



Universidad Zaragoza



Escuela de
Ingeniería y Arquitectura
Universidad Zaragoza

Proyecto Fin de Carrera
Ingeniería en Informática

CLASIFICACIÓN DE LOS ESTADOS DE REPOSO, PREPARACIÓN Y EJECUCIÓN DEL MOVIMIENTO DEL BRAZO A TRAVÉS DEL ELECTROENCEFALOGRAMA

Autor: Miguel Rodrigo Gómez

Directores: Luis Montesano del Campo
Javier Mínguez Zafra



Departamento de Informática
e Ingeniería de Sistemas

Zaragoza, Septiembre 2011

Índice general

A. Fundamentos teóricos	9
A.1. Estimador r^2	9
A.1.1. Coeficiente de correlación	10
A.2. Métodos lineales de clasificación	11
A.2.1. Introducción	11
A.2.2. LDA	11
A.2.3. Shrinkage	12
A.2.4. Pruebas con ERPS	14
B. Procesado de la señal	19
B.1. Sujeto 1	19
B.2. Sujeto 2	22
B.3. Sujeto 3	25
B.4. Sujeto 4	28
B.5. Sujeto 5	31
B.6. Sujeto 6	34
C. Resultados detallados con las tres clases simultáneas	37
C.1. Sujeto 1	37
C.1.1. LDA	37
C.1.2. Shrinkage	37
C.2. Sujeto 2	38
C.2.1. LDA	38
C.2.2. Shrinkage	38
C.3. Sujeto 3	39
C.3.1. LDA	39
C.3.2. Shrinkage	39
C.4. Sujeto 4	39
C.4.1. LDA	39

C.4.2. Shrinkage	40
C.5. Sujeto 5	40
C.5.1. LDA	40
C.5.2. Shrinkage	41
C.6. Sujeto 6	41
C.6.1. LDA	41
C.6.2. Shrinkage	41
D. Código de Matlab	43
D.1. Introducción	43
D.2. Procesado de la señal	43
D.2.1. Funciones Auxiliares	44
D.3. Clasificadores	44
D.3.1. Funciones Auxiliares	45
E. Artículo científico	47

Índice de figuras

A.1. correlación entre dos variables X e Y	11
B.1. Coeficiente r^2 de la señal de EEG en crudo del sujeto 1	19
B.2. Coeficiente de determinación de la señal de EEG con los distintos filtros seleccionados para el sujeto 1	20
B.3. Coeficiente r^2 para la selección del procesado de la señal para el sujeto 1	21
B.4. Coeficiente r^2 para la selección del procesado de la señal para el sujeto 2	22
B.5. Coeficiente r^2 para la selección del procesado de la señal para el sujeto 2	23
B.6. Tipos de señal seleccionadas para la clasificación para el sujeto 2	24
B.7. Separabilidad de las clases para el sujeto 3 con los datos en crudo	25
B.8. Filtros bipolares elegidos para el sujeto 3	26
B.9. Tipo de señal elegida para el sujeto tres en el caso de clasificación por parejas	27
B.10. Datos <i>raw</i> para el sujeto 4	28
B.11. Bandas de frecuencias seleccionadas para el sujeto 4	29
B.12. Tipo de procesado de la señal elegido para el sujeto número 4	30
B.13. Separabilidad de las clases para el sujeto 5 con los datos en crudo	31
B.14. r^2 de la señal de EEG del sujeto 5 con los filtros bipolares seleccionados	32
B.15. Tipo de procesado aplicado al EEG del sujeto quinto	33
B.16. Coeficiente de determinación para el sujeto 6 con los datos <i>raw</i>	34
B.17. Filtros bipolares elegidos para el sujeto 3	35
B.18. Filtros bipolares seleccionados para el sujeto 6	36

Índice de cuadros

A.1. Resultados de la clasificación de los ERPS para los cinco sujetos	14
A.2. Clasificación con LDA de los ERPS del Sujeto 1	15
A.3. Clasificación con Shrinkage de los ERPS del Sujeto 1	15
A.4. Clasificación con LDA de los ERPS del Sujeto 2	15
A.5. Clasificación con Shrinkage de los ERPS del Sujeto 2	15
A.6. Clasificación con LDA de los ERPS del Sujeto 3	16
A.7. Clasificación con Shrinkage de los ERPS del Sujeto 3	16
A.8. Clasificación con LDA de los ERPS del Sujeto 4	16
A.9. Clasificación con Shrinkage de los ERPS del Sujeto 4	16
A.10. Clasificación con LDA de los ERPS del Sujeto 5	17
A.11. Clasificación con Shrinkage de los ERPS del Sujeto 5	17
C.1. Promedio de clasificación para los tres estados usando "LDA"	37
C.2. Clasificación para los tres estados usando "shrinkage"	38
C.3. Promedio de clasificación para los tres estados usando "LDA"	38
C.4. Clasificación para los tres estados usando "shrinkage"	38
C.5. Promedio de clasificación para los tres estados usando "LDA"	39
C.6. Clasificación para los tres estados usando "shrinkage"	39
C.7. Promedio de clasificación para los tres estados usando "LDA"	40
C.8. Clasificación para los tres estados usando "shrinkage"	40
C.9. Promedio de clasificación para los tres estados usando "LDA"	40
C.10. Clasificación para los tres estados usando "shrinkage"	41
C.11. Clasificación para los tres estados usando "LDA"	41
C.12. Clasificación para los tres estados usando "shrinkage"	42

Apéndice A

Fundamentos teóricos

A.1. Estimador r^2

El "*coeficiente de determinación*", r^2 , es la proporción de la variabilidad existente en un conjunto de datos explicados por un modelo estadístico. Su principal propósito es la predicción de resultados sobre la base de otra información que esté relacionada con ésta. Además, proporciona una medida de cómo de bien se van a ajustar los datos futuros a las predicciones del modelo.

Hay muchas definiciones estadísticas de r^2 pero la que nos interesa es: el coeficiente de determinación equivale al cuadrado del coeficiente de correlación. En este caso, no es una medida, estrictamente hablando, de cómo de buenos son los datos del modelo, sino de cómo de buenos serán los datos construidos en comparación con los datos del modelo. En el caso que nos atañe, esta medida la usamos para saber cómo de buenos son los datos de la señal filtrada con respecto al modelo que teníamos anterior: la señal en crudo.

Si tenemos que $X_{raw} \in \mathbb{R}^{CXf}$, la señal en crudo en el espectro de frecuencias para un estado, e $Y_{raw} \in \mathbb{R}^{CXf}$, la señal en crudo para otro estado distinto también en el espectro frecuencial, el coeficiente de determinación es:

$$r_{raw}^2 = (\text{corr}(X_{raw}, Y_{raw}))^2, r_{raw}^2 \in [0, 1] \quad (\text{A.1})$$

Y con los distintos tipos de filtrados:

$$r_{filtered}^2 = (\text{corr}(X_{filtered}, Y_{filtered}))^2, r_{filtered}^2 \in [0, 1] \quad (\text{A.2})$$

El objetivo del procesado consistió en buscar el filtro que maximizar el valor de $r_{filtered}$ con respecto al calculado para r_{raw} .

A.1.1. Coeficiente de correlación

El "coeficiente de correlación" o "coeficiente de correlación momento-producto de Pearson", r , es la medida de la dependencia lineal entre dos variables X e Y. Su valor matemático es:

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X}) * (Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} * \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}}, r \in [-1, 1] \quad (\text{A.3})$$

Un valor de 1 implica que la ecuación lineal que describe la relación existente entre X e Y se ajusta a la perfección. Por otro lado un valor de 0 da a entender que no hay ninguna correlación entre los datos. Por último -1 indica que hay una relación inversa. En la **Figura A.1** podemos ver la representación visual:

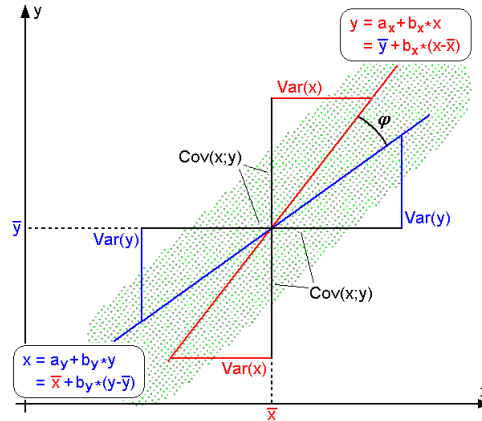


Figura A.1: correlación entre dos variables X e Y

A.2. Métodos lineales de clasificación

A.2.1. Introducción

El término "métodos lineales de clasificación" hace referencia a que los denominados "bordes de decisión" o "decision boundaries", que nos separan cada una de las K clases en las que se encuentra divididos los datos, son lineales. Cada una de las clases k , viene caracterizada por su función de densidad de probabilidad $f_k(X)$, siendo $X = (X_1, X_2, \dots, X_n)$ un conjunto de medidas de la clase. El borde de decisión entre la clase l y la clase m es el conjunto de puntos los cuales cumplen $f_l(x) = f_m(x)$. El método, a grandes rasgos, consiste en modelar las "funciones discriminantes" $\delta_k(x)$ para cada una de las clases y asignar x a la clase k que minimice el valor de esta función.

