$$

Podemos definir también que:

$$\mathbf{w}^\top = \mathbf{u}^\top \Sigma_T^{-1/2} \quad \text{donde} \quad \mathbf{u}^\top \mathbf{u} = 1 \quad (2.6)$$

Aplicando este par de definiciones sobre (2.4) llegamos a que:

$$\hat{\mathbf{u}} = \arg \max_{\mathbf{u}} \frac{\mathbf{u}^\top \Sigma_T^{-1/2} (\Sigma_L - \Sigma_R) \Sigma_T^{-1/2} \mathbf{u}}{\mathbf{u}^\top \Sigma_T^{-1/2} \Sigma_T \Sigma_T^{-1/2} \mathbf{u}} \quad \text{S.T.} \quad \mathbf{u}^\top \mathbf{u} = 1 \quad (2.7)$$

Donde:

$$\Sigma_T^{-1/2} \Sigma_T \Sigma_T^{-1/2} = I \quad \text{y} \quad \mathbf{u}^\top \mathbf{u} = 1 \quad (2.8)$$

Luego el denominador de la ecuación (2.7) es igual a uno, de tal manera que la ecuación queda:

$$\hat{\mathbf{u}} = \arg \max_{\mathbf{u}} \mathbf{u}^\top \Sigma_T^{-1/2} (\Sigma_L - \Sigma_R) \Sigma_T^{-1/2} \mathbf{u} \quad \text{S.T.} \quad \mathbf{u}^\top \mathbf{u} = 1 \quad (2.9)$$

Teniendo en cuenta que:

$$\Sigma_L + \Sigma_R = \Sigma_T \longrightarrow \Sigma_L - \Sigma_R = 2\Sigma_L - \Sigma_T = \Sigma_T - 2\Sigma_R \quad (2.10)$$

Y recordando (2.6), podemos reescribir el término de la ecuación (2.9) de las dos formas siguientes:

$$\begin{aligned} \mathbf{u}^\top \Sigma_T^{-1/2} (\Sigma_L - \Sigma_R) \Sigma_T^{-1/2} \mathbf{u} &= 2\mathbf{u}^\top \Sigma_T^{-1/2} \Sigma_L \Sigma_T^{-1/2} \mathbf{u} - \mathbf{u}^\top \Sigma_T^{-1/2} \Sigma_L \Sigma_T^{-1/2} \mathbf{u} = \\ &= 2\mathbf{u}^\top \Sigma_T^{-1/2} \Sigma_L \Sigma_T^{-1/2} \mathbf{u} - 1 = \\ &= 1 - 2\mathbf{u}^\top \Sigma_T^{-1/2} \Sigma_R \Sigma_T^{-1/2} \mathbf{u} \end{aligned} \quad (2.11)$$

Por lo que el problema de maximización nos queda:

$$\hat{\mathbf{u}} = \arg \max_{\mathbf{u}} 2\mathbf{u}^\top \underbrace{\Sigma_T^{-1/2} \Sigma_L \Sigma_T^{-1/2}}_{S_1} \mathbf{u} - 1 \quad \text{S.T.} \quad \mathbf{u}^\top \mathbf{u} = 1 \quad (2.12)$$

Estamos ante un problema de autovectores en el que para maximizar debemos de escoger el autovector asociado al mayor autovalor. El autovalor (λ) y autovector (ϕ) que buscamos satisfacen la siguiente ecuación:

$$S_1 \phi_j^{(1)} = \lambda_j^{(1)} \phi_j^{(1)} \quad (2.13)$$

Mediante este método obtenemos un filtro que maximiza la varianza entre las dos clases con lo que tendríamos una única dimensión para discriminar entre clases, ahora bien, el método de CSP consiste en hallar uno o más filtros espaciales que maximicen la diferencia entre las varianzas, y que además sean ortogonales entre sí.

Por lo que se propone que el siguiente problema que se resuelve sea el de minimización, en vez de maximización, es decir que desarrollando de la misma manera llegamos a:

$$\hat{\mathbf{u}} = \arg \min_{\mathbf{u}} 2\mathbf{u}^\top \Sigma_T^{-1/2} \Sigma_L \Sigma_T^{-1/2} \mathbf{u} - 1 \quad \text{S.T.} \quad \mathbf{u}^\top \mathbf{u} = 1 \quad (2.14)$$

Usando la ecuación (2.11) podemos escribir:

$$\hat{\mathbf{u}} = \arg \min_{\mathbf{u}} 1 - 2\mathbf{u}^\top \Sigma_T^{-1/2} \Sigma_R \Sigma_T^{-1/2} \mathbf{u} \quad \text{S.T.} \quad \mathbf{u}^\top \mathbf{u} = 1 \quad (2.15)$$

O lo que es lo mismo:

$$\hat{\mathbf{u}} = \arg \max_{\mathbf{u}} 2\mathbf{u}^\top \underbrace{\Sigma_T^{-1/2} \Sigma_R \Sigma_T^{-1/2}}_{S_2} \mathbf{u} \quad \text{S.T.} \quad \mathbf{u}^\top \mathbf{u} = 1 \quad (2.16)$$

Por lo que procediendo igual que en el caso anterior el siguiente filtro espacial vendría dado por el autovector asociado al máximo autovalor de S_2 , satisfaciendo que:

$$S_2 \phi_j^{(2)} = \lambda_j^{(2)} \phi_j^{(2)} \quad (2.17)$$

De esta manera se obtendría el segundo filtro. El tercer filtro vendría dado por el autovector asociado al segundo autovalor de mayor valor de S_1 , el cuarto filtro por el autovector asociado al segundo autovalor de mayor valor de S_2 y de esta manera se van obteniendo los filtros espaciales sucesivamente. Podemos tomar estas soluciones porque dado que las matrices S_1 y S_2 son simétricas sus autovectores son ortogonales entre sí, y vamos a ver que también los autovectores de ambas matrices son los mismos, por lo que todos los filtros que obtengamos mediante este procedimiento serán ortogonales.

Para demostrar que los autovectores de ambas matrices son los mismos nos vamos a basar en que:

$$S_2 = \Sigma_T^{-1/2} \Sigma_R \Sigma_T^{-1/2} = \Sigma_T^{-1/2} (\Sigma_T - \Sigma_L) \Sigma_T^{-1/2} = I - S_1 \quad (2.18)$$

Donde los autovalores y autovectores de la segunda clase satisfacen que:

$$S_2 \phi_j^{(2)} = (I - S_1) \phi_j^{(2)} = \lambda_j^{(2)} \phi_j^{(2)} \quad (2.19)$$

O lo que es lo mismo:

$$S_1 \phi_j^{(2)} = (1 - \lambda_j^{(2)}) \phi_j^{(2)} \quad (2.20)$$

De (2.13) y de (2.20) tenemos que

$$\phi_j^{(2)} = \phi_j^{(1)} \quad (2.21)$$

y que

$$\lambda_i^{(1)} = (1 - \lambda_j^{(2)}) \quad (2.22)$$

Entonces vemos que cada autovector de S_1 es igual a otro autovector de S_2 . Ya solo nos queda ver que teniendo en cuenta que el valor de los autovalores debe de estar acotado entre cero y uno¹ y debido a la relación entre los autovalores que nos da la ecuación (2.22) podemos decir que: si ordenamos los autovalores de la misma forma, por ejemplo de mayor a menor tenemos que el autovector asociado al mayor autovalor de S_1 es igual al autovector asociado al menor autovalor de S_2 , el segundo autovector de S_1 sería igual al penúltimo de S_2 y así sucesivamente. Por lo que podemos escribir que:

$$\phi_j^{(2)} = \phi_{C-j}^{(1)} \quad (2.23)$$

Donde C sería el número total de autovectores obtenidos, es decir el número de canales usados.

¹ Es así porque las matrices de covarianza son por definición semidefinidas positivas y de (2.22): $\lambda_i^{(1)} + \lambda_j^{(2)} = 1$

Una vez llegados a este punto vemos que no hace falta resolver los dos problemas, ya que resolviendo uno de ellos obtenemos las soluciones del otro directamente.

Finalmente, definimos la solución² del problema de dimensión N en (2.24), donde ϕ son los autovectores de S_1 :

$$\hat{\mathbf{W}} = \Sigma_T^{-1/2} [\phi_1 \dots \phi_{N/2}, \phi_{C-N/2} \dots \phi_C] \quad (2.24)$$

Teniendo en cuenta que hemos supuesto:

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_C \quad (2.25)$$

2.1.1 Extracción de características

Finalmente, una vez que se ha reducido la dimensión de la entrada usando los filtros \mathbf{w} solamente queda extraer las características que consiste unicamente en calcular la varianza de cada uno de los vectores de salida.

$$z_i = \text{var}(\mathbf{y}_i) = \text{var}(\mathbf{w}_i \mathbf{X}), \quad i \in [1, N] \quad (2.26)$$

Donde la operación $\text{var}(\cdot)$ hace referencia al operador varianza.

Como ya se ha comentado, de esta manera obtenemos una característica, si se realiza esta operación para cada solución obtenida con CSP se obtiene el vector de características que usará el clasificador para estimar la clase de cada ensayo. Es decir:

$$\mathbf{z} = [z_1, z_2, \dots, z_N] \quad (2.27)$$

En las Figura 2.1, Figura 2.2 y Figura 2.3 mostramos las características extraídas tras haber aplicado CSP con distinto número de dimensiones a unas señales de EEG que disponían de 22 sensores, durante las cuales el usuario pensaba en dos clases de movimiento.

² Podríamos tomar los autovectores de S_2 o de $\Sigma_T^{-1/2}(\Sigma_L - \Sigma_R)\Sigma_T^{-1/2}$ y las tres soluciones serían equivalentes

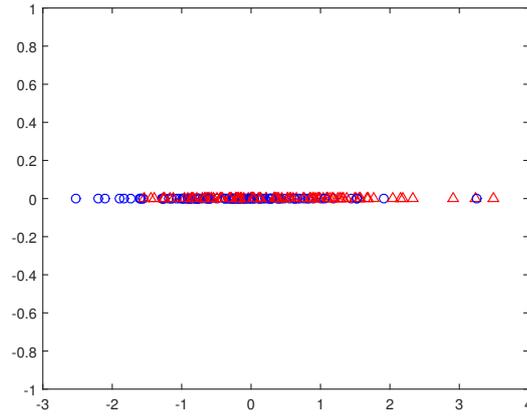


Figura 2.1 Ejemplo de CSP tomando una única dimensión.

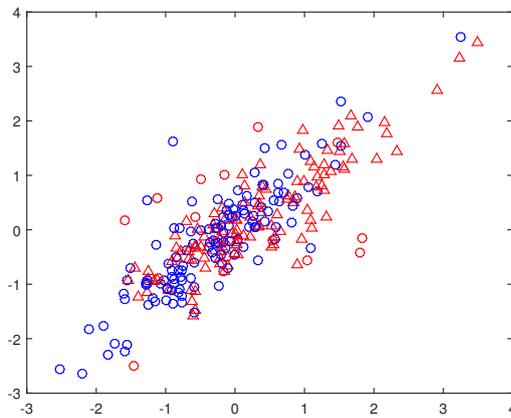


Figura 2.2 Ejemplo de CSP tomando dos dimensiones.

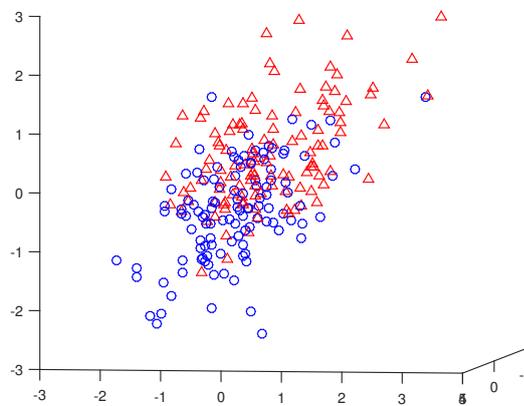


Figura 2.3 Ejemplo de CSP tomando una tres dimensiones.

2.2 LDA

LDA es un acrónimo de *Linear Discriminant Analysis* y se trata de una técnica de *Machine Learning* que se usa para realizar clasificación lineal supervisada. Los problemas de clasificación se tratan de problemas muy comunes, en los que tenemos una serie de observaciones que se pueden dividir en grupos o clases. El problema básicamente consiste en asignar la clase correcta a cada observación.

Los clasificadores pueden ser lineales y no lineales, y dentro de los no lineales podemos encontrar múltiples tipos como los polinómicos, gaussianos, etc.. En ambos tipos de clasificadores, podemos clasificar entre dos clases o más. Los clasificadores lineales son aquellos que basan su decisión según algún hiperplano³ asignando a cada clase un lado del subespacio evaluado. Los clasificadores lineales multiclase dividen el subespacio usando hiperplanos y asigna las clases a las regiones entre hiperplanos [22].

Existen multitud de técnicas de clasificación, entre las que cabe destacar SVM que permite realizar todo tipo de clasificaciones usando distintos kernels. En este caso se ha escogido LDA porque se trata de una técnica simple, fácil de comprender y muy estable, que nos va a permitir realizar el tipo de clasificación que necesitamos sin la utilización de ningún otro parámetro. Otras técnicas como SVM necesitan de parámetros de configuración que también debemos aprender, complicando enormemente el problema, ya que los parámetros óptimos pueden variar entre usuarios o experimentos.

En nuestro caso queremos entrenar el clasificador de tal manera que este sea capaz de estimar la clase a la que pertenece una observación a partir de observaciones de las cuales conocemos la clase, por lo que estamos ante un problema de clasificación supervisada. Los casos en los que no se conoce la clase de ninguna muestra a priori se conocen como problemas de clasificación no supervisada.

Por último destacar que vamos a realizar una clasificación entre dos clases, porque aunque los datos disponen de cuatro clases distintas entre las que diferenciar, todos los clasificadores se entrenan para hacer la clasificación entre dos clases.

LDA se basa en suponer que tenemos dos clases que siguen una distribución normal. Para cada clase se modelan los parámetros de la media y la varianza para conseguir la distribución que mejor la describa y posteriormente se usa el teorema de Bayes para calcular la probabilidad de pertenecer a cada una de las clases.

2.2.1 Teorema de Bayes

El teorema de Bayes es ampliamente usado en matemáticas y estadística y proporciona una relación entre la probabilidad de un evento dado (A) sabiendo que otro evento (B)

³ En geometría, un hiperplano es una extensión del concepto de plano. En un espacio unidimensional (como una recta), un hiperplano es un punto: divide una línea en dos líneas. En un espacio bidimensional (como el plano xy), un hiperplano es una recta: divide el plano en dos mitades.

ha ocurrido con la probabilidad de B sabiendo que ha ocurrido A . De manera que la probabilidad de $p(A|B)$ viene dada por:

$$p(A|B) = \frac{p(B|A)p(A)}{p(B)} \quad (2.28)$$

Centrandonos en nuestro problema, supongamos que queremos asignar una clase (c_i) de entre C clases, donde $C \geq 2$, a una observación (\mathbf{z}). Supongamos también que conocemos la probabilidad a priori de que se de cada una de las clases $p(C_i)$, entonces el teorema de Bayes define que:

$$p(c_i|\mathbf{z}) = \frac{p(c_i)p(\mathbf{z}|c_i)}{\sum_{l=1}^C p(c_l)p(\mathbf{z}|c_l)} \quad (2.29)$$

Esta formula puede representarse como:

$$\textit{Posteriori} = \frac{\textit{Prior} \cdot \textit{Verosimilitud}}{\textit{Evidencia}} \quad (2.30)$$

Donde hemos llamado *Prior* a la probabilidad a priori de la clase i , *Verosimilitud* al término que indica la probabilidad de obtener esta observación en concreto si la clase era i y hemos llamado *Evidencia* a un término que hace la función de factor de escala, y que regulariza el resultado, por lo que realmente no es importante a la hora de tomar una decisión.

Esta formula nos es de gran utilidad porque nos ayuda a tomar una decisión ya que tiene sentido asignar la clase con mayor probabilidad a posteriori.

$$\hat{c} = c_i \iff p(c_i|\mathbf{z}) \geq p(c_l|\mathbf{z}), \forall (l \neq i) \quad (2.31)$$

2.2.2 LDA para dos clases

Como se ha dicho en el apartado 2.2.1, la técnica LDA consiste en usar el teorema de Bayes para estimar la clase a la que pertenece cada observación suponiendo que las dos clases tienen una función de densidad de probabilidad gaussiana con igual covarianza y distintas medias. Esta suposición nos permite calcular tanto la *Verosimilitud* como la *Evidencia*, y supondremos conocido también el *prior* de cada clase.

De manera que tendremos:

$$\textit{Prior}(c_1) = p(c_1), \textit{Prior}(c_2) = 1 - p(c_1) \quad (2.32)$$

$$\textit{Verosimilitud} = p(\mathbf{z}|c_l) \quad (2.33)$$

$$\textit{Evidencia} = \sum_{l=1}^2 p(c_l)p(\mathbf{z}|c_l) \quad (2.34)$$

Donde $p(\mathbf{z}_n|c_l)$ es la probabilidad de que una nueva observación pertenezca a clase l .

Dado que hemos considerado que las observaciones tienen una función de densidad de probabilidad gaussiana, la probabilidad $p(\mathbf{z}_n|c_l)$ viene dada por:

$$p(\mathbf{z}_n|c_l) = \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} e^{j(-\frac{1}{2}(\mathbf{z}_n-\mu_l)^\top \Sigma^{-1}(\mathbf{z}_n-\mu_l))} \quad (2.35)$$

Entonces podemos decir que para una observación \mathbf{z}_n asignaremos la clase 1 si se cumple que:

$$p(c_1)p(\mathbf{z}_n|c_1) \geq (1-p(c_1))p(\mathbf{z}_n|c_2) \quad (2.36)$$

En vez de computar la condición anterior para cada muestra, estamos más interesados en encontrar la frontera entre las dos regiones de decisión por lo que pretendemos, es encontrar el hiperplano que cumpla la igualdad de la ecuación (2.36). Para hacerlo vamos a despejar \mathbf{z}_n de la ecuación (2.36), en la que se han simplificado los denominadores.

$$p(c_1)p(\mathbf{z}_n|c_1) = (1-p(c_1))p(\mathbf{z}_n|c_2) \quad (2.37)$$

Desarrollando esta ecuación:

$$1 = \frac{(1-p(c_1))p(\mathbf{z}_n|c_2)}{p(c_1)p(\mathbf{z}_n|c_1)} \quad (2.38)$$

$$1 = \frac{(1-p(c_1))e^{j(-\frac{1}{2}(\mathbf{z}_n-\mu_2)^\top \Sigma^{-1}(\mathbf{z}_n-\mu_2))}}{p(c_1)e^{j(-\frac{1}{2}(\mathbf{z}_n-\mu_1)^\top \Sigma^{-1}(\mathbf{z}_n-\mu_1))}} \quad (2.39)$$

$$0 = \log \left(\frac{(1-p(c_1))e^{j(-\frac{1}{2}(\mathbf{z}_n-\mu_2)^\top \Sigma^{-1}(\mathbf{z}_n-\mu_2))}}{p(c_1)e^{j(-\frac{1}{2}(\mathbf{z}_n-\mu_1)^\top \Sigma^{-1}(\mathbf{z}_n-\mu_1))}} \right) \quad (2.40)$$

$$0 = \log \left(\frac{(1-p(c_1))e^{j(-\frac{1}{2}(\mathbf{z}_n-\mu_2)^\top \Sigma^{-1}(\mathbf{z}_n-\mu_2))}}{p(c_1)e^{j(-\frac{1}{2}(\mathbf{z}_n-\mu_1)^\top \Sigma^{-1}(\mathbf{z}_n-\mu_1))}} \right) \quad (2.41)$$

$$0 = \log \left(\frac{1-p(c_1)}{p(c_1)} \right) - \frac{1}{2}(\mathbf{z}_n-\mu_2)^\top \Sigma^{-1}(\mathbf{z}_n-\mu_2) + \frac{1}{2}(\mathbf{z}_n-\mu_1)^\top \Sigma^{-1}(\mathbf{z}_n-\mu_1) \quad (2.42)$$

$$0 = (\mu_2-\mu_1)^\top \Sigma^{-1}\mathbf{z}_n - \frac{1}{2}\mu_2^\top \Sigma^{-1}\mu_2 + \frac{1}{2}\mu_1^\top \Sigma^{-1}\mu_1 + \log \left(\frac{1-p(c_1)}{p(c_1)} \right) \quad (2.43)$$

En la ecuación (2.43) obtenemos la solución al problema ya que esta es la ecuación general del hiperplano que separa las dos clases. A partir de esta ecuación⁴, obtenemos los *coeficientes* (**a**) y la *el termino independiente* (**b**). Están definidos por:

$$\mathbf{a} = (\mu_2-\mu_1)^\top \Sigma^{-1} \quad (2.44)$$

$$\mathbf{b} = \frac{1}{2}\mu_1^\top \Sigma^{-1}\mu_1 - \frac{1}{2}\mu_2^\top \Sigma^{-1}\mu_2 + \log \left(\frac{1-p(c_1)}{p(c_1)} \right) \quad (2.45)$$

⁴ Teniendo en cuenta que la ecuación general de un hiperplano viene dada por: $0 = \mathbf{a}\mathbf{z}_n + \mathbf{b}$

Se puede comprobar fácilmente que la ecuación (2.45) también se puede escribir como:

$$\mathbf{b} = -\frac{\mathbf{a}^\top(\boldsymbol{\mu}_1 + \boldsymbol{\mu}_2)}{2} + \log\left(\frac{1 - p(c_1)}{p(c_1)}\right) \quad (2.46)$$

La estimación de la clase de una observación se puede estimar simplemente evaluando el signo de:

$$\hat{c} = \text{sgn}(\mathbf{a}^\top \mathbf{z}_n + \mathbf{b}) \quad (2.47)$$

Por lo que, por la manera en la que hemos planteado el problema: si el resultado tiene signo positivo la clase estimada sería 1 y si el signo fuese negativo la clase estimada sería 2.