A.2.2. LDA

Este método se basa en tres principios:

1. Las características de cada clase siguen una distribución *Gaussiana*.
2. Las distribuciones *Gaussianas* de todas las clases tienen la misma matriz de covarianza, Σ .
3. Las distribuciones de las clases son conocidas. Este punto evidentemente nunca se cumple del todo puesto que siempre hay que estimar la media $\hat{\mu}$ y la covarianza $\hat{\Sigma}$.

Con estas tres asunciones, la función de densidad de la probabilidad de la clase k seguiría una distribución normal siendo su función discriminante:

$$\delta_k(X) = (X - \hat{\mu}_k)' * \hat{\Sigma}^{-1} * (X - \hat{\mu}_k) + \ln|\hat{\Sigma}| - 2 * \ln\pi_k \quad (\text{A.4})$$

Con

$$\pi_k = \frac{W_k}{W} \quad (\text{A.5})$$

Siendo W_k el número de muestras de la clase k y W el número total de muestras.

$\hat{\mu}_k$ es la media de las muestras de la clase k y $\hat{\Sigma}$ es la covarianza de todas las muestras de todas las clases. El primer miembro de la ecuación A.4 es la conocida como *distancia Mahalanobis*.

Por lo tanto una observación cualquiera x se clasificará a la clase $\hat{k}$ siguiendo la regla de clasificación:

$$\delta_{\hat{k}}(x) = \min_k \delta_k(x), \forall k \in K \quad (\text{A.6})$$

A.2.3. Shrinkage

El problema del LDA es que el estimador para la matriz de covarianza es la covarianza empírica, éste es impreciso para conjuntos de datos con un número de dimensiones muy elevada en comparación a las muestras, porque el número de parámetros desconocidos que hay que estimar es de orden cuadrático con respecto al número de dimensiones. Esto desemboca en otro error: los valores propios de la covarianza que son grandes son estimados más grandes todavía y los que son pequeños se estiman aún más pequeños.

Shrinkage es un método muy utilizado para compensar este error al estimar las matrices de covarianza, para lo cuál la matriz de covarianza estimada es:

$$\tilde{\Sigma}(\gamma) = (1 - \gamma) * \Sigma + \gamma * \nu * I \quad (\text{A.7})$$

Siendo $\gamma \in [0, 1]$ el parámetro de afinación, ν el promedio de la traza de los autovalores dividido entre la dimensión de los datos e " I " la matriz identidad. Tenemos que se puede descomponer en sus valores propios: $\tilde{\Sigma} = V * D * V'$, con V ortonormal y D diagonal. Sustituyéndolo en A.7:

$$\tilde{\Sigma}(\gamma) = (1 - \gamma) * V * D * V' + \gamma * \nu * I = (1 - \gamma) * V * D * V' + \gamma * \nu * V * I * V' \quad (\text{A.8})$$

$$\tilde{\Sigma}(\gamma) = V * ((1 - \gamma) * D + \gamma * \nu * I) * V' \quad (\text{A.9})$$

Esto implica que los valores propios en los extremos se modifican con respecto a ν . Además usando $\gamma = 1$ asume que la matriz de covarianza es esférica mientras que $\gamma = 0$ nos da como resultado LDA.

Para la obtención del parámetro de afinación γ se puede usar un método de validación cruzada, el cuál es muy costoso computacionalmente y, además, es complejo de programar. Para remediar esto, en algunas divulgaciones científicas [16], [22] se propone la selección automática del parámetro de afinación óptimo, γ^* :

Siendo $(x_k)_i$ el elemento i -ésimo del vector x_k y s_{ij} la i -ésima fila y la j -ésima columna de la matriz de covarianza, $\tilde{\Sigma}$, definimos:

$$z_{ij}(k) = ((x_k)_i - (\hat{\mu})_i) * ((x_k)_j - (\hat{\mu})_j) \quad (\text{A.10})$$

$$\gamma^* = \frac{n}{(n-1)^2} * \frac{\sum_{i,j=1}^d \text{var}_k(z_{ij}(k))}{\sum_{i,j} s_{ij}^2 + \sum_i (s_{ii} - \nu)^2} \quad (\text{A.11})$$

A.2.4. Pruebas con ERPS

En las primeras fases del proyecto, probamos lo óptimos que podrían ser éstos dos métodos de clasificación lineal, para comprobarlo, los utilizamos con unos datos de potenciales de error suministrados por el Grupo de Robótica de la Universidad de Zaragoza. Estos datos están descritos en [23].

Las diferencias principales que existen con los datos de movimiento son:

1. Dos clases nada más: feedback de error y feedback de no-error.
2. 34 canales en lugar de los 28 del movimiento.
3. A la señal se le aplica un filtrado CAR.
4. A la señal se le aplica un filtrado de paso bajo a 10 Hz.
5. Sólo hay cinco sujetos.

Los resultados que se obtuvieron para los cinco sujetos fueron:

SUJETO	PORCENTAJE	
1	81'50	81'01
2	74'12	73'11
3	81'18	76'99
4	79'17	78'67
5	75'84	73'83
PROMEDIO	78'36	76'72

Tabla A.1: Resultados de la clasificación de los ERPS para los cinco sujetos

Como podemos comprobar, obtuvimos unos muy buenos resultados con los dos clasificadores para las dos clases, lo que nos animó a probar estos mismos clasificadores en el problema más complejo del EEG relativo al movimiento del brazo derecho.

A continuación se muestran los resultados particulares de cada uno de los distintos sujetos:

1. SUJETO 1

- **LDA** El porcentaje de acierto para este sujeto es del 81'50 %

REAL VS CLASIFICADO	NO-ERROR	ERROR
NO-ERROR	80'99	19'01
ERROR	18'04	81'96

Tabla A.2: Clasificación con LDA de los ERPS del Sujeto 1

- **SHRINKAGE** El porcentaje de acierto para este sujeto es del 81'01 %

REAL VS CLASIFICADO	NO-ERROR	ERROR
NO-ERROR	81'34	18'66
ERROR	19'30	80'70

Tabla A.3: Clasificación con Shrinkage de los ERPS del Sujeto 1

2. SUJETO 2

- **LDA** El porcentaje de acierto para este sujeto es del 74'12 %

REAL VS CLASIFICADO	NO-ERROR	ERROR
NO-ERROR	68'20	31'80
ERROR	20'50	79'50

Tabla A.4: Clasificación con LDA de los ERPS del Sujeto 2

- **SHRINKAGE** El porcentaje de acierto para este sujeto es del 73'11 %

REAL VS CLASIFICADO	NO-ERROR	ERROR
NO-ERROR	67'49	32'51
ERROR	21'77	78'23

Tabla A.5: Clasificación con Shrinkage de los ERPS del Sujeto 2

3. SUJETO 3

- **LDA** El porcentaje de acierto para este sujeto es del 81'18 %

REAL VS CLASIFICADO	NO-ERROR	ERROR
NO-ERROR	83'44	16'55
ERROR	21'05	78'95

Tabla A.6: Clasificación con LDA de los ERPS del Sujeto 3

- **SHRINKAGE** El porcentaje de acierto para este sujeto es del 76'99 %

REAL VS CLASIFICADO	NO-ERROR	ERROR
NO-ERROR	77'03	22'97
ERROR	23'03	76'47

Tabla A.7: Clasificación con Shrinkage de los ERPS del Sujeto 3

4. SUJETO 4

- **LDA** El porcentaje de acierto para este sujeto es del 79'17 %

REAL VS CLASIFICADO	NO-ERROR	ERROR
NO-ERROR	80'67	19'33
ERROR	22'33	77'67

Tabla A.8: Clasificación con LDA de los ERPS del Sujeto 4

- **SHRINKAGE** El porcentaje de acierto para este sujeto es del 78'67 %

REAL VS CLASIFICADO	NO-ERROR	ERROR
NO-ERROR	80'67	19'33
ERROR	23'33	76'67

Tabla A.9: Clasificación con Shrinkage de los ERPS del Sujeto 4

5. SUJETO 5

- **LDA** El porcentaje de acierto para este sujeto es del 75'84 %

REAL VS CLASIFICADO	NO-ERROR	ERROR
NO-ERROR	79'87	20'23
ERROR	28'57	71'43

Tabla A.10: Clasificación con LDA de los ERPS del Sujeto 5

- **SHRINKAGE** El porcentaje de acierto para este sujeto es del 73'83 %

REAL VS CLASIFICADO	NO-ERROR	ERROR
NO-ERROR	77	23
ERROR	29'62	70'38

Tabla A.11: Clasificación con Shrinkage de los ERPS del Sujeto 5

Apéndice B

Procesado de la señal

B.1. Sujeto 1

El primer paso es obtener la señal en crudo en el espectro frecuencia y mostrar el coeficiente de correlación entre las tres clases por parejas:

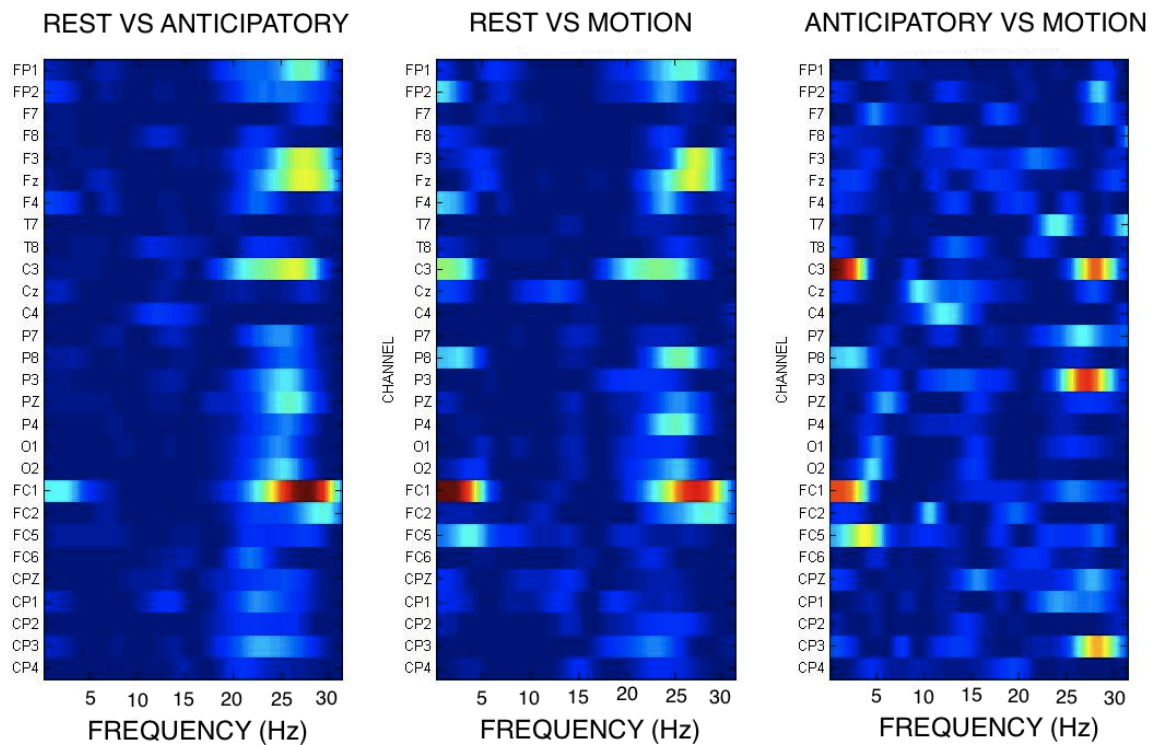


Figura B.1: Coeficiente r^2 de la señal de EEG en crudo del sujeto 1

Luego con la señal filtrada bipolarmente comprobamos el rango de frecuencias con mayor diferencia estadística. En la **Figura B.2** entre *reposo* - *preparación* tenemos el filtro bipolar *FC1*, entre *reposo* - *movimiento* tenemos de nuevo un filtro bipolar *FC1* y entre *preparación* - *movimiento* *P3*. Como se puede apreciar entre *reposo* - *preparación* y *reposo* - *movimiento* se produce una desincronización en la banda β , mientras que entre *preparación* - *movimiento* se produce en las bandas α y δ .

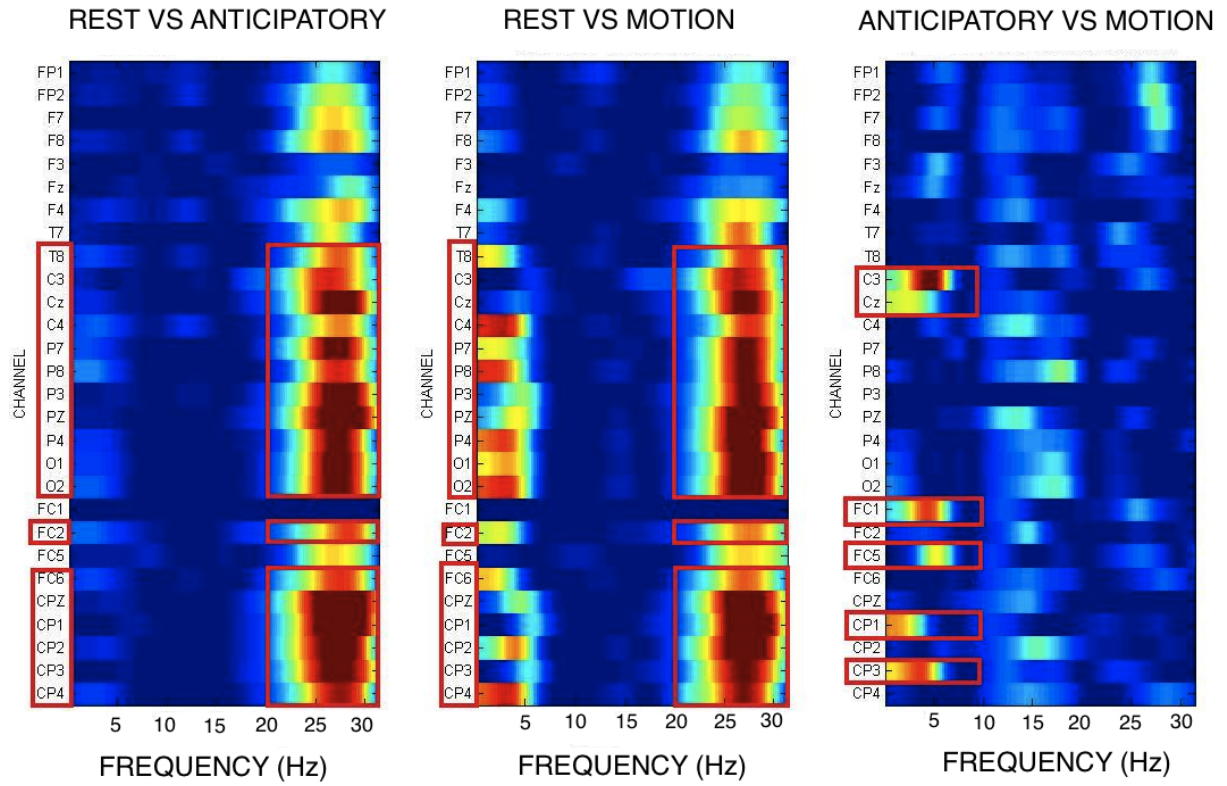


Figura B.2: Coeficiente de determinación de la señal de EEG con los distintos filtros seleccionados para el sujeto 1

En último caso, mostramos cuál del procesado de la señal se seleccionó para la clasificación final: señal suavizada para el *reposo* - *preparación* y para *reposo* - *movimiento* y con la señal filtrada frecuencialmente para el caso de *preparación* - *movimiento*.

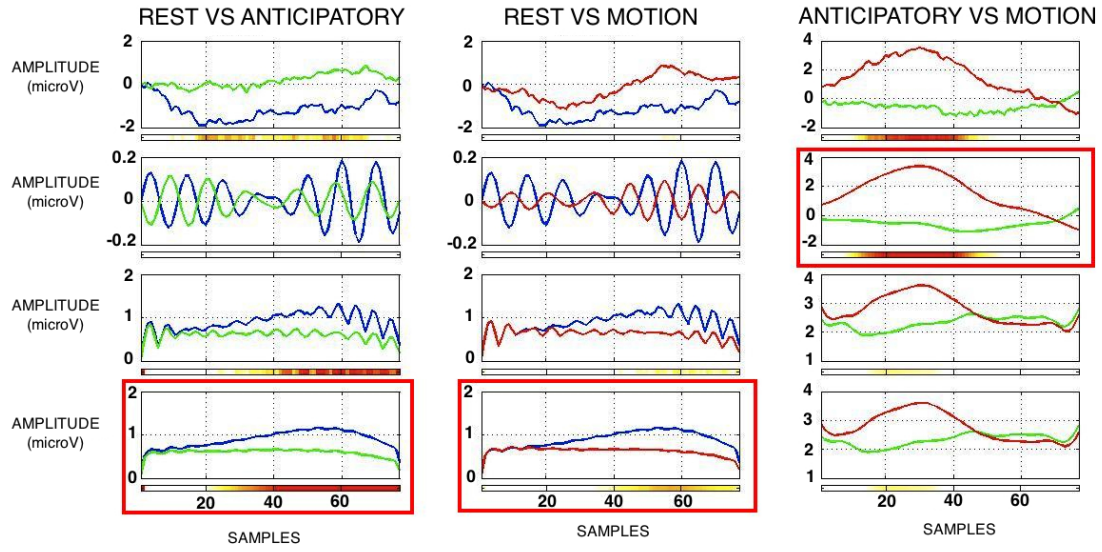


Figura B.3: Coeficiente r^2 para la selección del procesado de la señal para el sujeto 1