Para ilustrar los resultados que se obtienen usando LDA se ha incluido la Figura 2.4 en la que se muestran dos nubes de puntos de dos clases, y la recta que separa las dos regiones. En la Figura 2.5 se muestran estas mismas nubes de puntos junto con sus respectivas FDP, también se ha representado un plano que representa la separación del espacio tanto de las FDP como de las nubes de puntos.

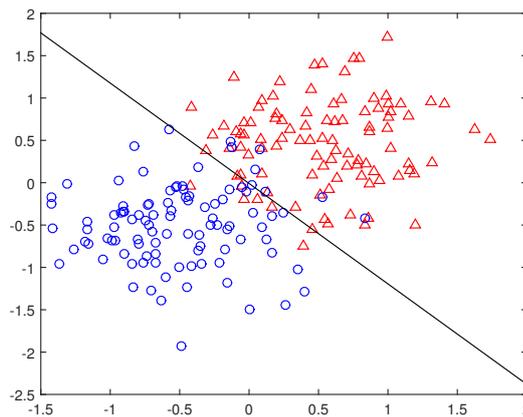


Figura 2.4 Nubes de puntos separadas con LDA.

2.3 Conclusiones

En este segundo capítulo hemos visto como todos los algoritmos usados en sistemas BCI se pueden dividir en dos bloques. Que siempre vamos a tener un bloque de extracción de características y otro bloque de clasificación. En este sentido se han estudiado las dos técnicas matemáticas en las que se apoyan la mayoría de los sistemas BCI para cada uno de los bloques que se han comentado.

Este estudio nos va a permitir entender el funcionamiento del algoritmo que se ha desarrollado en este trabajo y comprender la base de otros algoritmos como el que se ha explicado en el capítulo 5.

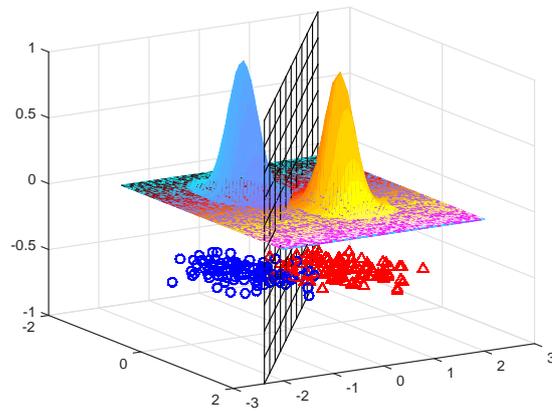


Figura 2.5 Nubes de puntos separadas con LDA y FDP de las nubes de puntos.

3 Conjunto de datos

Emancipate yourself from mental slavery, none but ourselves can free our minds.

BOB MARLEY

Para la realización de este trabajo, se han usado los datos 2a de la cuarta competición de BCI¹. Este conjunto de datos proporcionan señales de EEG que han sido adquiridas siguiendo los procedimientos adecuados en diferentes sujetos y nos permiten comprobar la validez de nuestro algoritmo. La documentación relacionada con la adquisición de los datos la podemos encontrar en [11].

3.1 Datos experimentales

El conjunto de datos con el que se ha trabajado está constituido por las señales de EEG de nueve sujetos. El experimento que se llevó a cabo consistía en que el paciente tenía que imaginar cuatro tipos diferentes de movimientos motores: el movimiento de la mano izquierda (*clase 1*), de la mano derecha (*clase 2*), de los dos pies a la vez (*clase 3*) y de la lengua (*clase 4*). Para cada paciente se grabaron dos sesiones en diferentes días. Cada sesión se compone de seis partes y cada parte a su vez se compone de 48 ensayos, 12 para cada tipo de movimiento. En total se tendrán 288 pruebas por sesión.

Al inicio de cada sesión se tomaban aproximadamente cinco minutos para estimar la influencia del movimiento ocular. Esto se realizaba mediante el Electrooculograma (EOG). El EOG es una prueba similar al EEG en la que los sensores se colocan alrededor de los ojos para grabar la actividad ocular del sujeto. Esta prueba se realiza para poder eliminar

¹ La BCI Competition IV se trata de una competición en la que se proporcionaron las grabaciones de EEG de unos sujetos determinados (la mitad de ellas estaban clasificadas y la otra mitad no) para que los participantes pudieran estimar en base al conjunto de datos que no estaba clasificado, la clase del movimiento imaginario que realizaba el paciente. Para más información consultar: <http://www.bbci.de/competition/iv>

los artefactos producidos por los movimientos de los ojos en la prueba de EEG. Lo que se hace al inicio de la sesión es realizar tres pruebas que consisten en estar 2 minutos con los ojos abiertos, mirando a una cruz en la pantalla, un minuto con los ojos cerrados y un minuto moviendo los ojos libremente. Los diferentes bloques que componen esta prueba se representan en la Figura 3.1:



Figura 3.1 Esquema de las sesiones.

A continuación, tras la prueba de EOG los usuarios se sentaban en un sillón delante de una pantalla de ordenador. Al principio de cada ensayo ($t = 0s$) se ilustraba una cruz en el centro de la pantalla en negro para así fijar su visión, además se avisaba al sujeto de que comenzaría un nuevo experimento con una señal sonora (*Beep*). Transcurridos dos segundos del comienzo de la prueba se ilustraba una flecha que apuntaba hacia la izquierda, derecha, abajo o arriba, que se corresponden con las cuatro clases de movimientos imaginarios. La flecha permanecía en la pantalla durante 1.25s y era la señal para que el usuario empezase a realizar el movimiento imaginario. El usuario debía de prolongar la realización del experimento hasta que desapareciese la cruz a los 6 segundos. Finalmente había un descanso durante el cual la pantalla permanecía en negro. Este procedimiento se ilustra en la Figura 3.2 que se muestra a continuación::

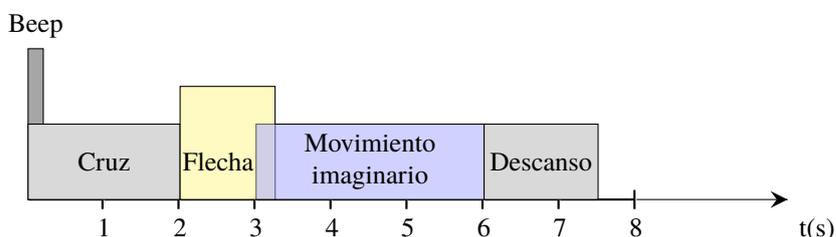


Figura 3.2 Esquema de los ensayos.

3.2 Grabación de los datos

El conjunto de datos utilizados consta de 25 señales, de las cuales 22 se corresponden con los electrodos situados alrededor de la cabeza del sujeto (con una distancia entre ellos de 3.5 cm). Las señales se obtuvieron usando el hueso mastoideo² derecho como tierra y el hueso mastoideo izquierdo como referencia. La señal se digitalizó con una frecuencia de 250Hz y se le aplicó un filtro paso banda entre 0.5Hz y 100Hz. Además se habilitó un filtro

² El hueso mastoideo o también llamado la apófisis mastoides o proceso mastoideo, es un hueso localizado justo detrás del oído.

notch³ para eliminar la influencia de la línea eléctrica. La sensibilidad del amplificador se estableció en $1\mu V$ y el montaje se representa en Figura 3.3.

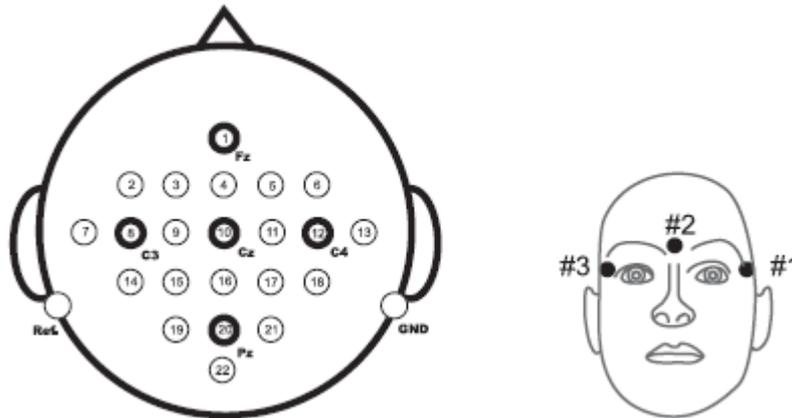


Figura 3.3 Esquema del montaje para la grabación de las señales EEG.

Como podemos intuir en la imagen de la derecha de la Figura 3.3, las otras tres señales restantes se corresponden con señales EOG y se han grabado para que puedan ser usadas en la eliminación de artefactos por el movimiento ocular, como ya se comentó al principio del apartado 3.1, pero no deben de ser usadas para estimar la clase de movimiento. Estas señales fueron grabadas en las mismas condiciones que las anteriores pero con una sensibilidad de $1mV$.

El conjunto de datos fue revisado por un experto y los ensayos con artefactos fueron marcados debidamente.

3.3 Descripción de los datos

Los datos que usaremos se encuentran en el formato estándar para señales biomédicas, es decir, en formato General Data Format (GDF). Para que puedan ser consumidos por Matlab es necesario el uso de un software adicional. Para ello usaremos Biosig⁴, una librería de software abierto para el procesado de señales que proporciona un toolbox para Matlab.

El conjunto de datos consta de un archivo por cada sujeto que realizó las pruebas (9 sujetos en total) y por cada sesión, obteniendo así un total de 18 archivos. Tras importar estos archivos a Matlab, obtenemos una matriz s que contiene las 25 señales proveniente de los sensores y un struct⁵ que contiene distinta información relativa al sujeto que ha realizado la prueba, del archivo, de las muestras, etc...

³ El filtro notch o también conocido como filtro suprime banda o de rechazo de banda, es un filtro electrónico que no permite el paso de señales cuyas frecuencias se encuentran comprendidas entre las frecuencias de corte superior e inferior.

⁴ <http://biosig.sourceforge.net/>

⁵ Tipo de dato de Matlab

La matriz de las señales s tiene dos dimensiones: tendremos una columna por cada sensor y cada columna tendrá entre 670000 y 680000 muestras aproximadamente, es decir, tendremos la grabación de la sesión completa que dura alrededor de 45 minutos, dato que tendremos que tener en cuenta a la hora de implementar el algoritmo ya que este deberá de recortar los trozos de interés de las señales, es decir, los ensayos.

Por otro lado en el struct, como ya hemos comentado antes, tenemos información referente a distintos aspectos. En total encontramos 42 campos, pero tan solo se han usado unos pocos en el desarrollo de este algoritmo. Los citamos a continuación:

- *EVENT*: Este campo es de tipo struct y consta a su vez de los siguientes campos:
 - *SampleRate*: Dato de tipo entero que nos proporciona la tasa de muestreo (en Hz) a la que han sido adquiridos los datos.
 - *TYP*: Dato de tipo matriz, con una columna en la que se indica el tipo de evento, para cada uno de los eventos ocurridos durante la señal. Los tipos de eventos que se pueden producir se exponen en Tabla 3.1.

Tabla 3.1 Tipos de eventos.

Evento	Descripción
276	Inicio de tiempo de espera con los ojos abiertos
277	Inicio de tiempo de espera con los ojos cerrados
768	Comiendo de un nuevo ensayo
769	Flecha hacia la izquierda (clase 1)
770	Flecha hacia la derecha (clase 2)
771	Flecha hacia abajo (clase 3)
772	Flecha hacia arriba (clase 4)
783	Clase desconocida
1023	Ensayo descartado
1072	Inicio de tiempo de espera moviendo los ojos
32766	Comienzo de una nueva parte (hay seis en cada sesión)

- *POS*: Dato de tipo matriz, con una columna del mismo tamaño que *TYP*. Indica la muestra temporal de la señal s en la que empiezan los eventos indicados en *TYP*.
- *DUR*: Dato de tipo matriz, con una columna del mismo tamaño que *TYP*. Indica la duración en número de muestras que tienen los eventos indicados en *TYP*.
- *Label*: Dato de tipo *cell* con 25 celdas que indican el tipo de señal (EEG o EOG) al que se corresponden cada una de las columnas de s . Además, en algunos casos también nos da información acerca del sensor del que proviene la señal, como puede ser el *FZ* o el *PZ*, aunque si se quiere saber con que sensor se corresponde cada una de las columnas es mejor consultar la Figura 3.3.

- *Classlabel*: Dato de tipo matriz, con una columna de 288 posiciones, una por cada ensayo. Cada posición contiene un 1, 2, 3 o 4, haciendo referencia a la clase de movimiento que se indicó. También puede contener el valor *NaN* si la clase es desconocida.
- *TRIG*: Dato de tipo matriz, con una columna de 288 posiciones, una por cada ensayo. Cada posición contiene la posición de la señal *s* en la que comienzan los ensayos.
- *ArtifactSelection*: Dato de tipo matriz, con una columna de 288 posiciones, una por cada ensayo. Cada posición contiene un 1 o un 0 indicando si el ensayo ha sido contaminado por artefactos.
- *CHANTYP*: Cadena de caracteres de 25 posiciones en las que se indican que canales son de EEG y cuales de EOG con las letras *E* y *O*.

Como se indicó en el apartado ?? cada sujeto realizó dos sesiones. En los datos descargados, los campos que indican la clase de movimiento solamente están completos en una sesión de cada sujeto, ya que recordemos que el objetivo era realizar una competición, por lo que la clase de movimiento de los ensayos de la otra sesión eran desconocidos en principio para poder determinarlo. Es por esta razón por la que el campo *Classlabel* está relleno con *NaN* y en las posiciones de *TYP* (referentes a la flecha que se represento) están fijadas al valor 783. Una vez que la competición acabó se proporcionó las clases de movimientos de los ensayos en un formato *.mat* para que pudiesen ser consumidos por Matlab directamente. Los archivos *.mat* contenían simplemente una matriz equivalente al campo *Classlabel* para poder comprobar la validez del algoritmo implementado.

3.4 Conclusiones

Este capítulo se considera de gran importancia si se va a analizar el algoritmo desarrollado y si se quiere conocer las circunstancias bajo las cuales se han tomado las señales que con las que se ha trabajado en este proyecto.

Conocer el entorno en el que se han realizado las grabaciones y el patrón de pruebas que se han realizado nos permite extraer los pedazos de señal que se consideran más importantes, dándonos la libertad de decidir donde empiezan y acaban los pedazos de interés. Además como ya se ha comentado el usar datos grabados correctamente y analizados por expertos reduce el grado de incertidumbre al que nos enfrentamos cuando se nos plantea cualquier problema y por último el que se traten de datos disponibles para todos los desarrolladores y que han sido ampliamente usados también nos permite contrastar los resultados con los de otras publicaciones.

4 Algoritmo implementado

Life is what happens while you are busy making other plans

JHON LENNON

El objetivo de este trabajo es implementar un algoritmo para la estimación de dos clases de pensamiento a partir de señales de EEG. En este capítulo se explica detalladamente el algoritmo desarrollado así como las diferentes partes que lo componen y los aspectos más relevantes del mismo.

4.1 Visión general

En la Figura 4.1 se ha representado el diagrama de bloques del algoritmo con el que se ha trabajado, que nos ha servido tanto para ir realizando las distintas pruebas, como para obtener los resultados finales que se exponen en el capítulo 6.

Tras la realización del trabajo se ha descubierto que una técnica similar al algoritmo que se ha desarrollado ya ha sido propuesta en [6] y se la conoce como Filter Bank Common Spatial Patterns (FBCSP).

El algoritmo ha sido realizado en *Matlab* empleando las técnicas explicadas en el capítulo 2. Básicamente se compone de dos bucles *for*. El primer bucle *for* (bloque superior a partir de ahora) se ha creado para analizar a cada usuario por separado, es decir, nos proporcionará una solución distinta para cada sujeto. En cuanto al segundo bucle *for* (bloque inferior a partir de ahora), se ha creado para calcular varias soluciones para un mismo usuario y escoger la mejor de ellas.

A rasgos generales, el programa comienza con la declaración de los parámetros que van a ser usados durante el desarrollo del algoritmo, tras lo cual se entra en el bloque superior

para trabajar con cada sujeto por separado. En el bloque superior, se preparan los datos para que puedan ser tratados en el bloque inferior, que comienza barajando los datos para obtener una solución distinta en cada ejecución. Las señales son pasadas por un banco de filtros para analizar cada banda por separado, a partir de este momento se trata cada banda como señales independientes. Se calcula el filtro espacial óptimo y se extraen las características para cada banda, tras lo cual se crea el clasificador usando todas las características de todas las bandas y se prueba la solución. Estos resultados son almacenados. Al salir de este bloque se escoge la solución que mejores resultados haya proporcionado y se emplea para estimar la clase de los movimientos imaginarios que se habían reservado previamente para ser usados únicamente como datos de test.

A continuación mostramos el diagrama de bloques del algoritmo implementado. Los distintos bloques de la Figura 4.1 han sido numerados en el orden en el que se van a ir explicando a lo largo de este capítulo para poder comprender mejor cual es la labor de cada uno de ellos

Como se expuso en el capítulo 3, en los datos que se han estado usando se encuentran cuatro clases de movimiento imaginarios, pero como se ha explicado en este capítulo, el algoritmo que se ha desarrollado solamente es capaz de distinguir entre dos clases, por lo que se ha optado por escoger dos de las clases para hacer las pruebas. En la Tabla 4.1 se muestran los pares de clases de movimientos que se han contemplado para probar el algoritmo.

Tabla 4.1 Pares de clases de movimiento.

Clases	Tipos de movimiento
1-2	Mano izquierda contra la mano derecha
1-3	Mano izquierda contra las dos piernas
1-4	Mano izquierda contra la lengua
2-3	Mano derecha contra las dos piernas
2-4	Mano derecha contra la lengua
3-4	Las dos piernas contra la lengua

4.2 Elección de parámetros

El programa comienza almacenando los valores que deseamos aplicar en la ejecución del algoritmo. En esta parte del programa se elige el número de bandas que se van a usar. El número de características que se van a extraer de cada banda, el orden del filtro que se usará para descomponer las bandas y número de CV que queremos realizar entre otros parámetros.

También se inicializan las variables que se usan en el primer bucle *for*.