B.2. Sujeto 2

A continuación en la **Figura B.4** el punto de partida del proceso: la señal *raw* del sujeto 2:

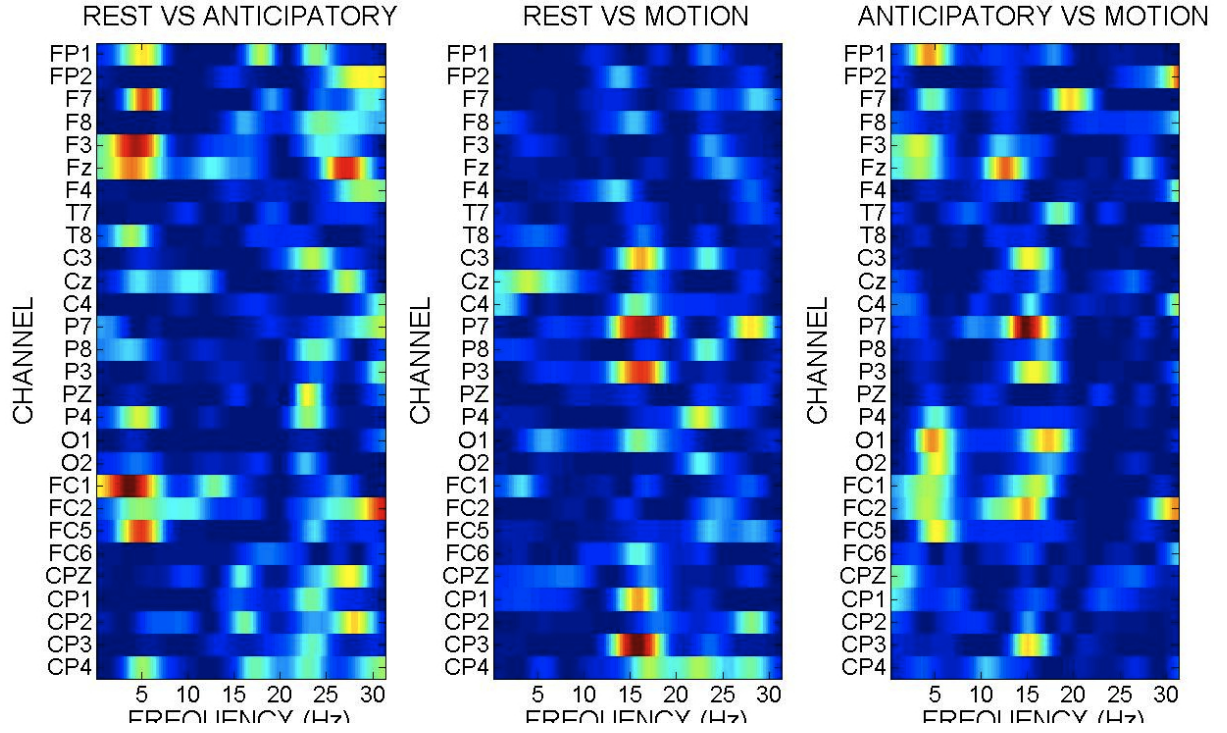


Figura B.4: Coeficiente r^2 para la selección del procesamiento de la señal para el sujeto 2

La **Figura B.5** muestra el coeficiente de determinación para los filtros bipolares seleccionados para el sujeto 2: bipolar Fz para el *reposo - preparación*, bipolar $FC2$ para los casos de *reposo - movimiento* y *preparación - movimiento*. Para *reposo - preparación* aparece la mayor diferencia estadística en la banda β mientras que en los otros dos casos aparecen entre las bandas α y β .

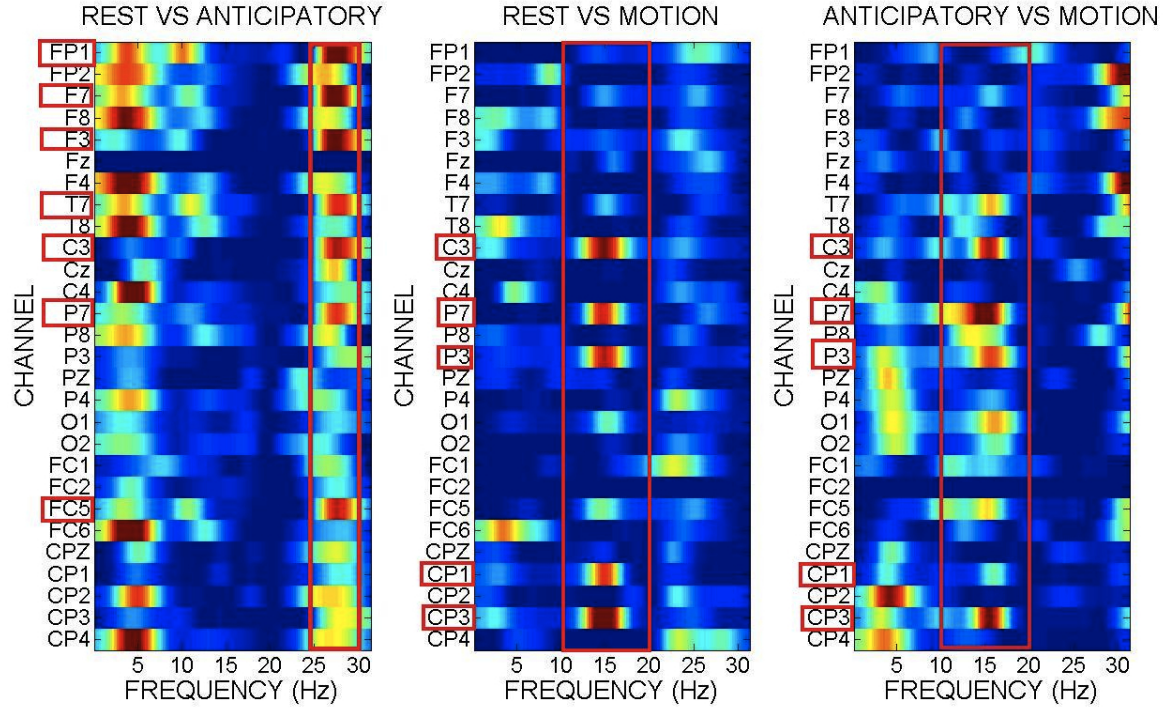


Figura B.5: Coeficiente r^2 para la selección del procesado de la señal para el sujeto 2

Por último, comparamos los 4 tipos de señales para clasificar: filtrado espacial, frecuencia, rectificado y suavizado y mostramos con cuál de ellos nos quedamos: en los tres casos nos quedamos con la señal filtrada únicamente con el filtro bipolar.

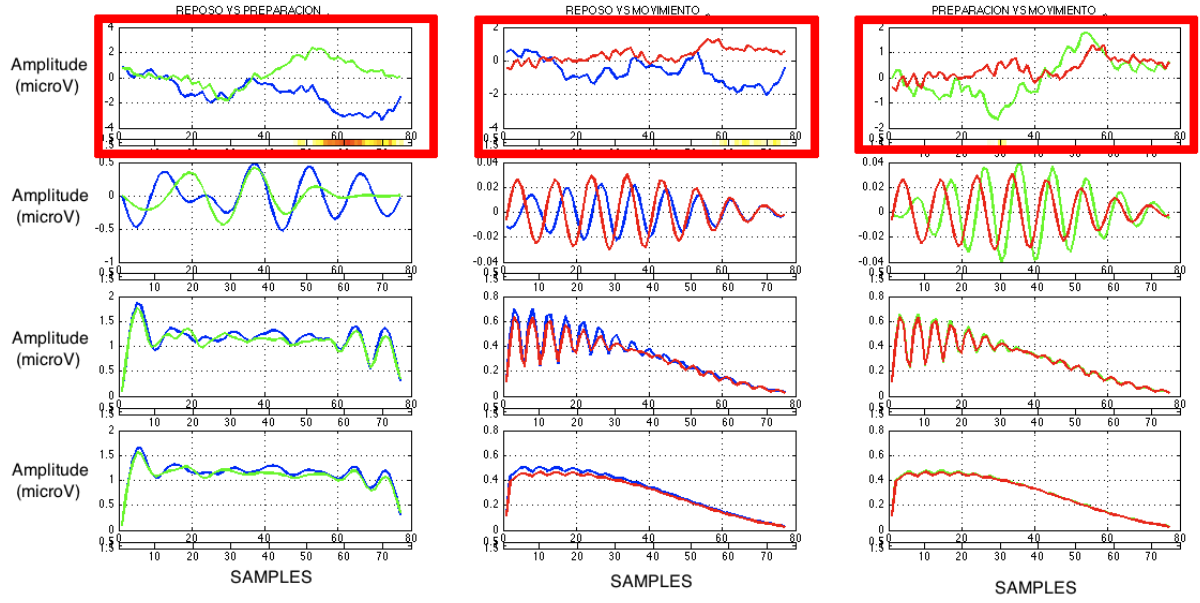


Figura B.6: Tipos de señal seleccionadas para la clasificación para el sujeto 2

B.3. Sujeto 3

Como en todos los casos primero calculamos el coeficiente de determinación del sujeto número tres para los datos sin ningún tipo de procesado:

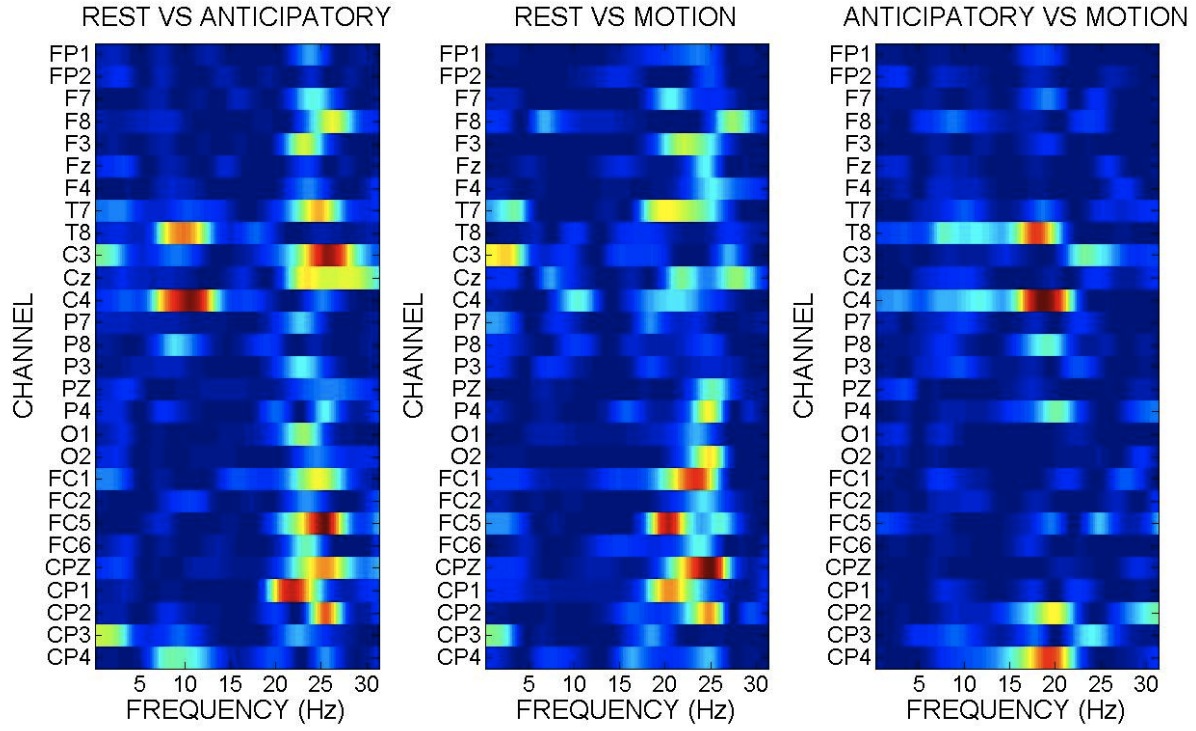


Figura B.7: Separabilidad de las clases para el sujeto 3 con los datos en crudo

Para este sujeto en particular se han elegido las siguientes bandas: en el caso de *preparación - movimiento* implementamos un filtro de paso banda entre 20 y 30 Herzios, para el caso de *reposo - movimiento* utilizamos un filtro de paso bajo a 5 Hz, mientras que para el caso de *reposo - movimiento* usamos un filtro de paso banda entre 20 y 30 Herios, como el primero. Seleccionamos para el caso de *reposo - movimiento* un filtro bipolar *CP3*, al igual que para el caso de *preparación - movimiento* el bipolar elegido fue con el canal *F7*.

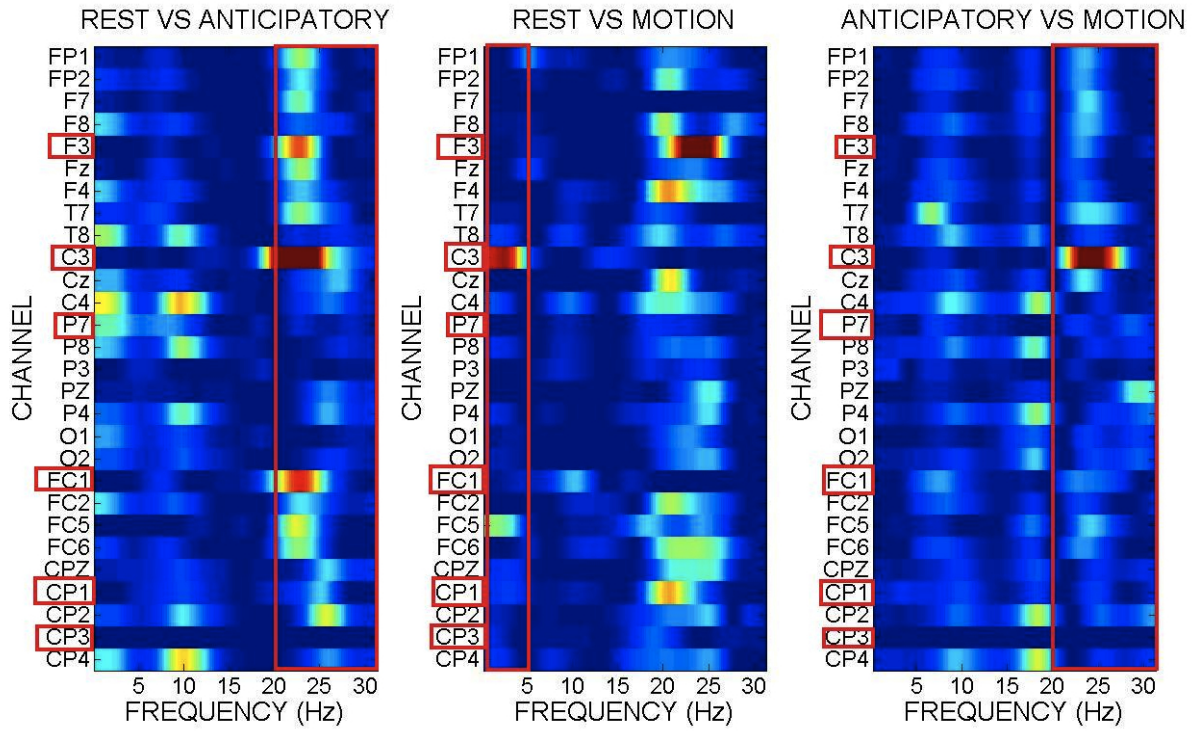


Figura B.8: Filtros bipolares elegidos para el sujeto 3

Este sujeto es peculiar puesto que para clasificar las clases por parejas se utilizó entre *reposo* - *preparación* y *preparación* - *movimiento* la señal filtrada espacialmente nada más y para el caso de *reposo* - *movimiento* un filtro entre el espectro de frecuencias seleccionadas, pero se da la particularidad que para el caso de tener tres clases es mejor utilizar, en vez de la señal filtrada bipolarmente, la señal suavizada. Este hecho apareció de manera empírica cuando se estaban probando a mejorar los resultados con otros procesados que no eran los que revelaba el análisis r^2 .