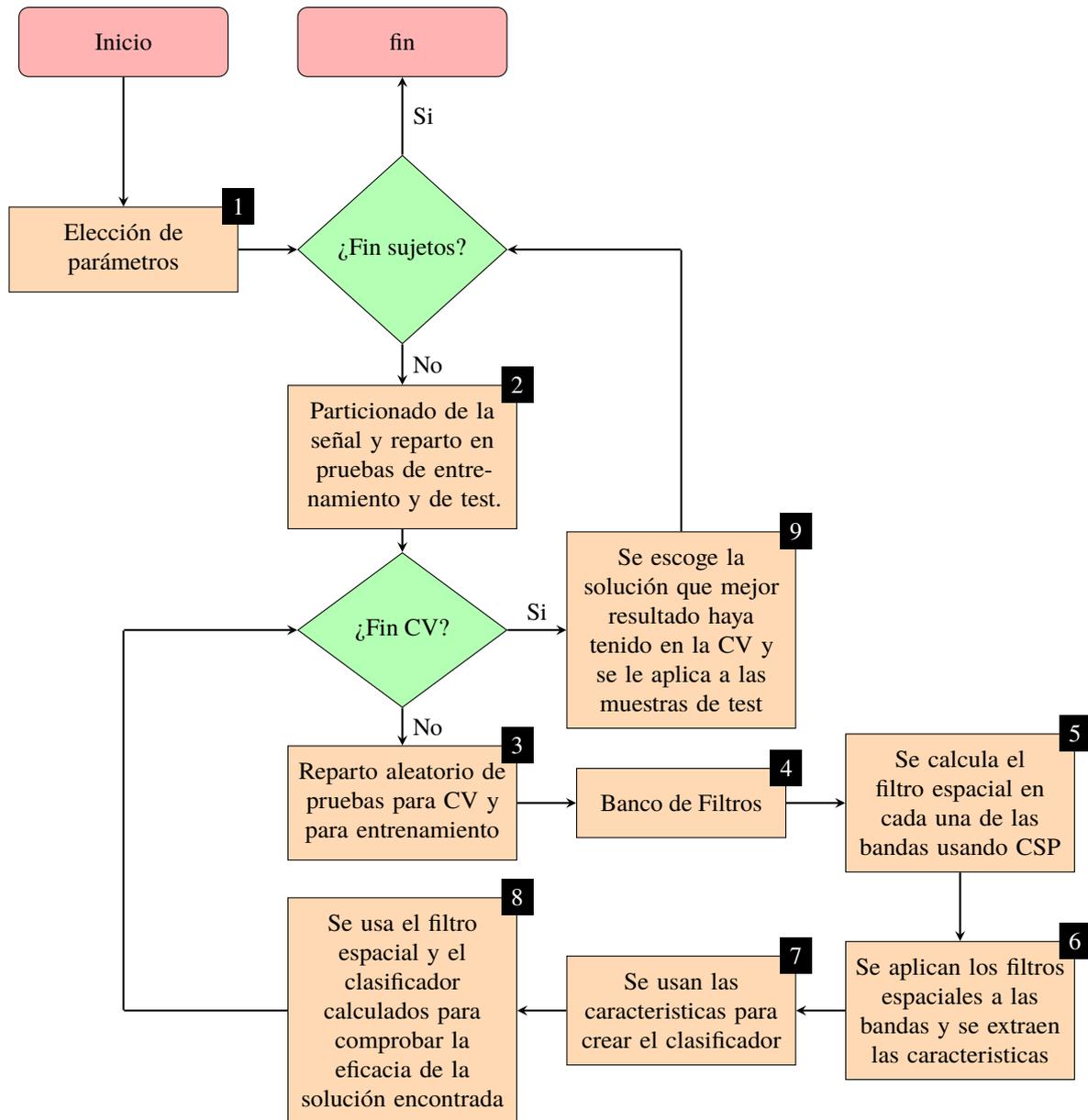


Figura 4.1 Diagrama de bloques del algoritmo desarrollado.

4.3 Partición de señales

Tras entrar en el primer bucle *for*, se preparan los datos de cada usuario para poderlos usar. En primer lugar se cargan las señales de las que extraeremos cada uno de los ensayos. Se cargan las dos sesiones de cada sujeto, la que originalmente era para entrenamiento y la que era para test junto con sus correspondientes clases y se le asigna a cada uno de los ensayos.

De cada ensayo solamente se toma un trozo temporal de la señal, porque como se explicó en la sección 3.1 del capítulo 3 el sujeto solamente está realizando el movimiento imaginario durante un periodo de tiempo del ensayo. En este trabajo se ha tratado de tomar el periodo

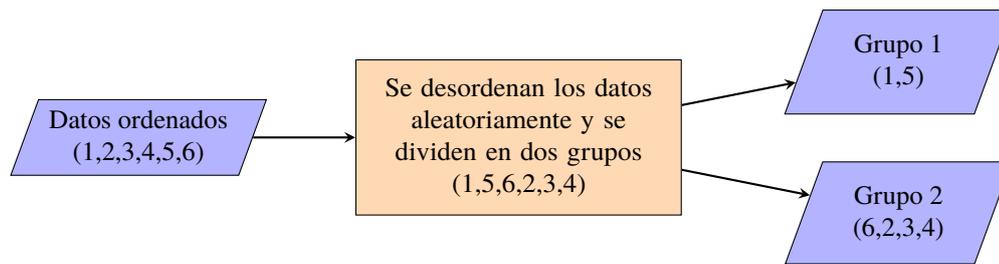


Figura 4.2 Esquema de método usado para separar en grupos aleatorios.

de tiempo que contiene más información y finalmente se ha optado por tomar la señal desde el segundo 2.3 hasta el segundo 5.3 después del inicio del ensayo.

Los trozos de ensayos son organizados en una matriz, en la cual se concatenan los trozos de las dos sesiones y se barajan para mezclarlos. Posteriormente, se separan en dos grupos (que contendrán ensayos de ambas sesiones), tal y como se muestra en la Figura 4.2. Uno de estos dos grupos se destinará a entrenamiento y CV, mientras que el segundo grupo se destinará a test. La cantidad de ensayos en cada uno de estos grupos dependerá de los valores establecidos en el Bloque 1, donde se especificó el porcentaje de ensayos que se destinarían a test y el porcentaje que se destinaría a CV. El grupo que contiene los ensayos que se destinarán a test, se reservará, mientras que el resto de los ensayos (el grupo 2) se pasará al bucle *for* inferior.

4.4 Iniciación de CV

Al empezar cada iteración del bloque inferior, partimos con los datos destinados a entrenamiento y a CV. Estos datos se vuelven a mezclar, siguiendo el mismo esquema que en la Figura 4.2. Así, dividimos la matriz que contiene los ensayos en otros dos grupos: el de CV y el de entrenamiento. En este caso, se reserva el grupo de ensayos de CV y se sigue trabajando con el grupo de entrenamiento. Así, en cada iteración del bucle, los grupos de CV y de entrenamiento serán diferentes.

Por último, se descartan las muestras correspondientes a las señales de EOG, de manera que de aquí en adelante, estaremos trabajando con grupos de 22 señales y no 25 como en un principio.

4.5 Banco de filtros

En este bloque, se toman las muestras de entrenamiento y a cada una por separado se le aplica un banco de filtros de análisis, descomponiendo la señal en tantas señales como número de bandas se indicase en el bloque 1.

El banco de filtros se ha implementado usando tantos filtros paso de banda como fueran

necesarios. Los filtros paso de banda que se han usado son filtros Butterworth¹ de orden variable y de tipo paso de banda. Estos filtros han sido calculados usando la función *butter* de Matlab.

Tabla 4.2 Bandas de frecuencias usadas.

Limite entre bandas	Descripción
8-30	Se toma una sola banda de frecuencia que contiene a las bandas Alfa y Beta.
8-14-30	Dos bandas de frecuencia: bandas Alfa y Beta.
4-8-14-30	Tres bandas: bandas Theta, Alfa y Beta.
4-8-15-30-40	Cuatro bandas: bandas Theta, Alfa, Beta y parte baja de la banda Gamma.
4-10-20-30-40	Cuatro bandas que cubren el rango 4Hz-30Hz, sin usar las bandas cerebrales.
8-12-16-20-24-28-32	Seis bandas equiespaciadas que cubren el rango 4Hz-32Hz

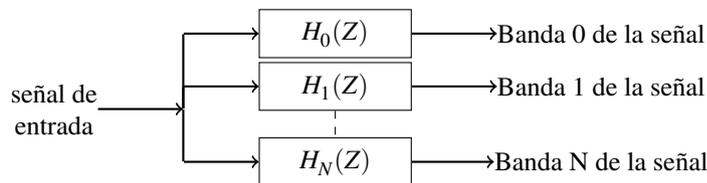


Figura 4.3 Esquema de banco de filtros.

En ningún caso se han tomado frecuencias superiores a los 40Hz ya que los cascos de EEG no tienen sensibilidad a frecuencias mayores de 40 Hz[20] [7].

4.6 Cálculo de filtros espaciales

Se toma cada una de las bandas de la señal y para ellas se calcula el filtro CSP como se explicó en el capítulo 2. En este bloque se recibe una matriz con todos los ensayos del grupo de training y otra con la clase (que indica el movimiento imaginario que se estaba realizando) de cada ensayo. Se separan las dos clases de ensayos y se calcula la covarianza de las señales correspondiente a cada clase. Hay que tener en cuenta que cada ensayo está compuesto por 22 señales, por lo que la matriz de covarianza de cada clase tendrá dimensión 22×22 . Tras esto se calcula la suma y la resta de ambas matrices, que son usadas para calcular los autovectores de las dos matrices conjuntas. Esto puede hacerse en Matlab fácilmente con el comando *eig(A,B)* donde *A* y *B* son las matrices suma y diferencia.

Finalmente se seleccionan tantos autovectores como características se hayan especificado,

¹ El filtro Butterworth es uno de los filtros electrónicos básicos, cuyo objetivo es obtener la respuesta más plana que sea posible hasta la frecuencia de corte. Es decir, la salida se mantiene constante casi hasta la frecuencia de corte, luego disminuye a razón de $20n$ dB por década, donde n es el número de polos del filtro.

escogiendo los autovectores asociados a los autovalores de menor y mayor valor y se invierte la matriz formada por los autovectores seleccionados para obtener el filtro espacial, que en realidad sería una matriz de *Características* \times 22.

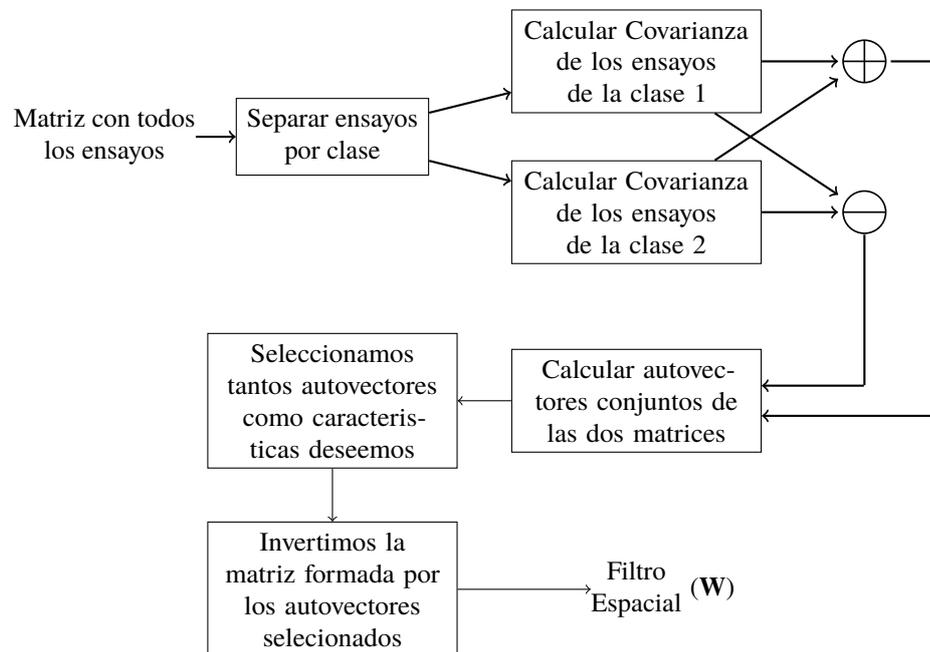


Figura 4.4 Esquema de cálculo de filtros espaciales (CSP).

4.7 Extracción de características

Una vez que se han calculado los filtros espaciales, se aplican a cada uno de los ensayos en cada una de las bandas que no es más que multiplicar ambas matrices. Tras esto, se calcula la varianza de este resultado. Obtenemos así tantas características como se hayan indicado por cada ensayo por cada banda y se almacenan en un vector.

El cálculo de la varianza se realiza porque, si recordamos la introducción del apartado 2.1, CSP era una solución cuando queríamos maximizar la diferencia entre varianzas.

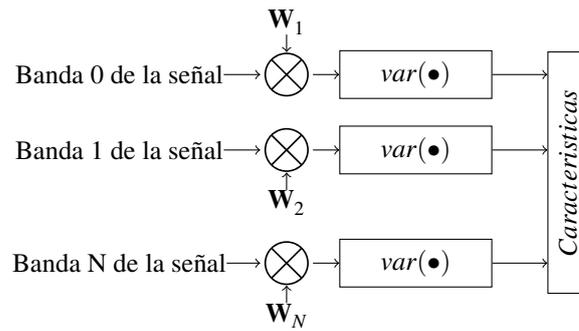


Figura 4.5 Esquema de cálculo de características.

En la solución final se tomaron cinco características por cada banda y dado que se tomaron cuatro bandas, en total se extraen 20 características por cada uno de los ensayos.

4.8 Entrenamiento del clasificador

El clasificador se crea tomando todas las características de todos los ensayos. Estas características son transformadas mediante la operación logaritmo. Esta transformación se realiza porque se ha comprobado que la distribución que siguen los puntos descritos por las características se normaliza [36] además de tener otras buenas propiedades como la convexidad [10][41][42].

El clasificador se calcula como se explica en el apartado 2.2.

4.9 Pruebas de CV

Una vez llegados a este punto, ya se dispone de todas las herramientas para realizar la clasificación. Tomamos los ensayos que fueron reservados para CV y le aplicamos el banco de filtros, el filtro espacial para extraer las características y el clasificador.

Las clases de las muestras reservadas para CV son conocidas, por lo que se contrastan los resultados obtenidos con el algoritmo con las verdaderas clases y se anota el porcentaje de acierto de nuestra solución.

4.10 Test

Tras haber realizado tantas CV como se indicase en el bloque 4.2, se escoge la solución que mejor resultados nos ha proporcionado y se sigue el mismo procedimiento del bloque 4.9, que se ilustra en la Figura 4.6.

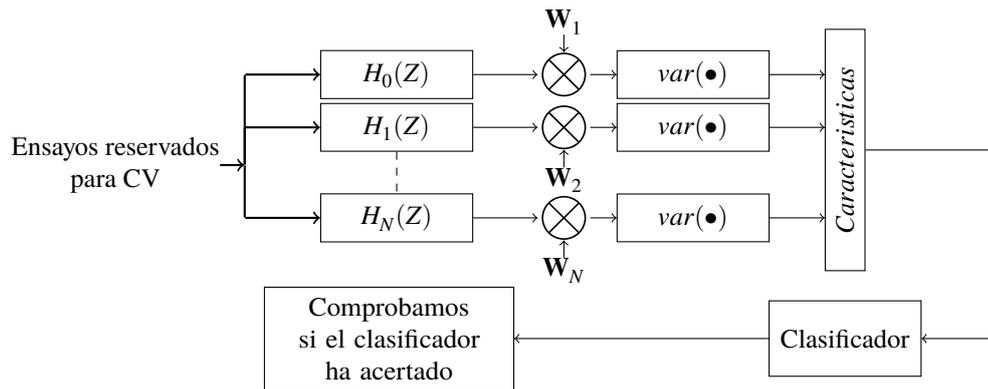


Figura 4.6 Esquema del proceso de estimación de la clase de movimiento.

4.11 Extensión a cuatro clases

Una vez desarrollado y probado el algoritmo para dos clases, nos hemos planteado si sería posible usarlo para distinguir entre las cuatro clases. Para este propósito se ha creado un sistema de votaciones simple, en el que se compara cada clase de movimiento imaginario con el resto y se escoge la clase con más votos. Lo explicamos con más detalle a continuación.

Procedemos de la misma manera que para distinguir dos clases con la diferencia de que en este caso cargamos todas las muestras de cada usuario (todas las clases) cuando antes solo cargábamos aquellas correspondientes a las clases que quisiésemos diferenciar. A continuación, se separan para CV, training y test y reservamos las de test, como hacíamos en el Bloque 2. Tras esto, se entrena un clasificador para cada par de las clases mostradas en la Tabla 4.1. Para esto, usamos tan solo las muestras de las clases que queremos diferenciar. Por ejemplo: para entrenar entre la mano izquierda y la mano derecha, se toman las muestras de la clase 1 y 2 de entre todas las que tenemos para training y CV. Hacemos lo mismo para todos los casos mostrados en la Tabla 4.1, por lo que realizamos en total seis clasificadores distintos.

Una vez que disponemos de nuestros seis clasificadores, tomamos las muestras de test que habíamos reservado y usamos los clasificadores con estas muestras para que determinen la clase a la que pertenece cada una. Tras pasar cada muestra por el clasificador, almacenamos el valor de la clase que se ha estimado, de manera que esta clase ganará un "voto". Así, tras pasar por los seis clasificadores, tendremos cuantos "votos" tiene cada clase, es decir, tendremos cuantas veces se ha estimado cada clase para una muestra. Tras el recuento de estos "votos", solamente tenemos cuatro tipos de resultados posible que se explican en la Tabla 4.3.

De los cuatro resultados posibles, en dos de ellos la clase estimada está claro que sería la clase con más votos, que serían los dos primeros casos de la Tabla 4.3. En el caso (2-2-2-0) consideramos que no es posible estimar la clase y se asigna la clase 1 (simplemente por sencillez de implementación). Por último puede darse el caso en que haya únicamente dos clases con el mismo número de votos que se trataría del caso mostrado en la última fila de la Tabla 4.3. En este caso, se asigna la clase que resultase ganadora en el clasificador que comparaba ambas clases. Esto se puede entender de forma gráfica en la Figura 4.7. Como

Tabla 4.3 Posibles resultados de la votación de clases.

Clases	Tipos de movimiento
3-2-1-0	El máximo número de votos que puede tener una clase es tres. Si una clase ha tenido tres votos significa que ha salido "ganadora" frente a las otras tres clases, por lo que no puede haber otra clase con tres votos. En todo caso, otra clase puede haber obtenido dos votos, quedando un solo voto que se repartiría entre alguna de las otras dos clases.
3-1-1-1	Otro caso es que una clase obtenga tres votos y los otros tres votos se repartan entre las otras tres clases.
2-2-2-0	Puede darse el caso de que haya tres clases que obtengan dos votos.
2-2-1-1	Por último puede haber dos clases con dos votos y otras dos clases con un solo voto.

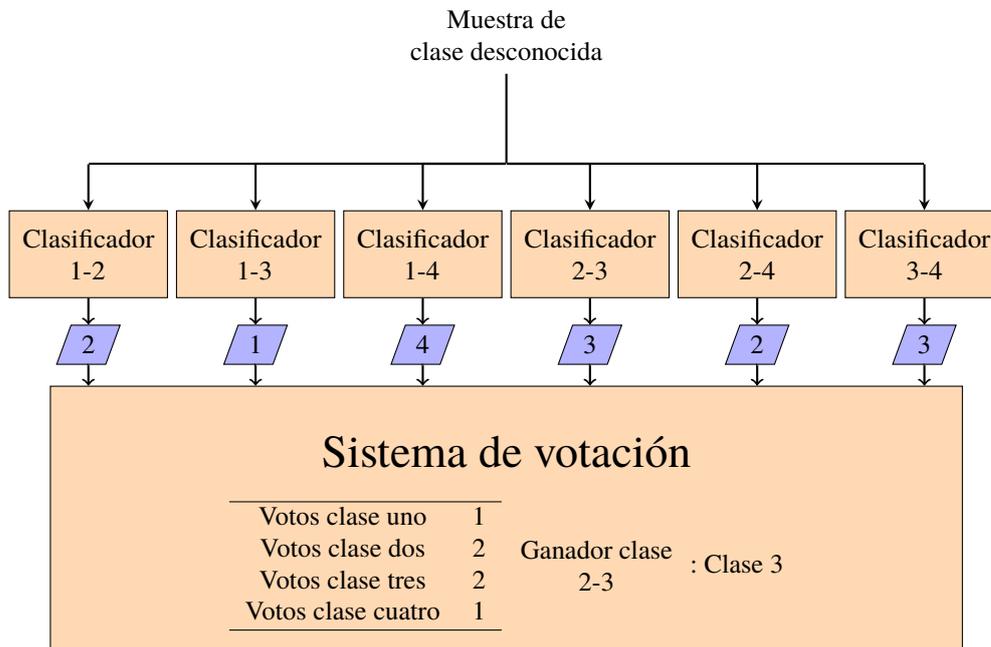


Figura 4.7 Ejemplo de ejecución del algoritmo multiclasificado cuando se da el caso en el que dos clases tienen el mismo número de votos.

podemos ver, tanto la clase 2 como la clase 3 han obtenido dos votos cada una. Para poder determinar la clase final, deberemos mirar el resultado que se ha obtenido en el clasificador de las clases 2-3, que en este caso, ha determinado que la clase estimada es la clase 3. De ahí que la clase ganadora sea la clase 3.

4.12 Conclusiones

A lo largo de este capítulo se ha explicado mediante el uso de esquemas y diagramas cada uno de los bloques de los que consta el algoritmo que se ha desarrollado, en el que se han podido distinguir las dos partes principales de extracción de características y de clasificación.

Gracias al estudio matemático que se ha realizado en el capítulo 2 de las principales técnicas que se han usado, en este capítulo nos hemos podido centrar en el funcionamiento general del programa, pudiendo verlo como una secuencia de operaciones, sin entrar en el detalle sobre ellas.

Por último se ha creado un programa basado en un sistema de votación que crea un clasificador multiclase a partir de varios clasificadores biclase como los que se han explicado durante el capítulo.

5 RSTFC: A Novel Algorithm for Spatio-Temporal Filtering and Clasificación of Single-Trial EEG

Some people feel the rain, others just get wet.

BOB MARLEY

Este capítulo está dedicado a la explicación del algoritmo propuesto en [35] que ha sido llamado Regularized Spatio-Temporal Filtering and Classification (RSTFC). Se trata de una técnica que ha sido estudiada e implementada durante el desarrollo de este trabajo con la intención de poder contrastar los resultados obtenidos mediante el algoritmo descrito en el capítulo 4.

Se ha escogido esta técnica porque consiste en una variante de CSP en la que se combina de una manera óptima la información espacial que nos proporcionan los diferentes sensores, con información temporal que puede ser vista como información frecuencial al tratarse de unidades inversas.

5.1 Metodología

El algoritmo RSTFC consiste en dos bloques: un módulo de extracción de características y otro de clasificación en el que se usa el método LDA (que ha sido estudiado con anterioridad en la sección 2.2).

En el bloque de extracción de características se establece la función objetivo mediante el uso de una matriz aumentada, que consiste en la matriz original de cada ensayo y la repetición de varias copias de esta misma matriz retrasada.

Para conseguir esto, en primer lugar se considera un filtro espacial:

$$s = [s_1, s_2, \dots, s_C]^\top \quad (5.1)$$

A su vez tendríamos el filtro temporal correspondiente a cada canal que se define como:

$$a_c = [a_{c1}, a_{c2}, \dots, a_{cN}]^\top \quad (5.2)$$

Estos filtros pueden representarse como funciones donde $f_s : \mathbb{R}^{C \times T} \rightarrow \mathbb{R}^{1 \times T}$ y $f_{a_c} : \mathbb{R}^{1 \times T} \rightarrow \mathbb{R}^{1 \times T}$ entonces:

$$X \rightarrow f_s(X) = \sum_{c=1}^C s_c X_c \quad (5.3)$$

$$X_c \rightarrow f_{a_c}(X_c) = \sum_{n=1}^N a_c X_c^{n-1} \quad (5.4)$$

Donde X_c^{n-1} denota la copia de los datos de un sensor de un ensayo, retrasada $(n-1)$ muestras temporales. Si se toman todos los canales podemos expresar la siguiente matriz:

$$\bar{X} = f_a(X) = \begin{pmatrix} f_{a_1}(X_1) \\ f_{a_2}(X_2) \\ \vdots \\ f_{a_c}(X_c) \end{pmatrix} \quad (5.5)$$

La operación combinada de realizar el filtrado temporal y espacial se puede expresar de mediante la función $f_w : \mathbb{R}^{C \times T} \rightarrow \mathbb{R}^{1 \times T}$ y de la siguiente forma:

$$X_c \rightarrow f_w(X_c) = f_s \cdot f_a(X) = f_s(\bar{X}) \\ \sum_{c=1}^C s_c \sum_{n=1}^N a_c X_c^{n-1} = \bar{w} \bar{X} \quad (5.6)$$

Donde $\bar{w} = [s_1 a_{11}, s_2 a_{12}, \dots, s_C a_{1C}, \dots, s_1 a_{N1}, s_2 a_{N2}, \dots, s_C a_{NC}]$ se trata de un filtro espacio-temporal y

$$\tilde{X} = f_a(X) = \begin{pmatrix} X^{(1)} \\ X^{(2)} \\ \vdots \\ X^{(N-1)} \end{pmatrix} \quad (5.7)$$

es una matriz de datos aumentada que añade señales retrasadas como si se tratasen de nuevos canales.

De esta manera y conforme lo visto en la ecuación (5.6) se puede ver la transformación espacio-temporal como una sola, con el vector \bar{w} representando los dos filtros. Se puede optimizar el vector \bar{w} como si se tratase de un filtrado espacial usando CSP tal y como

vimos en el apartado 2.1, usando en vez de las matrices de datos simples, las matrices aumentadas \tilde{X} .

Es decir, se optimiza el vector \bar{w} siguiendo el siguiente criterio:

$$\bar{w} = \arg \max_w \frac{w^\top \tilde{R}_1 w}{w^\top \tilde{R}_2 w} \quad (5.8)$$

Donde \tilde{R}_1 y \tilde{R}_2 denotan las covarianzas de las matrices aumentadas. Realmente el algoritmo RSTFC se basa en calcular la solución de Regularized Common Spatial Patterns (RCSP), ya que uno de los principales problemas de CSP es que tiende a sobreajustarse [25] [38], y RCSP soluciona este problema. Como hemos visto el algoritmo RSTFC consiste en añadir repeticiones retrasadas de las señales como si fueran nuevos canales, por lo que se va multiplicando el número de dimensiones con las que trabaja CSP. Esto repercute en que es mucho más probable cometer overfitting.

RCSP consiste unicamente en añadir un parámetro de regularización en el denominador de la ecuación (5.8), quedando la ecuación de la siguiente forma:

$$\bar{w} = \arg \max_w \frac{w^\top \tilde{R}_1 w}{w^\top (\tilde{R}_2 + \rho I) w} \quad (5.9)$$

Pero además, debido a la dimensión del filtro w , en este caso nos interesa tener tantos ceros como sea posible por lo que se regulariza usando la ecuación:

$$\bar{w} = \arg \max_w \frac{w^\top \tilde{R}_1 w}{w^\top \tilde{R}_2 w + \rho \|w\|} \quad (5.10)$$

Donde $\|w\|$ es la norma de w . En cualquier caso, la ecuación (5.10) ya no es un cociente de Rayleigh y no puede ser optimizado mediante el uso de autovectores, por lo que se ha de recurrir a otras técnicas. Los autores del algoritmo RSTFC proponen buscar el valor de \bar{w} mediante la optimización de la ecuación (5.11) que resulta equivalente a hacerlo mediante la ecuación (5.9):

$$\bar{w} = \arg \min_w w^\top \tilde{R}_2 w + \frac{1}{w^\top \tilde{R}_1 w} + \rho \|w\| \quad (5.11)$$

Para encontrar el valor óptimo de la ecuación anterior se propone usar la función `fminunc` de MATLAB.

Finalmente, el módulo de extracción de características se puede resumir como:

1. Se construye la matriz aumentada de datos para cada ensayo como se especifica en (5.7).
2. Se calculan las medias de las matrices de covarianza para cada clase \tilde{R}_1, \tilde{R}_2 .
3. Se optimiza el filtro \bar{w} mediante (5.11).

5.2 Conclusiones

En este capítulo se ha podido estudiar una técnica novedosa que fue publicada en 2015, ayudándonos a acercarnos al estado del arte en los sistemas BCI.

El algoritmo estudiado en este capítulo usa CSP para aprovechar la diversidad espacial que nos proporcionan los sensores a la vez que aprovecha diversidad temporal incorporando muestras retrasadas de las señales como si fuesen nuevos canales. Esto permite obtener eficientemente un filtro espacio-temporal óptimo bajo un único criterio.

Además en este capítulo hemos podido dar un breve repaso al método RCSP que resuelve uno de los principales problemas de CSP, el overfitting.

6 Resultados

Success is a result, not a goal.

GUSTAVE FLAUBERT

En este capítulo se exponen los resultados que se han obtenido tras probar el algoritmo desarrollado con los distintos conjuntos de muestras. Para exponer dichos resultados, además de usar el software Matlab con el cual hemos obtenido las gráficas que se incluyen aquí, se han usado las librerías de EEGLAB¹ que nos han ayudado a representar las figuras que simulan las cabezas de los pacientes y los pesos de los filtros espaciales.

Como se comentó en el capítulo 3, tenemos cuatro clases de movimiento y 288 ensayos por cada sesión. Además tenemos dos sesiones por cada usuario, es decir, que disponemos de un total de 5184 ensayos para hacer experimentos. En los casos en los que probemos el algoritmo para dos clases solamente, podremos usar la mitad de los ensayos, es decir, cogeremos solo los de las clases de movimiento que estemos evaluando.

Para poder obtener los mejores resultados posibles, es necesario optimizar los parámetros que el código usa. Puesto que podemos hacer hasta seis experimentos distintos (los mostrados en la Tabla 4.1), tendremos que optimizar cada uno de los seis problemas por separado. Los parámetros que se han optimizado están explicados en la Tabla 6.1 que se muestra a continuación:

Debido a la complejidad que tiene optimizar todos estos parámetros a la vez para cada caso, se ha optado por optimizarlos todos una vez, concretamente para el caso en que estemos tratando de distinguir entre los pensamientos de la mano derecha y de la mano izquierda. Para el resto de casos lo que hemos hecho es buscar solamente el valor óptimo de los parámetros de orden de los filtros y número de características por banda. Los parámetros que se han usado finalmente se exponen en la Tabla 6.2. Aun así, algunos de estos parámetros se variarán durante las pruebas para mostrar el efecto de los mismos sobre los resultados.

¹ <https://sccn.ucsd.edu/eeglab/>

Tabla 6.1 Parámetros optimizados.

Parámetro	Descripción
Orden de los filtros paso banda	El orden de los filtros se ha variado desde 3 hasta 10.
Bandas de frecuencia	Se han probado las bandas de frecuencia que se especifican en Tabla 4.2.
Inicio de ensayo	Este parámetro hace referencia a cuantos segundos después del comienzo del ensayo comienza la señal de interés. Se ha variado desde dos segundos hasta tres.
Duración	Se ha tratado de optimizar la duración del trozo de señal que tomamos. Se ha variado desde 2.5s hasta 3.5s.
Número de características	El número de características que se extrae de cada banda se ha variado desde 1 hasta 22.

Tabla 6.2 Parámetros usados.

Parámetro	Valor
Orden de los filtros paso banda	Variable
Bandas de frecuencia	Cuatro bandas de frecuencias límite (4-8-14-30-40)
Inicio de ensayo	2.3 segundos
Duración	3 segundos
Número de características	Variable
Porcentaje de ensayos usados para entrenamientos	80%
Porcentaje de ensayos usados para CV	10%
Porcentaje de ensayos usados para test	10%

Antes de exponer los resultados, creemos importante hacer una breve explicación acerca de las figuras que simulan las cabezas de los sujetos que se van a representar en este capítulo. En las figuras se ha representado una cabeza con varios puntos dentro, que representan los sensores del casco de EEG. La figura se ha coloreado para que pueda ser interpretada fácilmente. En ella se representan los pesos de los filtros espaciales calculados con CSP. Por lo que hay que destacar que realmente la información útil está solamente en los puntos que representan a los sensores, ya que los valores del resto de la superficie han sido calculados mediante interpolación. Por otro lado cada figura representa un filtro que extrae una característica. En la mayoría de los casos, se han extraído más de dos características, pero en este apartado solo representaremos los filtros que maximizan y minimizan el criterio de CSP en mayor medida. Por último, es importante destacar que todos estos diagramas se

corresponden con los filtros calculados para el sujeto 7, que ha sido elegido porque ha sido uno de los sujetos con los que mejores resultados se han obtenido y hemos considerado que representaría muy bien "como se supone que se debe de pensar en cada movimiento imaginario".

6.1 Mano izquierda - Mano derecha

En este apartado vamos a abordar el problema que consiste en estimar si los movimientos imaginarios de los ensayos corresponden con la mano derecha o con la mano izquierda.

En este problema se han usado cinco características y filtros paso de banda de orden seis. Para poder comparar la mejora que supone usar distintas bandas de frecuencia, vamos a repetir este experimento para algunos de los casos mostrados en la Tabla 4.2, en concreto: usando una sola banda, usando dos bandas y usando cuatro bandas.

6.1.1 Usando una sola banda (Alfa y Beta a la vez)

En este caso se trabaja con una sola banda de frecuencia, que abarca desde los 8Hz hasta los 30Hz. Esto se corresponde con las bandas de frecuencia Alfa y Beta de las ondas del cerebro, que es el rango de frecuencias con el que se suele trabajar en BCI [33], [39].

En este caso se ha estimado la clase correcta en un 80.16% de las muestras reservadas para test. En la Figura 6.1 vemos los resultados obtenidos en este caso, que como veremos más adelante en el apartado 6.1.2, distan mucho de los resultados obtenidos cuando se usan las bandas por separado. Además, en la figura Figura 6.1 también se pueden ver los aciertos por cada uno de los usuarios.

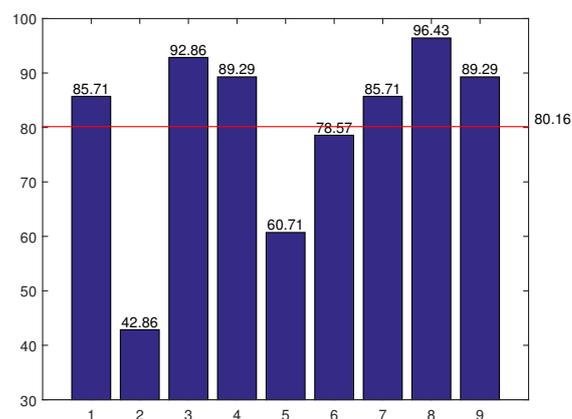


Figura 6.1 Mano izquierda - Mano derecha: Resultados por sujeto usando las bandas Alfa y Beta como una sola.

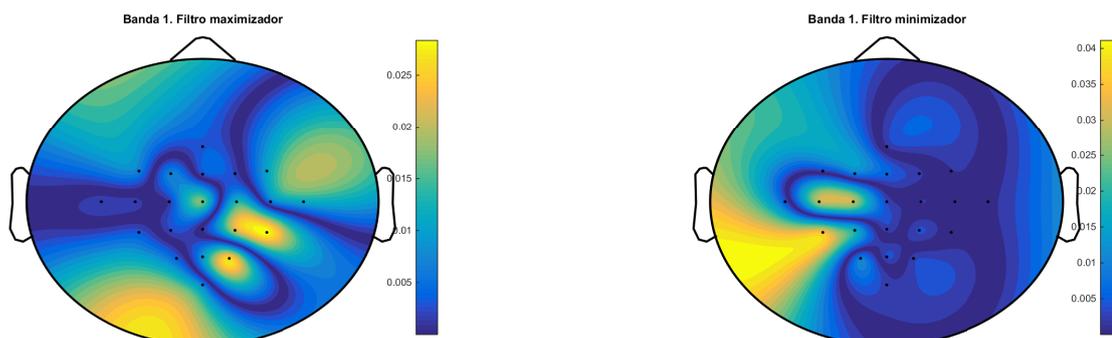


Figura 6.2 Mano izquierda - Mano derecha: Filtro maximizador y minimizador de las bandas Alfa y Beta a la vez. Donde se puede apreciar las zonas mas activas de la cabeza durante el pensamiento de estos dos movimientos.

6.1.2 Usando dos bandas (Alfa y Beta por separado)

Si en vez de usar una banda de frecuencia como en el caso anterior tomamos dos bandas, una desde 8Hz hasta 14Hz y otra desde 14Hz hasta 30Hz, vemos que el resultado mejora notablemente, obteniendo un 86.11 % el acierto de las clases estimadas.

Los filtros espaciales que se han calculado para cada banda específicamente se han representado en Figura 6.6 y Figura 6.7. Como podemos observar en la Figura 6.3, Los resultados han mejorado considerablemente y no ha habido ningún usuario al que el uso de dos bandas le haya perjudicado.

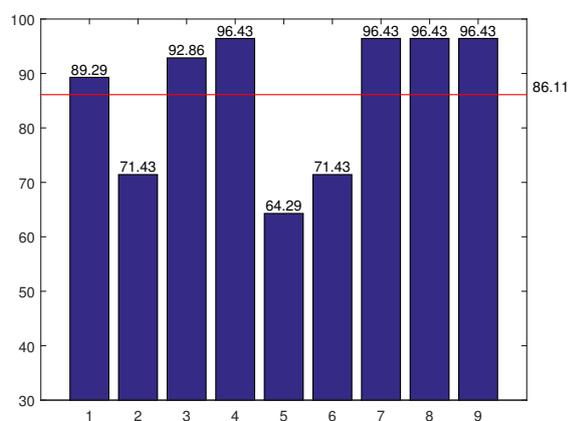


Figura 6.3 Mano izquierda - Mano derecha: Resultados por sujeto usando las bandas Alfa y Beta por separado.

6.1.3 Usando cuatro bandas

Si usamos las cuatro bandas de frecuencia el resultado mejora aun más. En la Figura 6.4 se aprecia que hay dos usuarios que alcanzan una tasa de estimación correcta del 100% de las muestras de test, consiguiendo en media un porcentaje de estimación correcta del 87.30%. Sin embargo, podemos apreciar que en varios sujetos se ha disminuido su tasa de estimación correcta, en concreto dos sujetos. Esto nos ha hecho plantearnos si merece la pena incluir las 4 bandas de frecuencia, pero al realizar varias pruebas con varias inicializaciones distintas, se ha comprobado que siempre conseguimos mejores resultados en este caso. Como muestra de ello, en la Tabla 6.3 se exponen los distintos resultados de las distintas pruebas que hemos hecho. Se han realizado tres inicializaciones distintas (tomando diferentes muestras para entrenamiento y test en cada caso, pero tomándolas iguales para las pruebas en las que se cambia el número de banda) y hemos calculado la media en porcentaje del número de muestras estimadas correctamente. Como se puede ver en dicha tabla, siempre conseguimos un porcentaje mayor en media para el caso de las cuatro bandas.

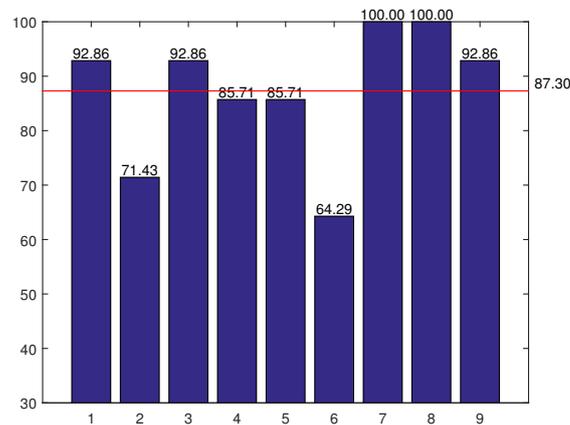


Figura 6.4 Mano izquierda - Mano derecha: Resultados por sujeto usando cuatro bandas.

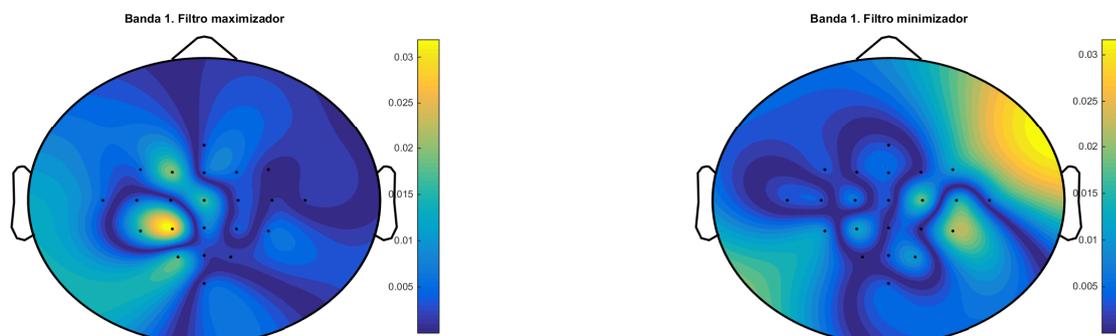


Figura 6.5 Mano izquierda - Mano derecha: Filtro maximizador y minimizador de la banda Theta.

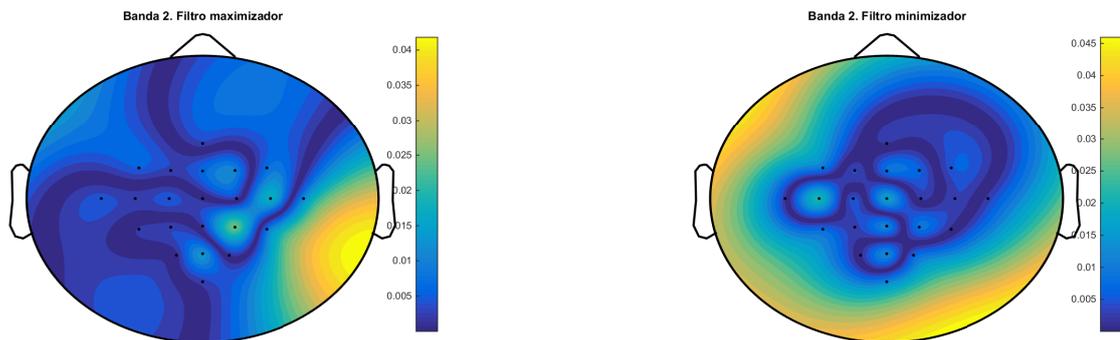


Figura 6.6 Mano izquierda - Mano derecha: Filtro maximizador y minimizador de la banda Alfa.