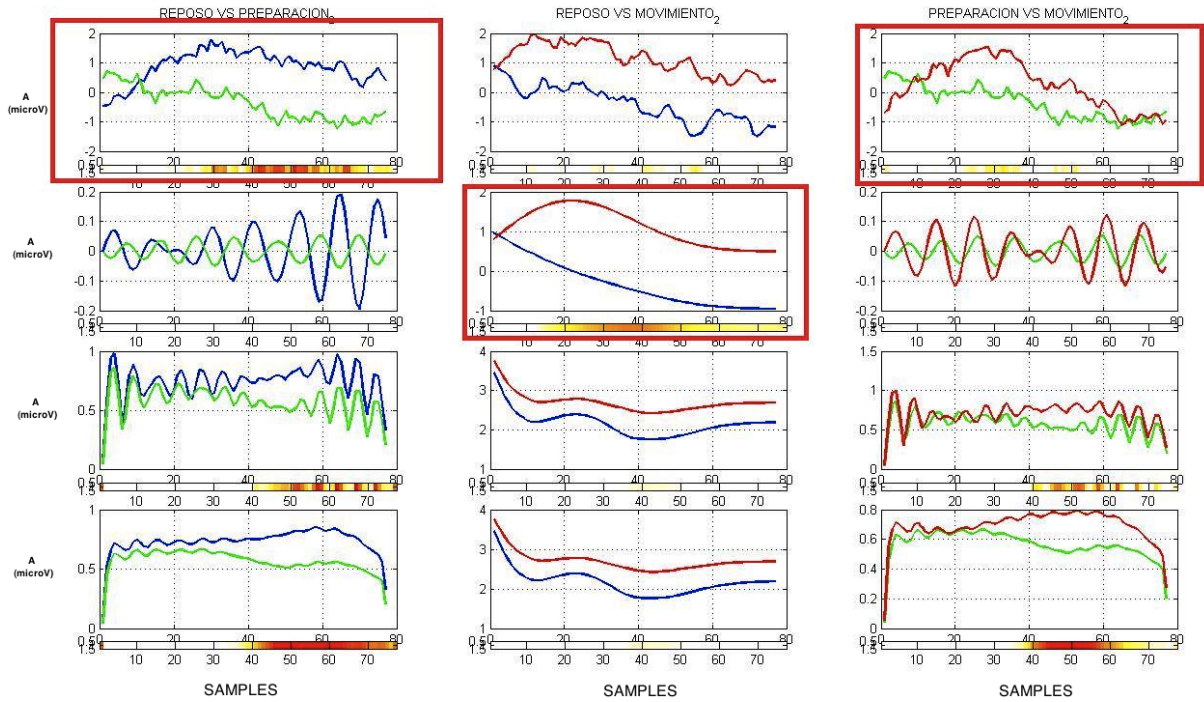


Figura B.9: Tipo de señal elegida para el sujeto tres en el caso de clasificación por parejas

B.4. Sujeto 4

Primero tenemos el coeficiente de determinación para el sujeto 4 y los datos del movimiento en *raw*:

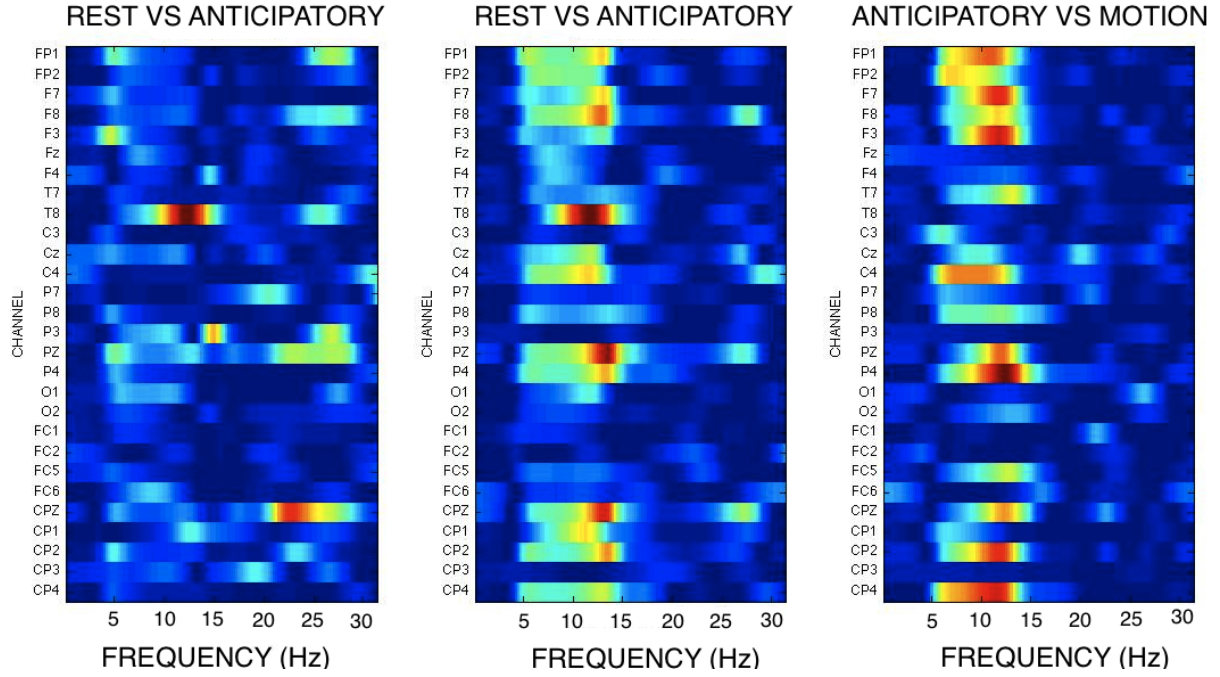


Figura B.10: Datos *raw* para el sujeto 4

En esta ocasión, para mejorar las características y la separabilidad de los estados de reposo, anticipación y ejecución del movimiento en sí, comprobamos que de entre todos los filtros el más idóneo para el caso de reposo - anticipación era un filtro bipolar CPz , al igual que para la anticipación - ejecución. Por otro lado, el *reposo - movimiento* necesitó otro filtro bipolar distinto, uno entre el canal $CP1$. Como podemos observar en la **Figura B.11**, aparecen desincronizaciones en la banda β de la señal, es decir entre 12 - 30 Hz, pero se diferencian en el punto de la banda, pues la primera pareja tiene la desincrinización en 20 - 25 Hz mientras que las otras dos lo tienen en 20 - 15 Hz.

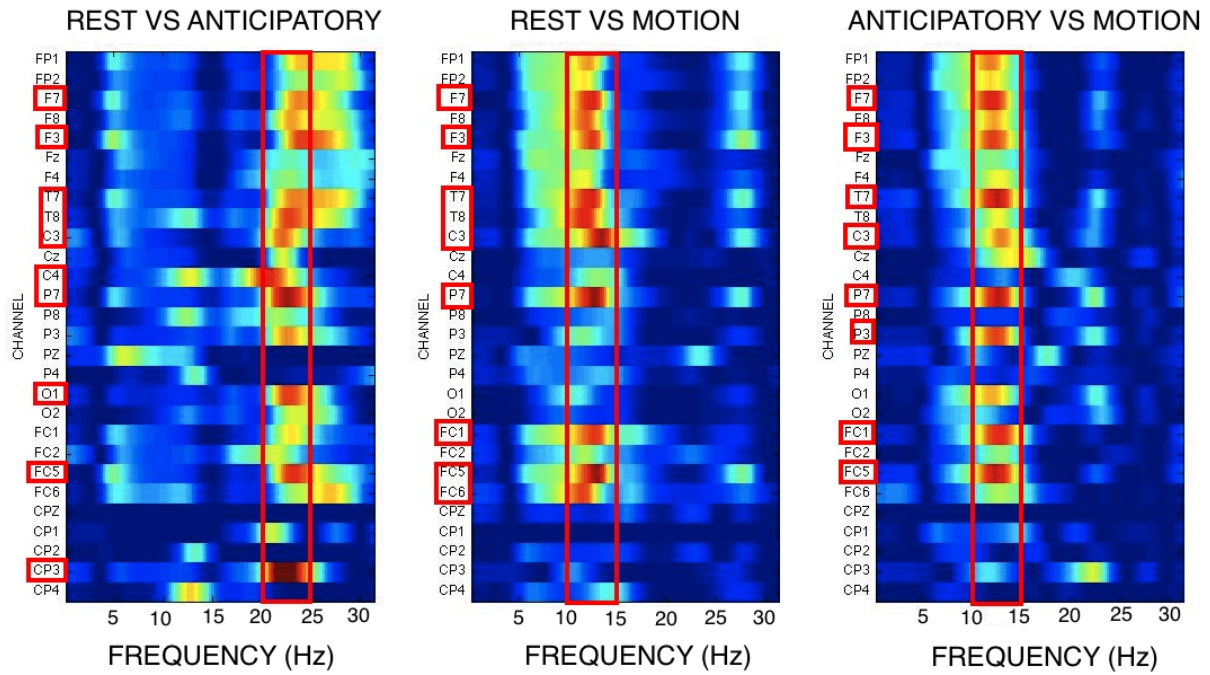


Figura B.11: Bandas de frecuencias seleccionadas para el sujeto 4

Para este sujeto en concreto, elegimos para clasificar la señal con el filtrado espacial exclusivamente para el *reposo - preparación* y para la preparación y el movimiento, por contra, para el caso de *preparación - movimiento* utilizamos la señal después de pasar por todo el procesado quedándonos con la señal suavizada.

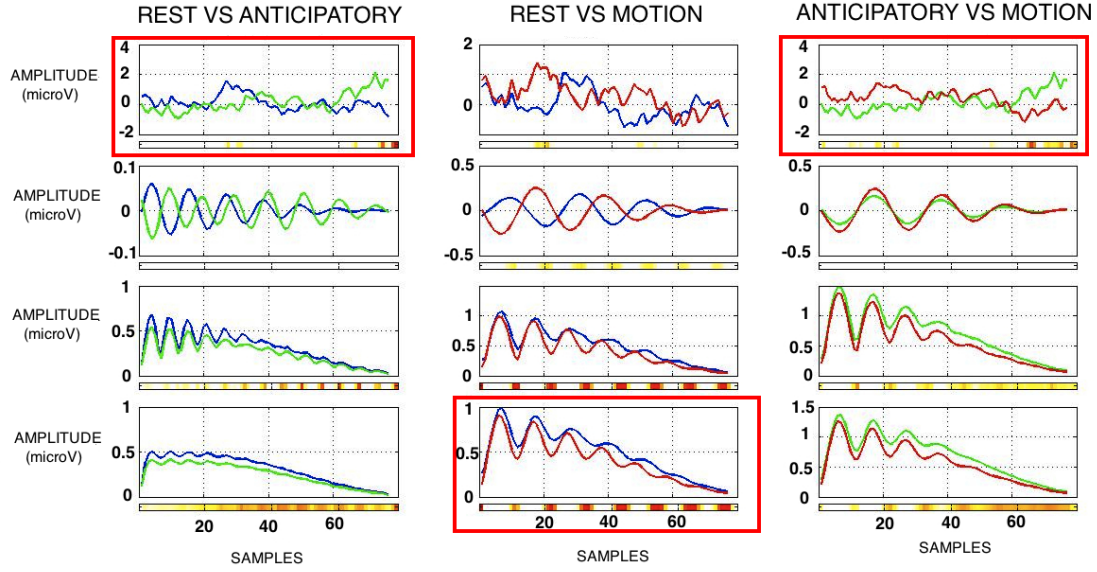


Figura B.12: Tipo de procesado de la señal elegido para el sujeto número 4

B.5. Sujeto 5

El sujeto 5 es el que peor resultados de clasificación dio y como podemos ver en los datos en crudo que se muestran en la **Figura B.13** no se intuye ninguna posible banda de actuación para realzar las características de la señal. Este EEG responde a un *"non-responder"*, puesto que éste no se parece en nada al de los otros sujetos.

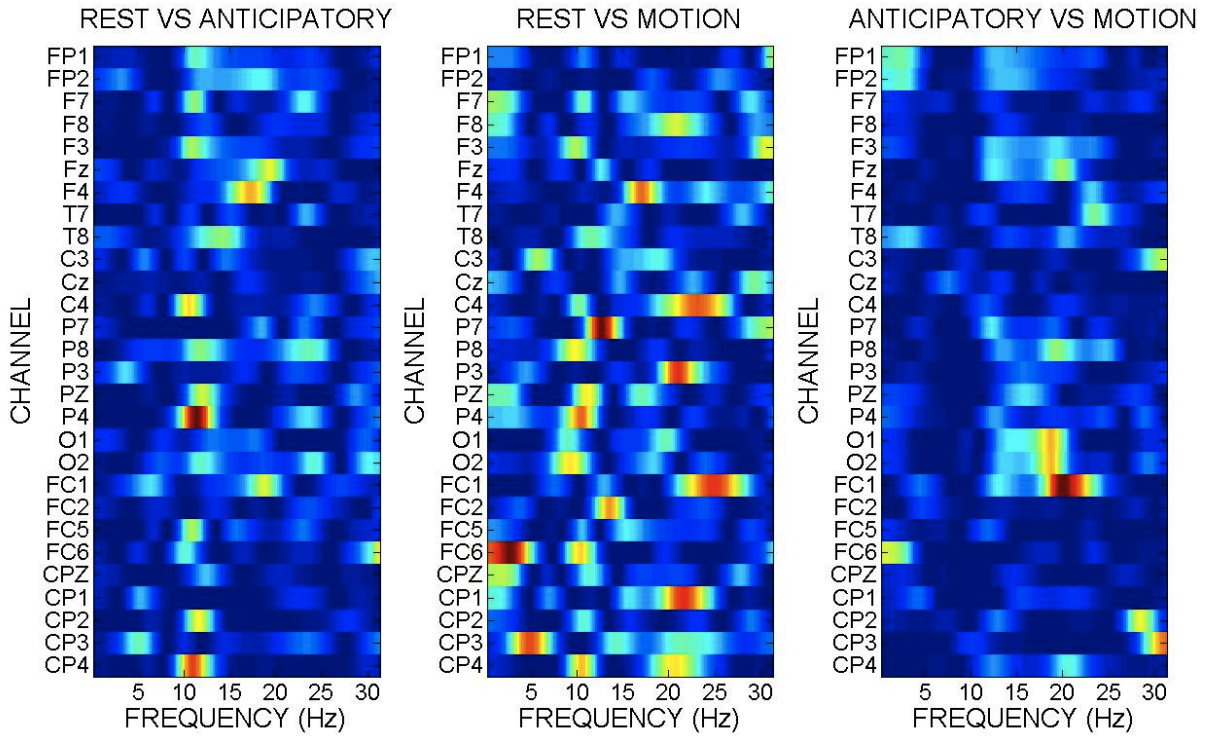


Figura B.13: Separabilidad de las clases para el sujeto 5 con los datos en crudo

Con los filtros seleccionados, observamos que para el caso de *reposo - preparación* va a ser una tarea ardua conseguir un resultado positivo, pues cada canal tiene una banda totalmente distinta a sus vecinos, a esto hay que sumar el que en la banda de *preparación - movimiento* a penas hay información significativa en canales que no sean el que nos interesa, *C3*.

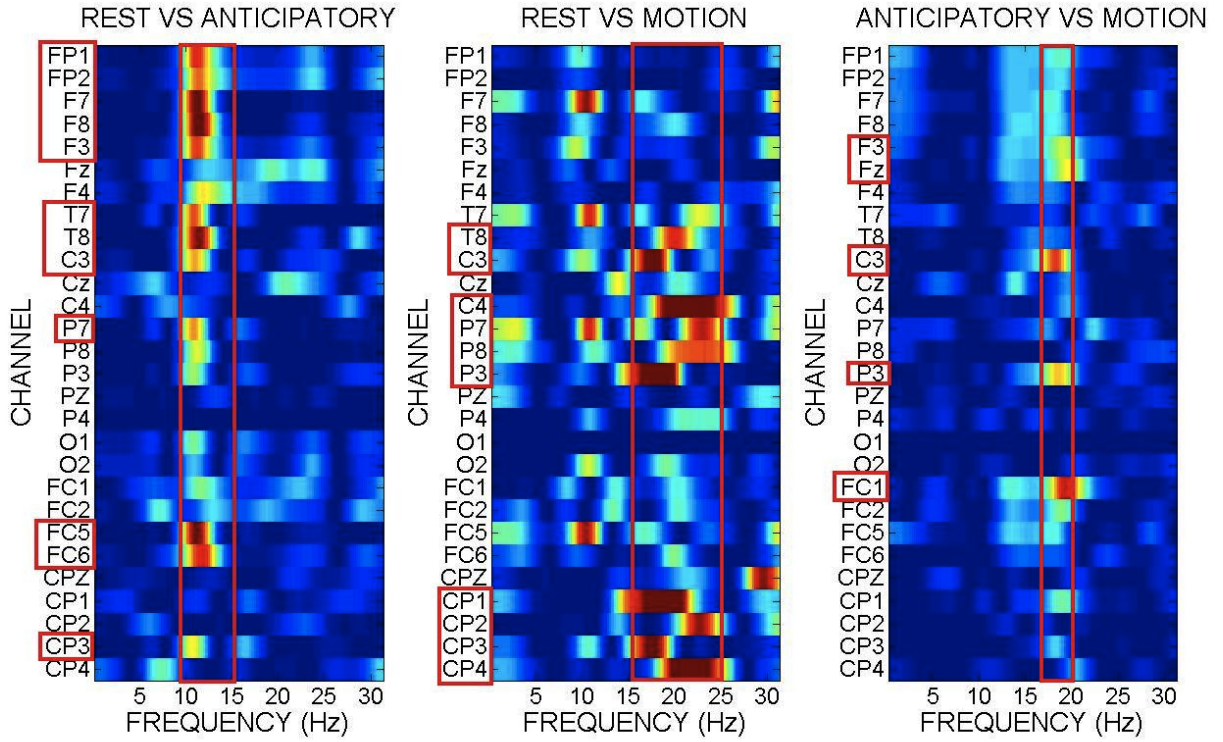


Figura B.14: r^2 de la señal de EEG del sujeto 5 con los filtros bipolares seleccionados

A todo lo anterior expuesto, a la hora de encontrar un valor significativo de r^2 en los tres tipos de parejas, descubrimos que no había diferencia significativa entre ellas. Por todos estos datos recopilados es por lo que podemos aventurar que el sujeto 5 es un claro candidato a ser "non-responder".