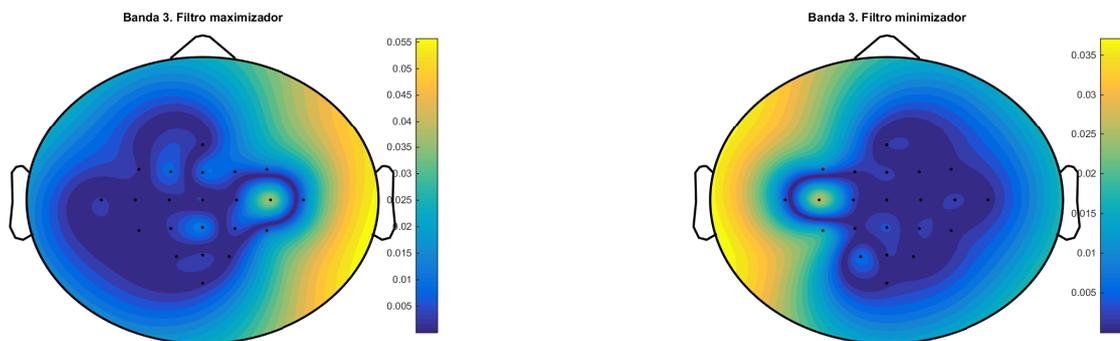


Figura 6.7 Mano izquierda - Mano derecha: Filtro maximizador y minimizador de la banda Beta.

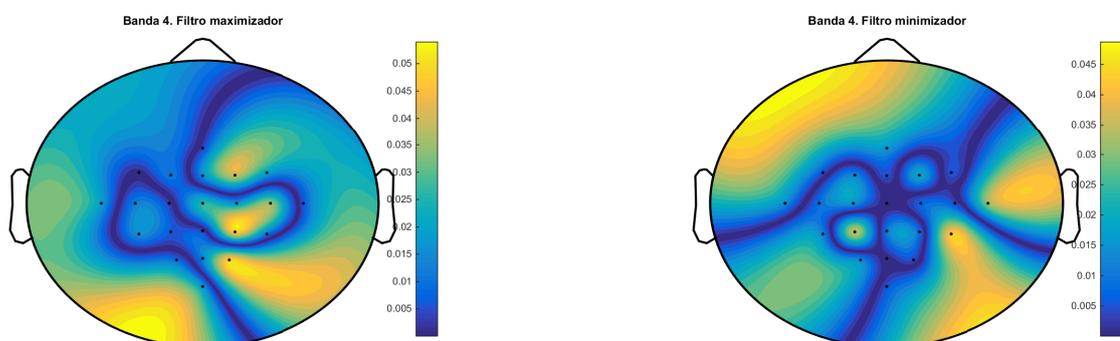


Figura 6.8 Mano izquierda - Mano derecha: Filtro maximizador y minimizador de la banda Gamma.

Tabla 6.3 Diferentes inicializaciones (usando diferentes semillas) con distinto numero de bandas de frecuencia.

Nº Bandas	Inicialización 1			Inicialización 2			Inicialización 3			
	1	2	4	1	2	4	1	2	4	
Usuarios	1	92.86	96.43	92.86	89.29	92.86	100	96.43	89.29	85.71
	2	57.14	64.29	57.14	57.14	53.57	57.14	60.71	75	75
	3	96.43	100	100	89.29	92.86	96.43	89.29	96.43	96.43
	4	75	71.43	67.86	85.71	78.57	60.71	82.14	78.57	75
	5	42.86	53.57	78.57	53.57	57.14	89.29	39.29	57.14	75
	6	57.14	75	64.29	67.86	78.57	78.57	75	64.29	60.71
	7	92.86	92.86	100	78.57	78.57	100	78.57	85.71	96.43
	8	92.86	89.29	92.86	96.43	96.43	96.43	100	100	100
	9	92.86	92.86	89.29	96.43	96.43	96.43	92.86	89.29	89.29
Media	77.78	81.75	82.54	79.37	80.56	86.11	79.37	81.75	83.73	

6.2 Mano izquierda - Ambos pies

Este problema consiste en estimar si las muestras de los ensayos son asociadas a movimientos imaginarios de la mano izquierda o de ambos pies a la vez, es decir, intentaremos estimar si es de clase 1 o 3.

Para este caso se vuelven a usar cinco características pero el filtro que se usa es de orden tres. Si solamente usamos una banda de frecuencia, la banda que va desde los 8Hz hasta los 30Hz, se obtiene un porcentaje de aciertos de 84,52%. En la Figura 6.9 se puede ver el porcentaje de aciertos por sujeto.

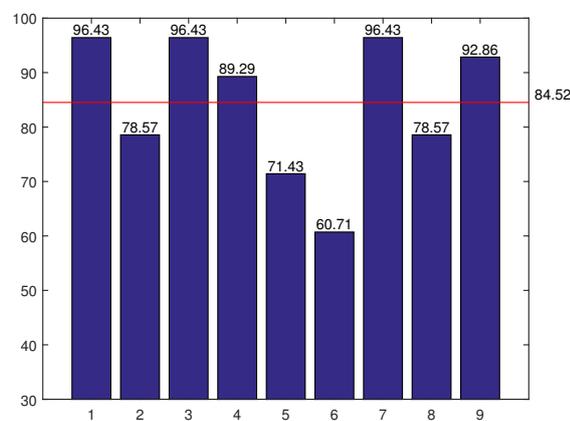


Figura 6.9 Mano izquierda - Ambos pies: Resultados por sujeto usando las bandas Alfa y Beta a la vez.

Los pesos de los filtros espaciales se han representado en la Figura 6.10.

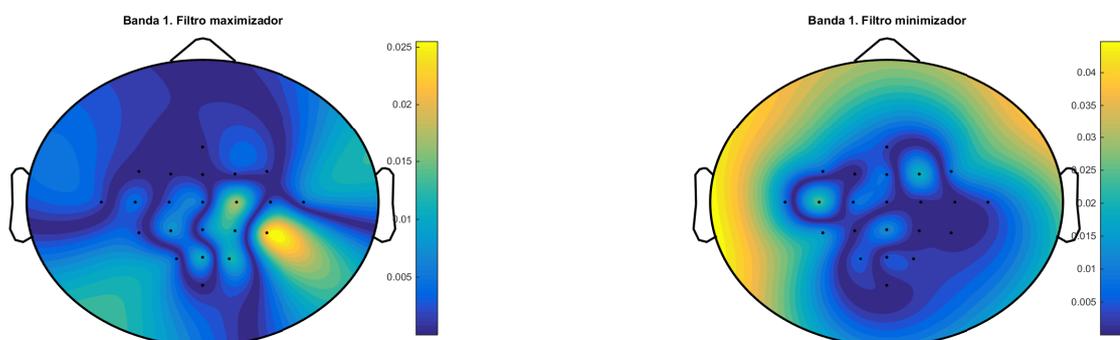


Figura 6.10 Mano izquierda - Ambos pies: Filtro maximizador y minimizador de las bandas Alfa y Beta a la vez.

Por otro lado, si usamos las cuatro bandas se obtiene un porcentaje de acierto de 91.87%. En la Figura 6.11 se han representado los porcentaje de aciertos por sujeto.

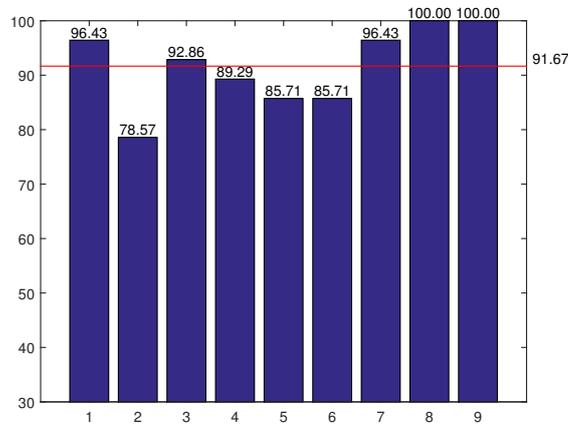


Figura 6.11 Mano izquierda - Ambos pies: Resultados por sujeto usando cuatro bandas (Theta, Alfa, Beta y Gamma).

6.3 Mano izquierda - Lengua

Este problema consiste en estimar si las muestras de los ensayos son asociadas a movimientos imaginarios de la mano izquierda o de la lengua, es decir, intentaremos estimar si es de clase 1 o 4.

Para este caso se usan cuatro características y el filtro que se usa es de orden tres. Si solamente usamos una banda de frecuencia, la banda que va desde los 8Hz hasta los 30Hz, se obtiene un porcentaje de aciertos de 81.75%. En la Figura 6.12 se puede ver el porcentaje de aciertos por sujeto.

Los pesos de los filtros espaciales se han representado en la Figura 6.13.

Por otro lado, si usamos las cuatro bandas comprobamos que los resultados mejoran notablemente, obteniéndose un porcentaje de acierto del 87.30%. En la Figura 6.14 se han representado los porcentaje de aciertos por sujeto.

6.4 Mano derecha - Ambos pies

Este problema consiste en estimar si las muestras de los ensayos son asociadas a movimientos imaginarios de la mano derecha o ambos pies, es decir, intentaremos estimar si es de clase 2 o 3.

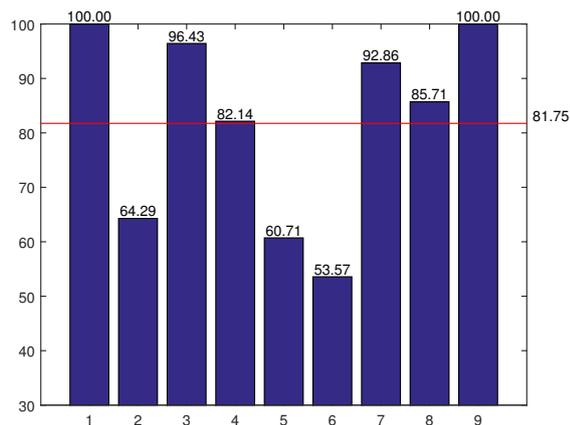


Figura 6.12 Mano izquierda - Lengua: Resultados por sujeto usando las bandas Alfa y Beta a la vez.

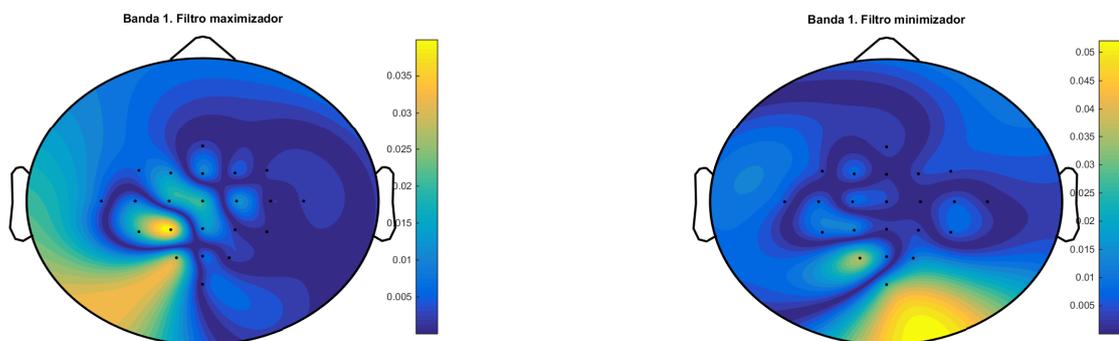


Figura 6.13 Mano izquierda - Lengua: Filtro maximizador y minimizador de las bandas Alfa y Beta a la vez.

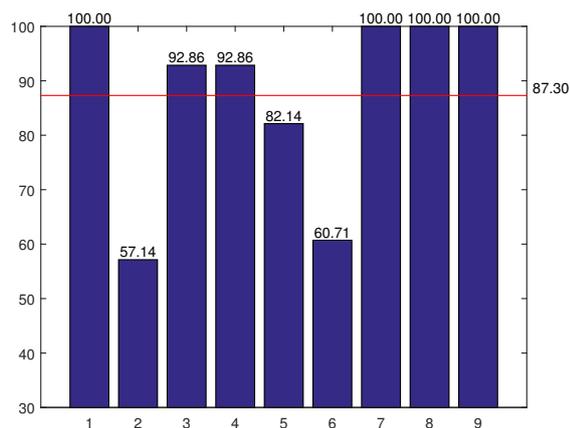


Figura 6.14 Mano izquierda - Lengua: Resultados por sujeto usando cuatro bandas (Theta, Alfa, Beta y Gamma).

Para este caso se usan cuatro características y el filtro que se usa es de orden siete. Si solamente usamos una banda de frecuencia, la banda que va desde los 8Hz hasta los 30Hz,

se obtiene un porcentaje de aciertos de 81.35%. En la Figura 6.15 se puede ver el porcentaje de aciertos por sujeto.

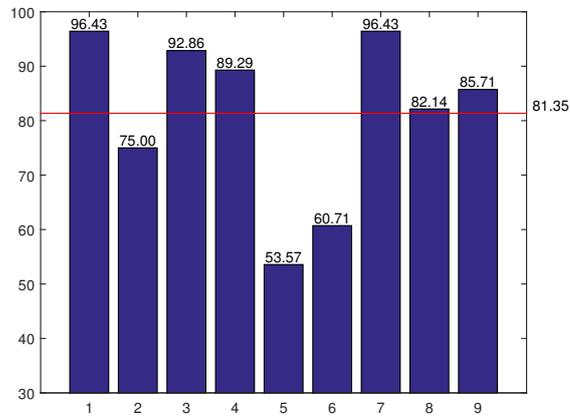


Figura 6.15 Mano derecha - Ambos pies: Resultados por sujeto usando las bandas Alfa y Beta a la vez.

Los pesos de los filtros espaciales se han representado en la Figura 6.13.

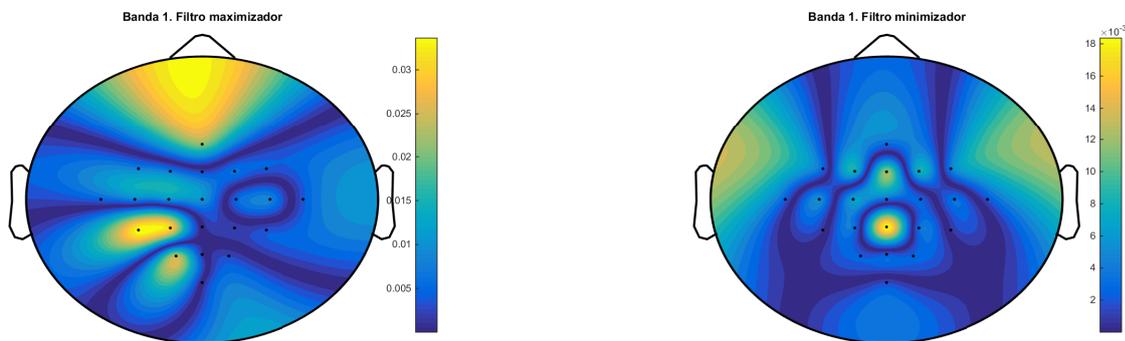


Figura 6.16 Mano derecha - Ambos pies: Filtro maximizador y minimizador de las bandas Alfa y Beta a la vez.

Si usamos las cuatro bandas comprobamos que los resultados mejoran notablemente, obteniéndose un porcentaje de acierto de 86.11%. En la Figura 6.17 se han representado los porcentaje de aciertos por sujeto.

6.5 Mano derecha - Lengua

Este problema consiste en estimar si las muestras de los ensayos son asociadas a movimientos imaginarios de la mano derecha o de la lengua, es decir, intentaremos estimar si es de

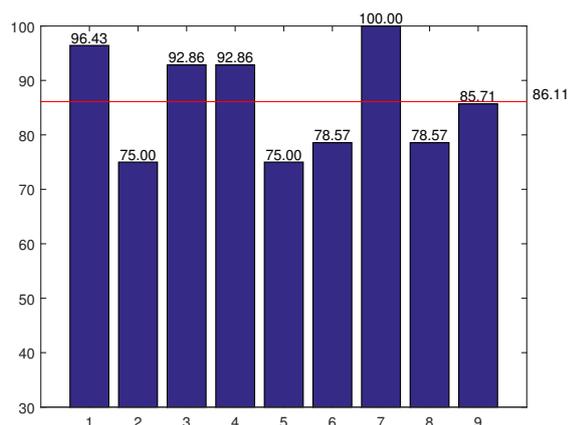


Figura 6.17 Mano derecha - Ambos pies: Resultados por usuario usando cuatro bandas (Theta, Alfa, Beta y Gamma).

clase 2 o 4.

Para este caso se usan seis características y el filtro que se usa es de orden 4. Si solamente usamos una banda de frecuencia, la banda que va desde los 8Hz hasta los 30Hz, se obtiene un porcentaje de aciertos de 84.92%. En la Figura 6.18 se puede ver el porcentaje de aciertos por sujeto.

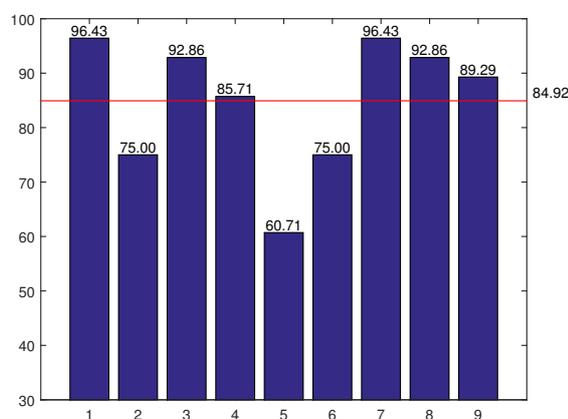


Figura 6.18 Mano derecha - Lengua: Resultados por sujeto usando las bandas Alfa y Beta a la vez.

Los pesos de los filtros espaciales se han representado en la Figura 6.13.

Si usamos las cuatro bandas comprobamos que los resultados mejoran notablemente, obteniéndose un porcentaje de acierto de 87.70%. En la Figura 6.20 se han representado los porcentaje de aciertos por sujeto.

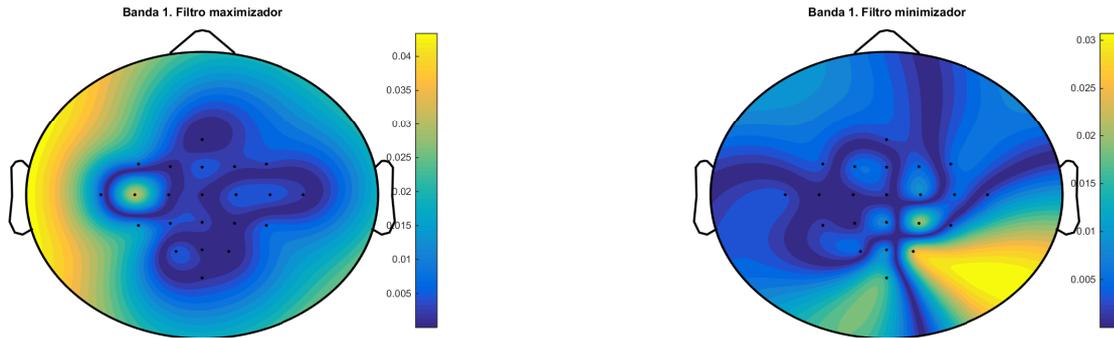


Figura 6.19 Mano derecha - Lengua: Filtro maximizador y minimizador de las bandas Alfa y Beta a la vez.

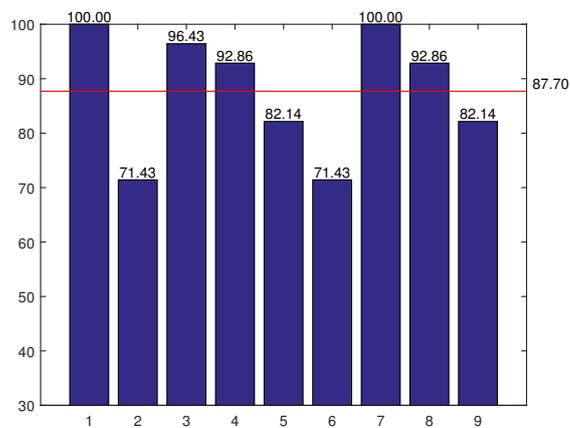


Figura 6.20 Mano derecha - Lengua: Resultados por sujeto usando cuatro bandas (Theta, Alfa, Beta y Gamma).

6.6 Ambos pies - Lengua

Este problema consiste en estimar si las muestras de los ensayos son asociadas a movimientos imaginarios de ambos pies o de la lengua, es decir, intentaremos estimar si es de clase 3 o 4.

Para este caso se usan seis características y el filtro que se usa es de orden 4. Si solamente usamos una banda de frecuencia, la banda que va desde los 8Hz hasta los 30Hz, se obtiene un porcentaje de aciertos de 80.16%. En la Figura 6.21 se puede ver el porcentaje de aciertos por sujeto.

Los pesos de los filtros espaciales se han representado en la Figura 6.13.

Si usamos las cuatro bandas comprobamos que los resultados mejoran notablemente, obteniéndose un porcentaje de acierto de 84.92%. En la Figura 6.23 se han representado los porcentaje de aciertos por sujeto.

En este caso se han obtenido resultados un poco mas bajos que en el resto de casos. Lo

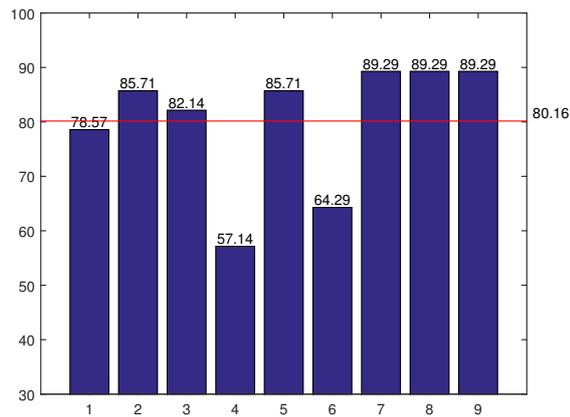


Figura 6.21 Ambos pies - Lengua: Resultados por sujeto usando las bandas Alfa y Beta a la vez.

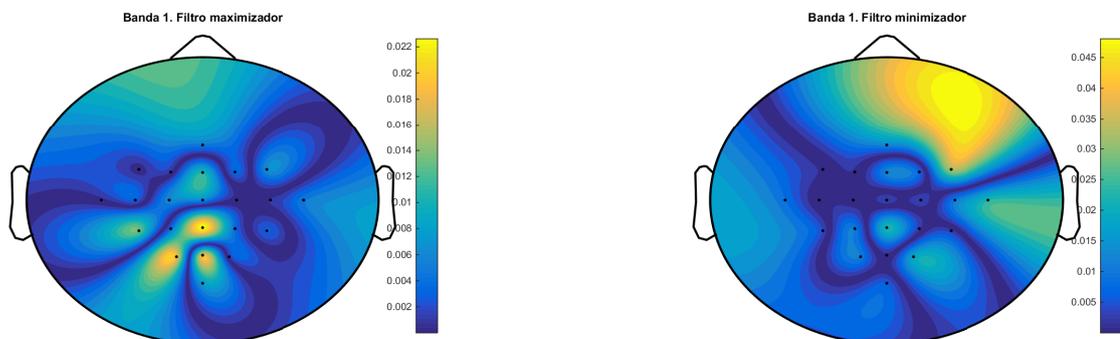


Figura 6.22 Ambos pies - Lengua: Filtro maximizador y minimizador de las bandas Alfa y Beta a la vez.

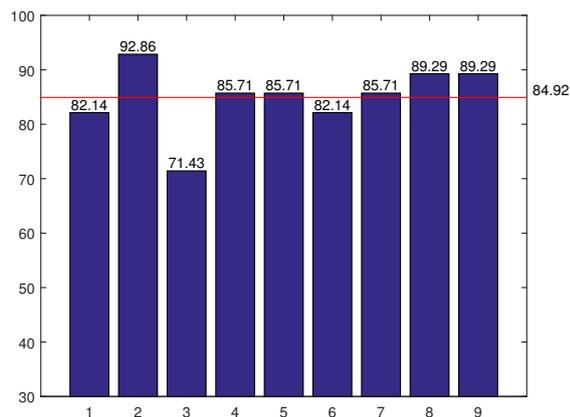


Figura 6.23 Ambos pies - Lengua: Resultados por sujeto usando cuatro bandas (Theta, Alfa, Beta y Gamma).

cual puede resultar curioso porque si observamos la Figura 1.5 vemos estas dos clases movimiento imaginario se asocian a dos zonas del cerebro muy separadas.

6.7 Multiclase

El último experimento que se ha realizado ha sido la prueba del algoritmo que se explica en el apartado 4.11. Para probarlo se ha usado cada uno de los parámetros que se describen en la Tabla 6.2 y los valores que se han tomado para los parámetros que aparecen como variables en la tabla son los que se han ido comentando a lo largo del capítulo para cada par de clases.

Finalmente encontramos un porcentaje de estimación de muestras correctas del 72.7%.

En la Figura 6.24 se ha representado el porcentaje de aciertos de cada sujeto en la prueba multiclase.

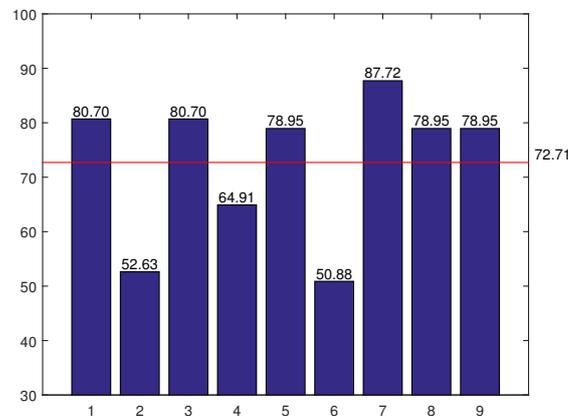


Figura 6.24 Multiclase: porcentaje de aciertos por sujeto.

En la Figura 6.25 se ha representado el porcentaje de muestras que han obtenido las diferentes combinaciones de votos que se explican en la Tabla 4.3.

Además, en este caso, al tratarse de un problema más complejo hemos realizado una serie de cálculos adicionales. Se ha representado la sensibilidad² y la precisión³ de cada una de las clases para cada uno de los sujetos. Se considera que esta información podría resultar útil para saber qué clases de movimiento debería de practicar más cada usuario, o usar esta información para estimar la clase de las muestras en las que se han obtenido el mismo número de votos en tres clases (2-2-2-0).

Por último en la tabla Tabla 6.4 se muestra la matriz de confusión de las clases.

² La sensibilidad se define como: $Sensibilidad = \frac{VP}{VP+FN}$. Donde VP son verdaderos positivos y FN son falsos negativos. Nos da la medida del porcentaje de las muestras de cada clase que se han acertado.

³ La precisión se define como: $Precision = \frac{VP}{VP+FP}$. Donde FP son falsos positivos. Es decir, nos da una medida de entre todas las veces que se ha clasificado con una clase determinada cuantas veces se ha acertado.

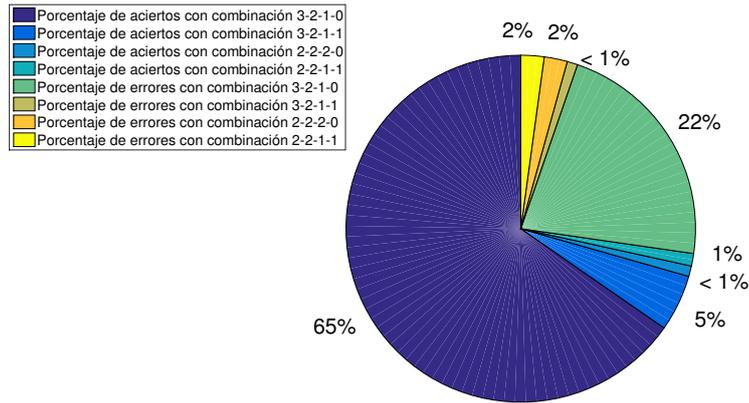


Figura 6.25 Diagrama de porcentaje de clases de votos en algoritmo multiclase.

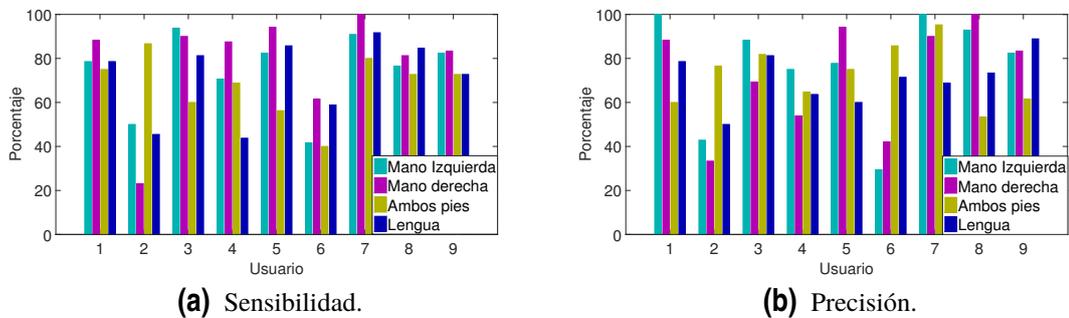


Figura 6.26 Sensibilidad y Precisión por clase y sujeto en el caso multiclase..

Tabla 6.4 Matriz de confusión del caso multiclase. En las filas se representan las clases reales, mientras que en las columnas las clases asignadas.

Real \ Estim	Estim				Total	Acierto	Error
	1	2	3	4			
1	14	1	1	1	17	82.35 %	17.65 %
2	0	15	3	0	18	83.33 %	16.67 %
3	2	1	8	0	17	72.73 %	27.27 %
4	1	1	1	8	17	72.73 %	27.27 %

6.8 Regularized Spatio-Temporal Filtering and Classification

En esta sección se presentan los resultados que se han obtenido para cada uno de los pares de movimientos imaginarios mediante el uso del algoritmo RSTFC que se ha explicado en el capítulo 5.

Para hacer estas pruebas, se ha tomado como parámetro de regularización $\rho = 0.1$ y el número de copias usadas para crear la matriz ampliada se ido variando entre 10 y 19, escogiéndose en cada caso el número de veces que mejor resultado ha dado en CV.

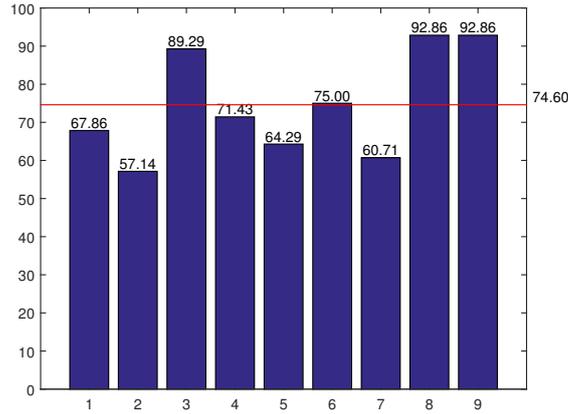


Figura 6.27 RSTFC: Mano izquierda, mano derecha..

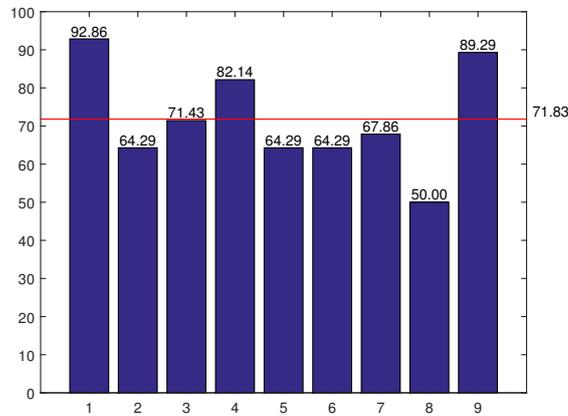


Figura 6.28 RSTFC: Mano izquierda, pies..

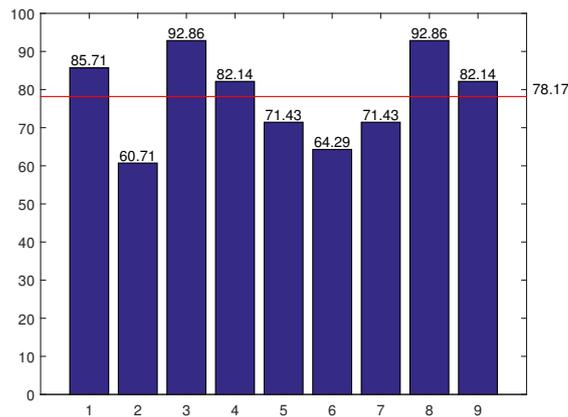


Figura 6.29 RSTFC: Mano izquierda, lengua..

Como se puede observar, los resultados obtenidos usando el algoritmo RSTFC distan mucho de los resultados obtenidos usando del algoritmo desarrollado en este trabajo, así como de los resultados que los autores expresan en el artículo en el que fue publicado.

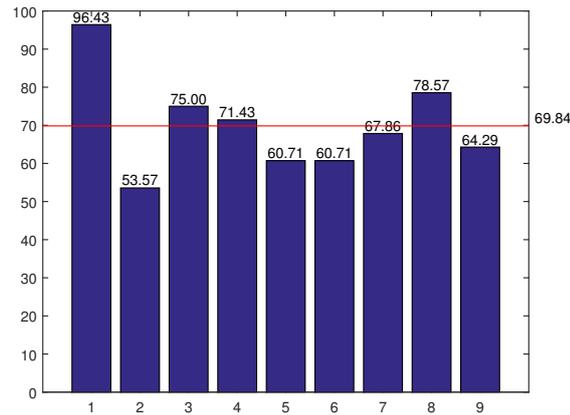


Figura 6.30 RSTFC: Mano derecha, pies..

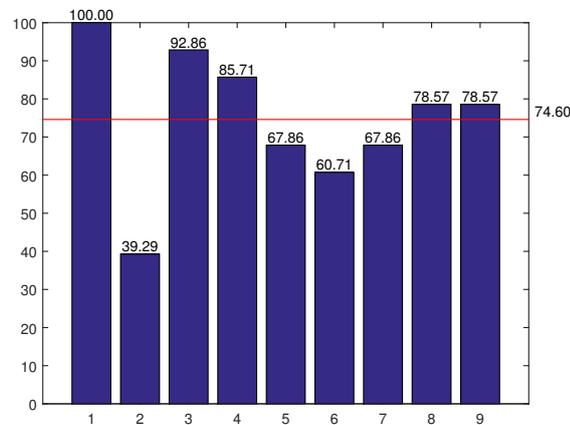


Figura 6.31 RSTFC: Mano derecha, lengua..

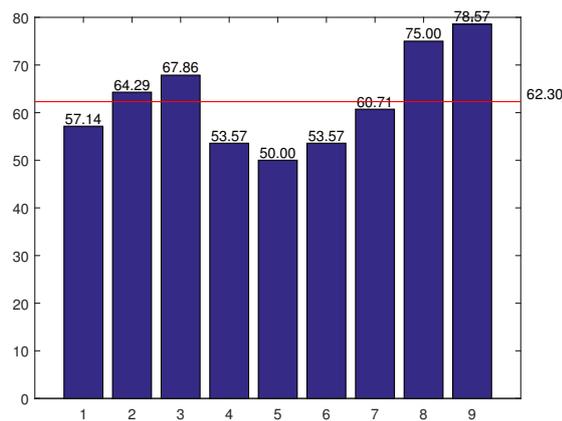


Figura 6.32 RSTFC: Pies, lengua..

Esta diferencia se achaca a que el algoritmo ha sido reprogramado por el autor de este trabajo a partir del artículo, de manera que no se han usado los códigos originales. Esto se ha debido a que existen varios parámetros desconocidos en la implementación que

los autores han realizado, destacando por ejemplo, que en el artículo no se especifica el parámetro de regularización ρ que se ha usado o cuantas copias retrasadas de las señales se usan para formar la matriz aumentada.

6.9 Conclusiones

En este capítulo se han expuesto los resultados que se han obtenido usando el algoritmo que se ha desarrollado en este trabajo junto con los obtenidos usando el algoritmo RSTFC. Se han presentado gráficas de resultados y se han hecho breves comentarios sobre lo que se ha obtenido en cada caso.

En el capítulo siguiente se detallan algunas de las conclusiones que se han tomado a partir de los resultados obtenidos.

7 Análisis de los resultados

If you can't explain it to a six year old, you don't understand it yourself.

ALBERT EINSTEIN

Tras los resultados expuestos en el capítulo 6, se ha considerado necesario analizar los mismos con detenimiento, ya que se pueden extraer múltiples conclusiones de todas las pruebas que se han realizado. A lo largo de este capítulo se comentan las conclusiones más importantes extraídas de los resultados así como la posible justificación de los resultados que se han obtenido.

7.1 Descomposición en bandas de frecuencia

La idea original de dividir la señal en distintas bandas de frecuencia surgió a partir del estudio de las ondas del cerebro. En el apartado 1.3.1 se vio que cada una de las bandas de frecuencia del cerebro actúa por separado y se asocian a estados mentales diferentes. Es más, cada banda del cerebro aparece en zonas distintas del mismo y se considera que tienen orígenes diferentes.

Aunque puede haber cierta relación entre las bandas() por ejemplo en la banda Beta se cueñan muchos armónicos de la onda μ [29]), tiene sentido analizarlas por separado, ya que cada una de las bandas nos puede proporcionar nueva información que puede o no, ser de utilidad.

Se ha podido comprobar que la mayor parte de la información útil para estimar el movimiento imaginario que está realizando un sujeto se encuentra en la banda Alfa, ya que como se ha mencionado en el apartado 1.4.1, las frecuencias en las que se dan las ondas μ coinciden con el rango de frecuencias de las ondas Alfa. En la Figura 7.1 se puede comprobar como en la banda de las ondas Alfa se encuentra información suficiente para

realizar un buen sistema de BCI, ya que en esta banda se dan la mayor parte de los ERD y de los ERS, pero a la vista de los resultados obtenidos se puede afirmar que el resto de bandas le aportan información extra y que la banda Alfa no contiene la totalidad de la información útil para la clasificación del tipo de movimiento imaginario.

En la figura Figura 7.1 también se ha representado el resultado de tomar todas las bandas a la vez (es decir, coger una sola banda de frecuencia que abarque las cuatro bandas) que, como era de esperar, vemos que empeora los resultados con respecto a tomar solamente la banda Alfa, porque aunque partimos de más información, esta es mezclada e integrada en toda la banda por lo que la información resulta contaminada por más ruido o artefactos.

Las Figura 6.5, Figura 6.6, Figura 6.7 y Figura 6.8 (*figuras en las que se representan los diferentes patrones espaciales de los filtros en cada banda*) se han representado para ver como en cada banda la información reside en partes diferentes del cerebro, ya que cada banda tiene un patrón espacial distinto. Estas figuras también nos sirven para comprobar la similitud que existe entre los patrones de los filtros que se obtienen cuando tomamos las bandas Alfa y Beta a la vez y cuando las tomamos por separado. Cuando tomamos las dos bandas a la vez se escoge un filtro minimizador que resulta muy parecido al de la banda Beta y se toma un filtro maximizador muy parecido al de la banda Alfa, mientras que los patrones de las otras dos bandas difieren bastante de los anteriores.

Otro aspecto a destacar en el uso de las distintas bandas de frecuencia es que en este trabajo se han usado siempre el mismo número de bandas para cada uno de los sujetos y, como se puede observar en las gráficas de los resultados, algunos sujetos empeoran los resultados al usar más bandas de frecuencia, por lo que vemos que tendría sentido personalizar el número de bandas que cada sujeto use.

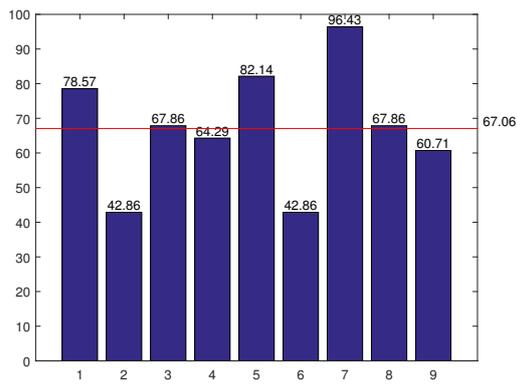
7.2 Zonas de cerebro

Ya que se han realizado pruebas con las cuatro clases de movimientos imaginario, se ha pensado que resultaría interesante ver los patrones espaciales de los filtros que se han obtenido para cada par de clases. Este es el motivo por el que se han ido representado todos los diagramas de las cabezas en cada una de las pruebas. any sentence

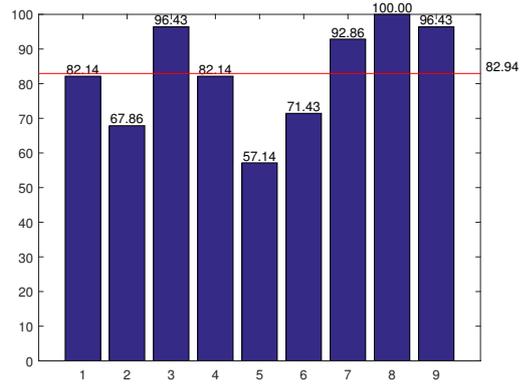
Para interpretar estos diagramas resulta muy útil la Figura 1.5 y recordar que el lado derecho del cerebro controla la parte izquierda del cuerpo y viceversa.

Con esta información, se puede identificar, ver e interpretar como está funcionando el algoritmo CSP. Podemos ver que en las clasificaciones en las que estamos intentando estimar un movimiento imaginario de en la mano izquierda, el algoritmo CSP le da mucha importancia al sensor C4, mientras que en las clasificaciones en las que se está intentando estimar un movimiento imaginario de la mano derecha, el sensor C3 cobra un peso muy alto en el algoritmo de CSP.

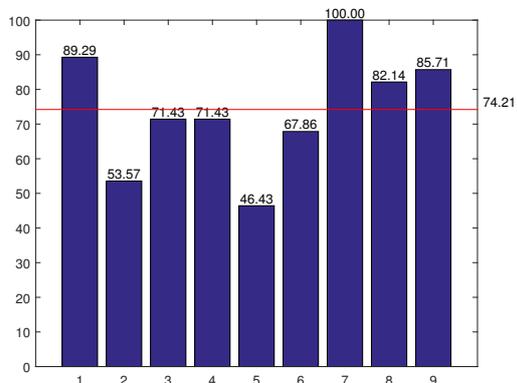
Al pensar en el movimiento de ambos pies, se activan en mayor medida los sensores que están al lado izquierdo y próximos al eje central de la cabeza (el eje que va desde la nariz



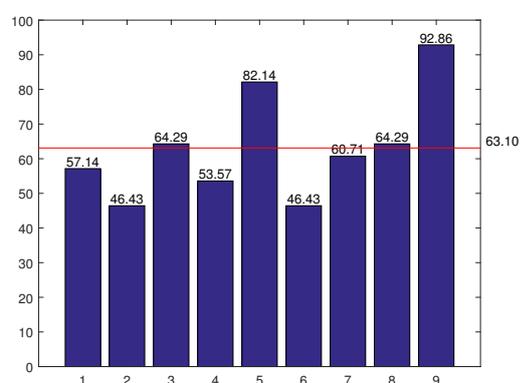
(a) Usando solamente la banda Theta.



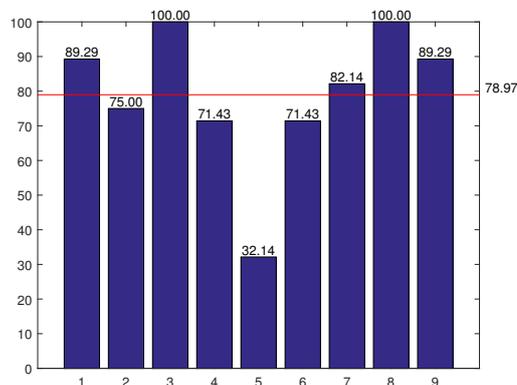
(b) Usando solamente la banda Alfa.



(c) Usando solamente la banda Beta.



(d) Usando solamente la banda Gamma.



(e) Usando las cuatro bandas como una sola.

Figura 7.1 Mano izquierda - Mano derecha: Porcentaje de aciertos de cada una de las bandas por separado y de todas a la vez.

hacia la nuca), lo que nos hace pensar que durante el pensamiento imaginario de los dos pies pensaba con más empeño en el pie derecho (ya que el lado izquierdo parece tener más actividad). Cuando intentamos estimar el movimiento entre los pies y cualquiera de las dos manos, los pesos de los filtros calculados con el algoritmo de CSP parecen repartirse entre

dos sensores de la zona próxima al lóbulo frontal y uno próximo al lóbulo occipital. Sin embargo cuando lo comparamos con el movimiento de la lengua, parece que cobran más importancia los sensores próximos al lóbulo occipital, ignorando los sensores cercanos al lóbulo frontal, lo cual puede tener sentido porque como vamos a ver a continuación, esta zona se asocia también al pensamiento de la lengua, por lo que no serviría para diferenciar entre lengua y pies.

A pesar de que en la figura Figura 1.5 podamos ver que la zona del cerebro que controla la lengua se sitúa en la zona más cercana a las orejas, en los casos en los queremos comparar el movimiento imaginario de la lengua contra el de las dos manos, cobra mucha importancia en el algoritmo CSP, el sensor FZ y el sensor de al lado de FZ que se encuentre en el mismo lado del cuerpo que la mano en la que se piensa. Esto nos resulta muy curioso ya que la zona del cerebro que controla la lengua queda muy alejada del sensor FZ. Sin embargo, cuando comparamos los movimientos de la lengua con los pies, el algoritmo CSP le da gran importancia al sensor más cercano a la oreja izquierda, aunque el sensor FZ sigue manteniendo un peso importante.

Repasando los resultados obtenidos en el apartado 6.6, vemos que el sujeto 4 ha obtenido mejor porcentaje en la clasificación entre la lengua y los pies por lo que es probable que domine mejor estos tipos de movimientos que el sujeto 7. En la Figura 7.2 se han representado los filtros espaciales de este sujeto. Observandola, podemos ver que parece que usa más la zona del cerebro más próxima a las orejas y descarga importancia del sensor FZ (dándosela al sensor de al lado) cuando piensa en la lengua. Para pensar en los pies, sin embargo, usa la misma zona que el sujeto 7.

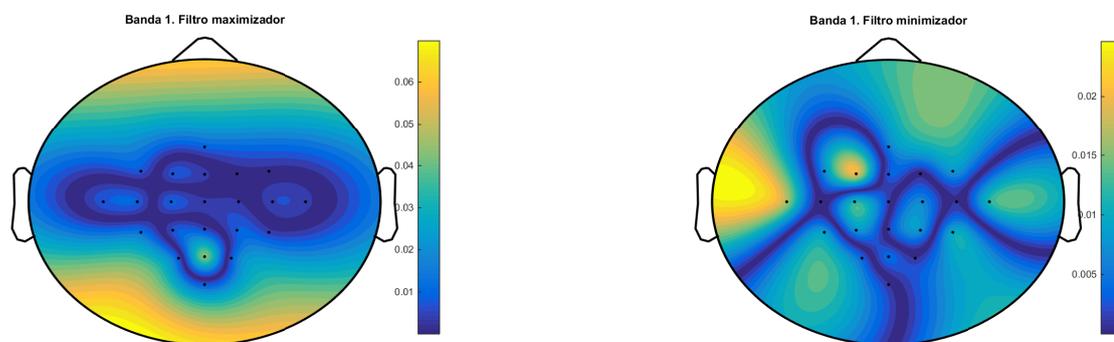


Figura 7.2 Ambos pies - Lengua del sujeto 4: Filtro maximizador y minimizador de las bandas Alfa y Beta a la vez.

7.3 Porcentajes de acierto

Durante el capítulo 6 también se han ido representando gráficas que representan el porcentaje de aciertos que obtenía cada uno de los sujetos en cada una de las pruebas, donde podemos comprobar que algunos de los sujetos siempre obtenían mejores resultados que otros.

Esto es un indicador claro de que para usar correctamente un sistema BCI hace falta tener cierta habilidad o pericia, lo que significa que cualquier persona puede entrenarse para usarlo y que con la práctica se puede ir mejorando en su uso.

En la Figura 7.3 se ha representado la media de aciertos que cada uno de los usuarios ha obtenido entre todos los experimentos (de dos clases) que se han realizado pudiendo comprobar que el sujeto 7 es el que mejor resultados consigue en media. Podemos suponer que los sujetos que consiguen mejores resultados son aquellos que tienen más pericia con el sistema.

En la Figura 7.3 también se ha representado la media en total que se ha conseguido entre todos los usuarios y todos los experimentos, obteniéndose un 87.29%. Otro aspecto a destacar es que podemos ver que todos los usuarios han obtenido un porcentaje de acierto superior al 70%, por lo que todos los sujetos podrían hacer uso de un sistema BCI con este diseño.

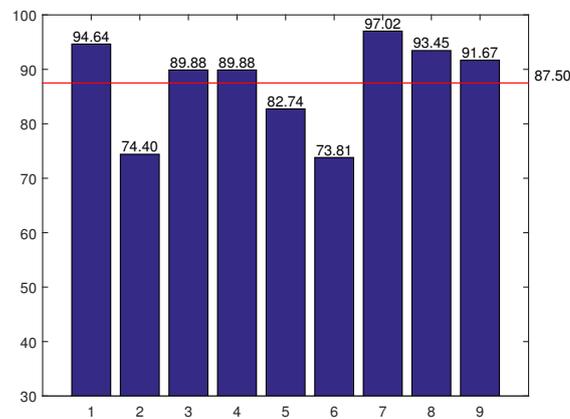


Figura 7.3 Media de aciertos que de cada uno de los usuarios ha obtenido entre todos los experimentos de dos clases.

7.4 Multiclase

Con respecto a los resultados del experimento multiclase, se pueden hacer un par de análisis que expondremos a continuación.

7.4.1 Regiones de decisión

Mientras que en los casos en los que solo comparábamos dos clases siempre obteníamos un porcentaje de acierto mayor del 80% (usando cuatro bandas de frecuencia), al comparar las cuatro clases, tenemos que los resultados han empeorado considerablemente. Esto se debe a que como se explicó en el apartado 4.11 se divide el mismo subespacio en cuatro

regiones mientras que en el caso de dos clases solamente se dividía en dos. En un principio podríamos pensar que al salir una clase ganadora al compararla con las otras tres, esta tendría un porcentaje de acierto mayor, porque no solo se ha comparado con una sino con tres clases distintas. Sin embargo, esto no es así.

Para ilustrar lo que se está explicando se ha creado el escenario sintético de la Figura 7.4. En este escenario se han representado cuatro clases de muestras y todas las fronteras posibles entre ellas: rojo-verde, verde-azul, rojo-azul, etc... En las fronteras (líneas que dividen el cuadrado) hemos incluido unos cuadrados a un lado y al otro de la línea para indicar el color de las clases que separa esa frontera.

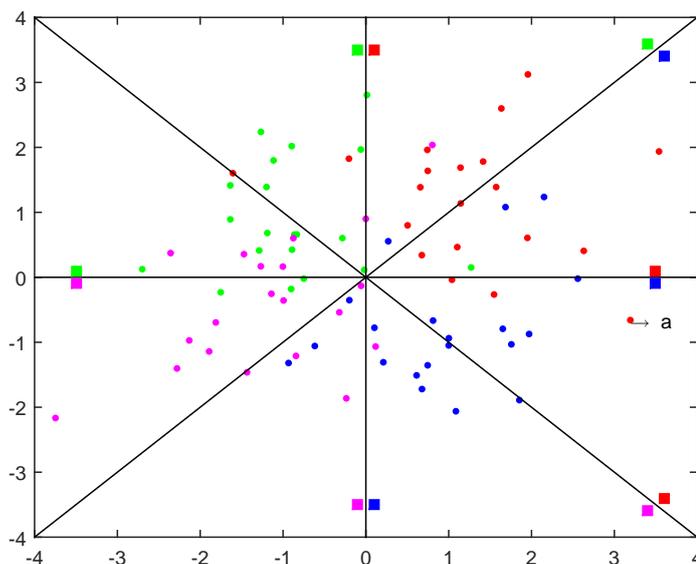


Figura 7.4 Escenario sintético para explicar el problema de la clasificación multiclase.

Vamos a explicar lo que sucede usando la muestra marcada como "a" y las fronteras marcadas en la figura. Si solamente tuviésemos clase roja y verde, la muestra *a* se clasificaría como roja. Esto lo podemos ver, cogiendo la Figura 7.4 y quedándonos tan solo con las muestras rojas y verdes y la frontera que separa dichas muestras. La figura Figura 7.5 sería el resultado. Como podemos ver, puesto que la muestra "a" está en el lado derecho, es decir, el lado rojo, se clasificaría como roja.

Lo mismo pasaría si solamente tenemos muestras rojas y magenta. Sin embargo, al comparar las muestras rojas y azules, nos equivocaríamos al estimar la clase, ya que se estimaría la clase azul. De manera que en los casos en los que solo comparamos dos clases la habríamos clasificado bien dos de las tres veces.

Para el caso multiclase debemos de hacer todas las comparaciones posibles, por lo, además de las comparaciones que ya hemos citado anteriormente, deberíamos de hacer la clasificación entre la clase azul y verde, que daría como resultado azul; y entre la clase azul y magenta, que volvería a dar como resultado la clase azul. Por último se compararían

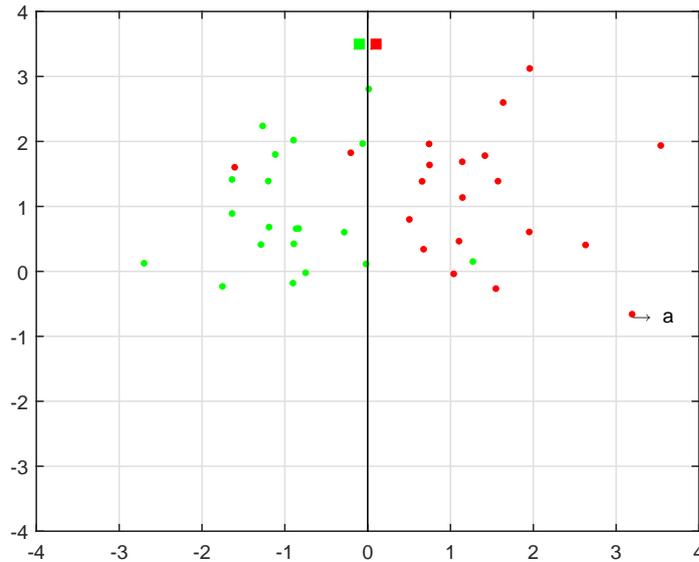


Figura 7.5 Comparación muestras rojas y verdes en escenario sintético.

las clases magenta y verde que darían como resultado la clase magenta.

Tras realizar todas las comparaciones, vemos que la clase roja ha tenido dos votos, la clase azul tres votos, la clase magenta un solo voto y la clase verde ninguno. De manera que la clase estimada sería la azul.

En la Figura 7.4 las cuatro regiones de clasificación serían cada uno de los cuatro cuadrantes, es decir se puede comprobar que si una muestra cae en el primer cuadrante se clasificaría como roja, en el segundo como azul, etc... Es fácil ver que en cualquiera de los casos en los que comparamos solo dos clases es más probable acertar que en el caso multiclase.

Por otro lado, al mirar el problema como una división del subespacio, nos puede ayudar a entender el diagrama de la Figura 6.25 ya que nos da a entender que las regiones que se corresponden con votos de (3-2-1-0) son mucho mayores que el resto de regiones, ya que un 87 % de las muestras han caído en una región de este tipo.

En la Figura 6.25 también podemos comprobar que tratar de estimar la clase de las muestras en las que se ha obtenido en dos clases los mismos votos (2-2-1-1) no ha resultado muy útil ya que ha resultado en que se ha fallado más veces de las que se ha acertado.

7.4.2 Aciertos por clase

Es interesante también analizar que movimientos mentales resultan más fáciles de realizar. En la Figura 6.26 se puede ver que a cada sujeto le cuesta más un movimiento u otro, y esta información podría resultar útil para que cada sujeto practique más los movimientos que peor se le dan. Por ejemplo, se podría decir que el sujeto 2 necesita practicar especialmente

los movimientos de la mano derecha, porque en la gráfica de la sensibilidad vemos que el clasificador ha tenido muy baja tasa de aciertos para este movimiento además de que en la gráfica de la precisión, de las veces que se ha clasificado un ensayo como clase 2, muy pocas veces ha acertado, por lo que deducimos que el clasificador para ese movimiento no está funcionando bien.

Por otro lado, tenemos los casos en los que el clasificador funciona correctamente. Por ejemplo, algunas clasificaciones han acertado todas las muestras de la clase objetivo teniendo mucha precisión, como la clase 2 del sujeto 7, o la clase 1 de este mismo sujeto, que aunque no se han encontrado todos los ensayos de la clase 1, todos los ensayos clasificados como clase 1 lo eran realmente.

Por último, en la Figura 7.6 se ha representado la media de entre todos los sujetos de la sensibilidad¹ y de la precisión² para cada clase. En esta figura vemos que aunque para los pensamientos de la mano derecha se obtiene una sensibilidad mayor, todas las clases obtienen porcentajes similares, por lo que se deduce que unas clases de movimiento no son más fáciles de realizar que otras, si no que simplemente hay sujetos a los que se les da mejor realizar una clase de movimiento que otro.

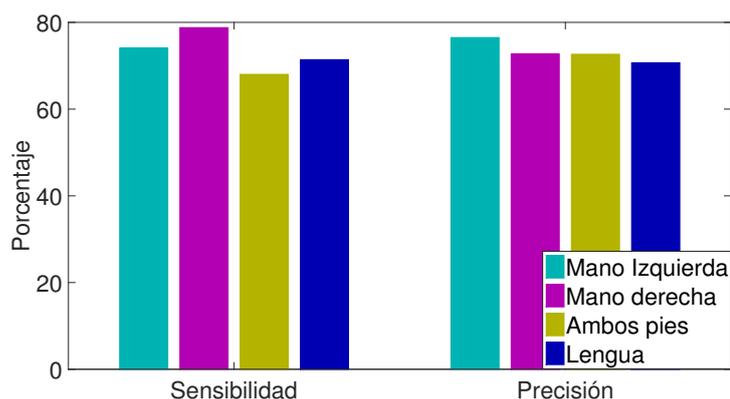


Figura 7.6 Media de sensibilidad y precisión de cada clase entre todos los sujetos.

7.5 Conclusiones

En este capítulo se han analizado en profundidad algunos de los resultados obtenidos en el capítulo 6. Hemos hecho un análisis sobre las ventajas de separar en frecuencia las distintas bandas de las ondas celebrables así como que bandas contienen más información.

¹ La sensibilidad se define como: $Sensibilidad = \frac{VP}{VP+FN}$. Donde VP son verdaderos positivos y FN son falsos negativos. Nos da la medida del porcentaje de las muestras de cada clase que se han acertado.

² La precisión se define como: $Precision = \frac{VP}{VP+FP}$. Donde FP son falsos positivos. Es decir, nos da una medida de entre todas las veces que se ha clasificado con una clase determinada cuantas veces se ha acertado.

Hemos analizado las zonas del cerebro más activas mientras que el sujeto piensa en cada uno de los movimientos imaginarios, y la importancia de como cada persona puede pensar de manera distinta concentrando la información en zonas distintas del cerebro.

Por último se ha analizado el caso multiclase explicando la división del subespacio que se realiza al contemplar más de dos clases.

8 Conclusiones

sometimes the most simple conclusion is also the most correct.

JULIE GARWOOD

Las comunicaciones por BCI suponen un gran avance en la actualidad ya que posibilitan a las personas controlar máquinas (ordenadores) a través de sus pensamientos. La utilidad que podemos encontrar en estos sistemas radica en las ventajas y mejoras que puede aportar en la calidad de vida de personas que sufren algún tipo de discapacidad motora. Para ellos, pequeños avances como podría ser cambiar el canal de la TV mediante sus pensamiento, pueden suponer grandes mejoras.

Durante la realización de este trabajo ha quedado patente que las comunicaciones por BCI son un gran reto para la ciencia ya que abarcan muchas ramas de la ingeniería, entre ellas encontramos ingeniería biomédica, ingeniería electrónica o procesado de señales. Como ya hemos dicho, este tipo de tecnologías pueden resultar de gran utilidad y se espera que sufran grandes avances en el futuro, pero para alcanzarlo es necesario avanzar en muchas direcciones a la vez.

Los sistemas BCI se tratan de sistemas muy complejos, que además tienen el problema añadido de que están dirigidos a personas, por lo que tenemos que tener en cuenta que cada persona es única y distinta de otra. Esto implica (en referencia a este trabajo) que cada persona puede pensar en movimientos de manera diferente. En este trabajo se ha entrenado un sistema BCI para cada persona, pero existen muchos intentos de realizar sistemas que puedan trabajar con la información de múltiples usuarios para hacer que la fase de entrenamiento del sistema sea más corta.

Por un lado, podemos plantearnos la pregunta de si una mayor comprensión del cerebro podría ayudar en el avance de este campo ya que tras estudiar los distintos modelos de comunicaciones BCI (onda P300 y ondas μ) nos planteamos que quizás sea posible encontrar otro modelo en el que el usuario sea capaz de utilizar un sistema BCI a la vez

que realiza otras tareas, al igual que somos capaces de conducir o de manejar un ordenador mientras mantenemos una conversación. Este pensamiento surge del hecho de que en los modelos que se conocen, el usuario debe de estar completamente concentrado en la tarea que realiza para que el ordenador sea capaz de registrar correctamente las ondas surgidas a raíz de sus pensamientos.

Por otro lado, también planteamos si será posible crear tecnologías que permitan una mejor captación de la señal sin necesidad de realizar implantes como en el caso de ECoG, para obtener medidas más limpias y representativas del estado mental de las personas.

En este trabajo se ha profundizado en el tratamiento de señales que como se ha visto es otra parte fundamental de todo sistema BCI. Creemos que en esta línea también se pueden realizar grandes avances. Actualmente se está investigando para tratar de mejorar en todo lo posible el algoritmo CSP ya que su aplicación en los sistemas BCI ha significado una gran aportación convirtiéndose en la técnica más usada para la extracción de características en las señales.

Además de estudiar en detalle el algoritmo CSP, se ha desarrollado una mejora sobre el mismo, que ha consistido en tomar distintas bandas de frecuencias en función de las frecuencias de las distintas ondas cerebrales (Theta, Alfa, Beta y la parte baja de la banda Gamma), aplicar el método de CSP en cada una de ellas por separado y extraer las características que se usaran más tarde para clasificar, en la literatura, a este tipo de técnica se le conoce como FBCSP. Realizando esta mejora se ha conseguido mejorar los resultados de un 82.14 % a un 87.5 % de ensayos acertados¹ llevándonos a la conclusión de que incluir la información proveniente de las distintas bandas del cerebro supone una mejoría.

Se han representado las zonas del cerebro más importantes durante cada tipo de movimiento imaginario, que en realidad se tratan de los coeficientes de los filtros CSP calculados, particularizando el caso para cada banda de frecuencia. Esto nos ha ayudado a sacar conclusiones sobre la manera de pensar en los movimientos imaginarios y qué bandas de frecuencia son más importantes en este sentido, ya que cobraban especial importancia las bandas alpha y beta.

También hemos podido ver que incluso en el caso multiclase donde se consiguen porcentajes de aciertos mucho menores, se han obtenido porcentajes superiores al 70 % en casi todos los usuarios (que es el límite teórico que se establece para que la comunicación en un sistema sea aceptable) habiendo solamente tres usuarios que han obtenido una tasa de acierto inferior a este porcentaje e incluso un usuario que ha obtenido una tasa de acierto del 87 %.

Finalmente, a partir de la información extraída de [35], se ha implementado el algoritmo RSTFC con la intención de comparar los resultados obtenidos con el algoritmo CSP y con el algoritmo desarrollado a lo largo de este trabajo, pero finalmente no ha sido posible ya que no disponíamos de algunos parámetros fundamentales del programa.

Como línea de investigación en el futuro se propone estudiar la manera de incorporar

¹ Para calcular estos porcentajes se ha hecho la media de porcentajes de aciertos de cada caso, mano izquierda-mano derecha, mano izquierda-pies, mano izquierda-lengua...

las distintas bandas de frecuencias intentando conseguir la selección de los mejores filtros espaciales de entre todas las bandas pudiendo escoger un número distinto de filtros de cada banda. Esto podría permitir ahorrar en características que ayudarían a alejarse del overfitting, ya que en este trabajo se han tomado al rededor de cinco filtros por cada banda proporcionando alrededor de 20 características. También creemos que podría ser de utilidad añadir un preprocesado que elimine un mayor número de artefactos y que limpie la señal. Por último, creemos que este algoritmo podría ser más versátil si se pudiese incorporar información de otros usuarios reduciendo el número de muestras de entrenamiento que se usen para cada sujeto.

Índice de Figuras

1.1	Esquema BCI.	3
1.2	EEG de una crisis tónico-clónica.	5
1.3	Lóbulos del cerebro.	6
1.4	Esquema de una neurona.	7
1.5	Correspondencia entre las zonas del lóbulo parietal y las zonas del cuerpo que controlan.	9
1.6	Tabla para deletrear usada en [17].	10
2.1	Ejemplo de CSP tomando una única dimensión	18
2.2	Ejemplo de CSP tomando dos dimensiones	18
2.3	Ejemplo de CSP tomando una tres dimensiones	18
2.4	Nubes de puntos separadas con LDA	22
2.5	Nubes de puntos separadas con LDA y FDP de las nubes de puntos	23
3.1	Esquema de las sesiones	26
3.2	Esquema de los ensayos	26
3.3	Esquema del montaje para la grabacion de las señales EEG	27
4.1	Diagrama de bloques del algoritmo desarrollado	33
4.2	Esquema de método usado para separar en grupos aleatorios	34
4.3	Esquema de banco de filtros	35
4.4	Esquema de cálculo de filtros espaciales (CSP)	36
4.5	Esquema de cálculo de características	37
4.6	Esquema del proceso de estimación de la clase de movimiento	38
4.7	Ejemplo de ejecución del algoritmo multiclase cuando se da el caso en el que dos clases tienen el mismo número de votos	39
6.1	Mano izquierda - Mano derecha: Resultados por sujeto usando las bandas Alfa y Beta como una sola	47
6.2	Mano izquierda - Mano derecha: Filtro maximizador y minimizador de las bandas Alfa y Beta a la vez. Donde se puede apreciar las zonas mas activas de la cabeza durante el pensamiento de estos dos movimientos	48
6.3	Mano izquierda - Mano derecha: Resultados por sujeto usando las bandas Alfa y Beta por separado	48
6.4	Mano izquierda - Mano derecha: Resultados por sujeto usando cuatro bandas	49

6.5	Mano izquierda - Mano derecha: Filtro maximizador y minimizador de la banda Theta	49
6.6	Mano izquierda - Mano derecha: Filtro maximizador y minimizador de la banda Alfa	50
6.7	Mano izquierda - Mano derecha: Filtro maximizador y minimizador de la banda Beta	50
6.8	Mano izquierda - Mano derecha: Filtro maximizador y minimizador de la banda Gamma	50
6.9	Mano izquierda - Ambos pies: Resultados por sujeto usando las bandas Alfa y Beta a la vez	52
6.10	Mano izquierda - Ambos pies: Filtro maximizador y minimizador de las bandas Alfa y Beta a la vez	52
6.11	Mano izquierda - Ambos pies: Resultados por sujeto usando cuatro bandas (Theta, Alfa, Beta y Gamma)	53
6.12	Mano izquierda - Lengua: Resultados por sujeto usando las bandas Alfa y Beta a la vez	54
6.13	Mano izquierda - Lengua: Filtro maximizador y minimizador de las bandas Alfa y Beta a la vez	54
6.14	Mano izquierda - Lengua: Resultados por sujeto usando cuatro bandas (Theta, Alfa, Beta y Gamma)	54
6.15	Mano derecha - Ambos pies: Resultados por sujeto usando las bandas Alfa y Beta a la vez	55
6.16	Mano derecha - Ambos pies: Filtro maximizador y minimizador de las bandas Alfa y Beta a la vez	55
6.17	Mano derecha - Ambos pies: Resultados por usuario usando cuatro bandas (Theta, Alfa, Beta y Gamma)	56
6.18	Mano derecha - Lengua: Resultados por sujeto usando las bandas Alfa y Beta a la vez	56
6.19	Mano derecha - Lengua: Filtro maximizador y minimizador de las bandas Alfa y Beta a la vez	57
6.20	Mano derecha - Lengua: Resultados por sujeto usando cuatro bandas (Theta, Alfa, Beta y Gamma)	57
6.21	Ambos pies - Lengua: Resultados por sujeto usando las bandas Alfa y Beta a la vez	58
6.22	Ambos pies - Lengua: Filtro maximizador y minimizador de las bandas Alfa y Beta a la vez	58
6.23	Ambos pies - Lengua: Resultados por sujeto usando cuatro bandas (Theta, Alfa, Beta y Gamma)	58
6.24	Multiclase: porcentaje de aciertos por sujeto	59
6.25	Diagrama de porcentaje de clases de votos en algoritmo multiclase	60
6.26	Sensibilidad y Precisión por clase y sujeto en el caso multiclase.	60
6.27	RSTFC: Mano izquierda, mano derecha.	61
6.28	RSTFC: Mano izquierda, pies.	61
6.29	RSTFC: Mano izquierda, lengua.	61
6.30	RSTFC: Mano derecha, pies.	62
6.31	RSTFC: Mano derecha, lengua.	62
6.32	RSTFC: Pies, lengua.	62
7.1	Mano izquierda - Mano derecha: Porcentaje de aciertos de cada una de las bandas por separado y de todas a la vez	67
7.2	Ambos pies - Lengua del sujeto 4: Filtro maximizador y minimizador de las bandas Alfa y Beta a la vez	68
7.3	Media de aciertos que de cada uno de los usuarios ha obtenido entre todos los experimentos de dos clases	69
7.4	Escenario sintético para explicar el problema de la clasificación multiclase	70
7.5	Comparación muestras rojas y verdes en escenario sintético	71

7.6	Media de sensibilidad y precisión de cada clase entre todos los sujetos	72
-----	---	----

Índice de Tablas

1	Tabla de signos matemáticos usados en el trabajo	XIII
3.1	Tipos de eventos	28
4.1	Pares de clases de movimiento	32
4.2	Bandas de frecuencias usadas	35
4.3	Posibles resultados de la votación de clases	39
6.1	Parámetros optimizados	46
6.2	Parámetros usados	46
6.3	Diferentes inicializaciones (usando diferentes semillas) con distinto numero de bandas de frecuencia	51
6.4	Matriz de confusión del caso multiclase. En las filas se representan las clases reales, mientras que en las columnas las clases asignadas	60

Glosario

BCI Brain Computer Interface. III, V, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 13, 22, 43, 47, 65, 68, 69, 75, 79

CSP Common Spatial Patterns. III, V, 2, 13, 15, 17, 35, 36, 41, 42, 43, 45, 66, 68, 79

CV Cross Validation. 32, 34, 37, 38, 45, 60

ECoG Electrocoorticografía (del inglés Electrocorticography). 3, 4, 76

EEG Electroencefalograma (del inglés Electroencephalography). 3, 4, 7, 8, 9, 10, 25, 27, 28, 29, 31, 35, 45, 79

EOG Electrooculograma. 25, 26, 27, 28, 29, 34

ERD evento relacionado con la desincronización. 8, 9, 65

ERS evento relacionado con la sincronización. 8, 65

FBCSP Filter Bank Common Spatial Patterns. 31, 76

FDP Función de densidad de probabilidad. 22, 79

fMRI functional Magnetic Resonance Imaging. 3

fNRI functional Near-Infrared Spectroscopy. 3

GDF General Data Format. 27

LDA Linear Discriminant Analysis. III, 13, 18, 19, 20, 22, 41

MEG Magnetoencefalografía. 3

PCA Principal Component Analysis. 13

RCSP Regularized Common Spatial Patterns. 43, 44

RSTFC Regularized Spatio-Temporal Filtering and Classification. 41, 43, 59, 60, 63, 76

SVM Support Vector Machine. 2, 19

Bibliografía

- [1] <http://bio100.class.uic.edu/lecturesf04am/lect22.htm>.
- [2] <https://www.dreamstime.com/photos-images/brain-icon.html#details6421863>.
- [3] *Niedermeyer's electroencephalography; basic principles, clinical applications, and related fields, 6th ed. (online access included)*, Reference and Research Book News (2011), no. 1.
- [4] M. Ahn, S. Ahn, J.H. Hong, H. Cho, K. Kim, B.S. Kim, J.W. Chang, and S.C. Jun, *Gamma band activity associated with bci performance: Simultaneous meg/leeg study*, *Frontiers in Human Neuroscience* **7** (2013), no. DEC, cited By 12.
- [5] Kristen LaMarca Disha Gupta Aleksandra Vuckovic, Jaime A. Pineda and Christoph Guger, *Interaction of bci with the underlying neurological conditions in patients: pros and cons*, *Frontiers Journals, Frontiers in Neuroengineering*, 2015.
- [6] Kai Keng Ang, Zheng Yang Chin, Chuanchu Wang, Cuntai Guan, and Haihong Zhang, *Filter bank common spatial pattern algorithm on bci competition iv datasets 2a and 2b*, *Frontiers in Neuroscience* **6** (2012), 39.
- [7] T. Ball, M. Kern, I. Mutschler, A. Aertsen, and A. Schulze-Bonhage, *Signal quality of simultaneously recorded invasive and non-invasive eeg*, *NeuroImage* **46** (2009), no. 3, 708–716, cited By 94.
- [8] Helen Barbas, *Frontal cortex*, pp. 1289–1334, Springer New York, New York, NY, 2013.
- [9] I.R. Bell, C.M. Baldwin, and G.E. Schwartz, *Illness from low levels of environmental chemicals: Relevance to chronic fatigue syndrome and fibromyalgia*, *American Journal of Medicine* **105** (1998), no. 3 A, cited By 67.
- [10] B. Blankertz, R. Tomioka, S. Lemm, M. Kawanabe, and K. r. Muller, *Optimizing spatial filters for robust eeg single-trial analysis*, *IEEE Signal Processing Magazine* **25** (2008), no. 1, 41–56.
- [11] C. Brunner, R. Leeb, G. R. Muller-Putz, A. Schlogl, and G. Pfurtscheller, *Bci competition 2008 - graz data set a*, Tech. report, Institute for Knowledge Discovery, Graz University of Technology, Austria and Institute for Human-Computer Interfaces, Graz University of Technology, Austria, 2008.

- [12] S.W. Choi, S.E. Chi, S.Y. Chung, J.W. Kim, C.Y. Ahn, and H.T. Kim, *Is alpha wave neurofeedback effective with randomized clinical trials in depression? a pilot study*, *Neuropsychobiology* **63** (2010), no. 1, 43–51, cited By 38.
- [13] Shirley Coyle, Tomás Ward, Charles Markham, and Gary McDarby, *On the suitability of near-infrared (nir) systems for next-generation brain–computer interfaces*, *Physiological Measurement* **25** (2004), no. 4, 815.
- [14] S.M. Coyle, T.E. Ward, and C.M. Markham, *Brain-computer interface using a simplified functional near-infrared spectroscopy system.*, *Journal of neural engineering* **4** (2007), no. 3, 219–226, cited By 218.
- [15] R. Debnath and E.A. Franz, *Perception of hand movement by mirror reflection evokes brain activation in the motor cortex contralateral to a non-moving hand*, *Cortex* **81** (2016), 118–125, cited By 0.
- [16] M. Diykh and Y. Li, *Complex networks approach for eeg signal sleep stages classification*, *Expert Systems with Applications* **63** (2016), 241–248, cited By 0.
- [17] E. Donchin, K.M. Spencer, and R. Wijesinghe, *The mental prosthesis: Assessing the speed of a p300-based brain- computer interface*, *IEEE Transactions on Rehabilitation Engineering* **8** (2000), no. 2, 174–179, cited By 623.
- [18] K. Fukunaga and W. L. G. Koontz, *Application of the karhunen-loève expansion to feature selection and ordering*, *IEEE Transactions on Computers* **C-19** (1970), no. 4, 311–318.
- [19] Henry Gray and Warren H Lewis, *Anatomy of the human body*, Philadelphia, Lea and Febiger., <http://www.biodiversitylibrary.org/bibliography/20311> — First ed. published in London in 1858 under title: Anatomy, descriptive and surgical.
- [20] Christoph Guger, Brendan Z. Allison, and Günter Edlinger, *Brain–computer interface research a state-of-the-art summary*, pp. 93–95, Springer, 2014.
- [21] Eva Maria Hammer, Sebastian Halder, Benjamin Blankertz, Claudia Sannelli, Thorsten Dickhaus, Sonja Kleih, Klaus-Robert Müller, and Andrea Kübler, *Psychological predictors of smr-bci performance*, *Biological Psychology* **89** (2012), no. 1, 80 – 86.
- [22] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An introduction to statistical learning: with applications in r*, Springer Texts in Statistics, Springer, 2013.
- [23] Caitlin Williams Kiley and W. Martin Usrey, *Cortical processing of visual signals*, pp. 655–674, Springer New York, New York, NY, 2013.
- [24] C.W. Kiley and W.M. Usrey, *Cortical processing of visual signals*, 2013, cited By 0.
- [25] J.-Y. Kim, S.-M. Park, K.-E. Ko, and K.-B. Sim, *Optimal eeg channel selection for motor imagery bci system using bpsa and ga*, *Advances in Intelligent Systems and Computing* **208 AISC** (2013), 231–239, cited By 2.
- [26] A. Kübler, N. Neumann, J. Kaiser, B. Kotchoubey, T. Hinterberger, and N.P. Birbaumer, *Brain-computer communication: Self-regulation of slow cortical potentials for verbal communication*, *Archives of Physical Medicine and Rehabilitation* **82** (2001), no. 11, 1533–1539, cited By 0.

- [27] S.M. LaConte, S.J. Peltier, and X.P. Hu, *Real-time fmri using brain-state classification*, *Human Brain Mapping* **28** (2007), no. 10, 1033–1044, cited By 120.
- [28] Bradley C. Lega, Joshua Jacobs, and Michael Kahana, *Human hippocampal theta oscillations and the formation of episodic memories*, *Hippocampus* **22** (2012), no. 4, 748–761.
- [29] Sergio Liberczuk, *Clasificación mediante mapas autoorganizados de señales de eeg orientado al control de interfaces cerebro computadora*.
- [30] Dennis J. McFarland and Jonathan R. Wolpaw, *Brain-computer interfaces for communication and control*, *Commun. ACM* **54** (2011), no. 5, 60–66.
- [31] R. Chris Miall, *Cortical motor control*, pp. 1187–1208, Springer New York, New York, NY, 2013.
- [32] S.T. Morgan, J.C. Hansen, and S.A. Hillyard, *Selective attention to stimulus location modulates the steady-state visual evoked potential*, *Proceedings of the National Academy of Sciences of the United States of America* **93** (1996), no. 10, 4770–4774, cited By 244.
- [33] X. Pei and C. Zheng, *Classification of left and right hand motor imagery tasks based on eeg frequency component selection*, 2008, cited By 2, pp. 1888–1891.
- [34] M. Penttilä, J.V. Partanen, H. Soinen, and P.J. Riekkinen, *Quantitative analysis of occipital eeg in different stages of alzheimer's disease*, *Electroencephalography and Clinical Neurophysiology* **60** (1985), no. 1, 1–6, cited By 219.
- [35] F. Qi, Y. Li, and W. Wu, *Rstfc: A novel algorithm for spatio-temporal filtering and classification of single-trial eeg*, *IEEE Transactions on Neural Networks and Learning Systems* **26** (2015), no. 12, 3070–3082, cited By 5.
- [36] H. Ramoser, J. Muller-Gerking, and G. Pfurtscheller, *Optimal spatial filtering of single trial eeg during imagined hand movement*, *IEEE Transactions on Rehabilitation Engineering* **8** (2000), no. 4, 441–446.
- [37] Supratim Ray, Nathan E. Crone, Ernst Niebur, Piotr J. Franaszczuk, and Steven S. Hsiao, *Neural correlates of high-gamma oscillations (60–200 hz) in macaque local field potentials and their potential implications in electrocorticography*, *The Journal of Neuroscience* **28** (2008), no. 45, 11526–11536.
- [38] L. Roijendijk, S. Gielen, and J. Farquhar, *Classifying regularized sensor covariance matrices: An alternative to csp*, *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **24** (2016), no. 8, 893–900, cited By 0.
- [39] W. Samek, M. Kawanabe, and K.-R. Muller, *Divergence-based framework for common spatial patterns algorithms*, *IEEE Reviews in Biomedical Engineering* **7** (2014), 50–72, cited By 29.
- [40] Saeid Sanei and J.A. Chambers, *Introduction to eeg*, pp. 1–34, John Wiley and Sons Ltd., 2007.
- [41] Bernhard Schölkopf, John Platt, and Thomas Hofmann, *Logistic regression for single trial eeg classification*, pp. 1377–1384, MIT Press, 2007.

- [42] R. Tomioka and K. Aihara, *Classifying matrices with a spectral regularization*, vol. 227, 2007, cited By 1, pp. 895–902.
- [43] M. Toscani, T. Marzi, S. Righi, M.P. Viggiano, and S. Baldassi, *Alpha waves: A neural signature of visual suppression*, *Experimental Brain Research* **207** (2010), no. 3-4, 213–219, cited By 15.
- [44] J.J. Vidal, *Toward direct brain-computer communication.*, *Annual review of biophysics and bioengineering* **2** (1973), 157–180, cited By 323.
- [45] J.R. Wolpaw and D.J. McFarland, *Multichannel eeg-based brain-computer communication*, *Electroencephalography and Clinical Neurophysiology* **90** (1994), no. 6, 444–449, cited By 278.
- [46] S. Yan, Q. Yu, and H. Wang, *Meg classification based on band power and statistical characteristics*, 2015, cited By 0, pp. 255–258.