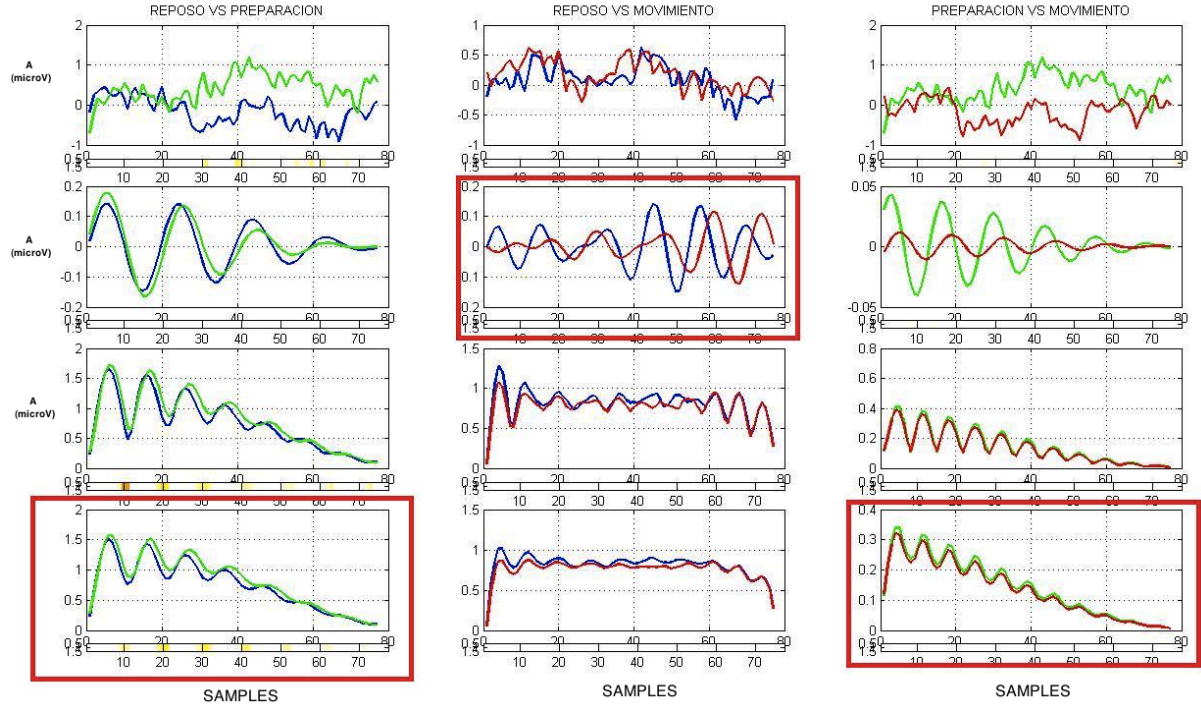


Figura B.15: Tipo de procesado aplicado al EEG del sujeto quinto

B.6. Sujeto 6

Primero tenemos el punto de partida común a todos los sujetos, que es la señal de EEG relativa al movimiento del brazo derecho sin ningún tipo de tratamiento en el espectro de frecuencias:

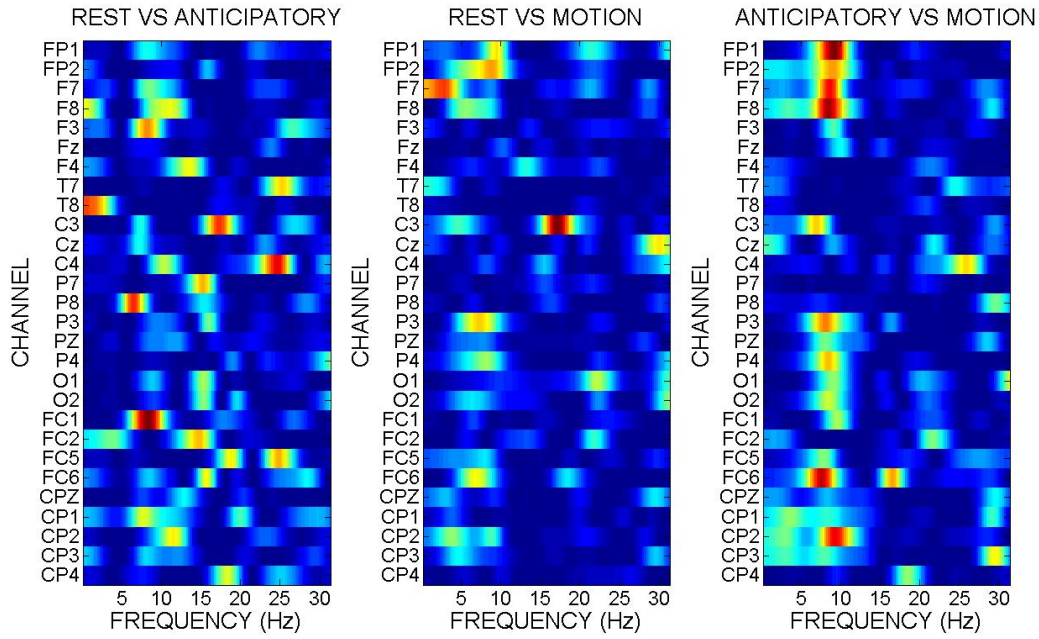


Figura B.16: Coeficiente de determinación para el sujeto 6 con los datos *raw*

Tras un estudio pormenorizado de todos los filtros posibles al final elegimos, como en todos los casos, unos filtros bipolares, que para este sujeto en particular fueron: *P8* para la pareja reposo - anticipación, para la segunda pareja, *reposo - movimiento*, seleccionamos un filtro *FP1* y, por último, también un filtrado *FP1* para la preparación y el movimiento.

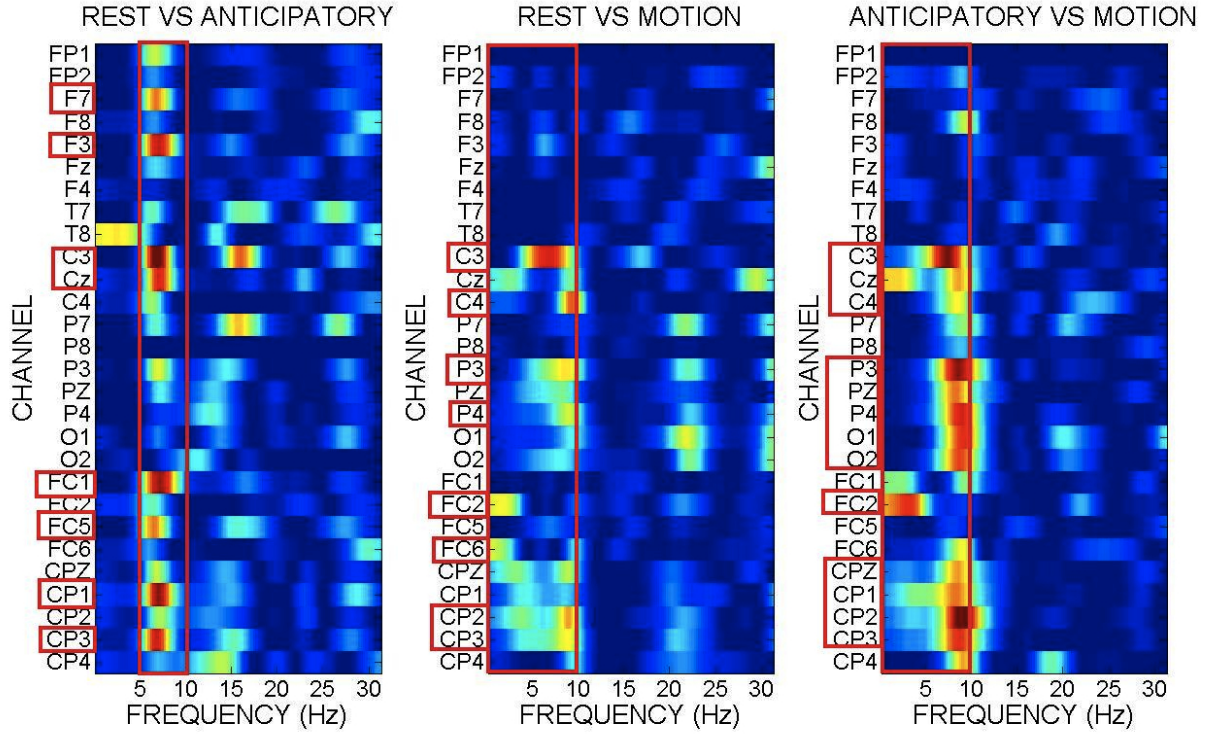


Figura B.17: Filtros bipolares elegidos para el sujeto 3

Para este sujeto en particular, como podemos apreciar en la **Figura B.18** en las tres parejas en liza seleccionamos la señal filtrada con sus respectivos bipolares como señal con las mejores características posibles para su clasificación.

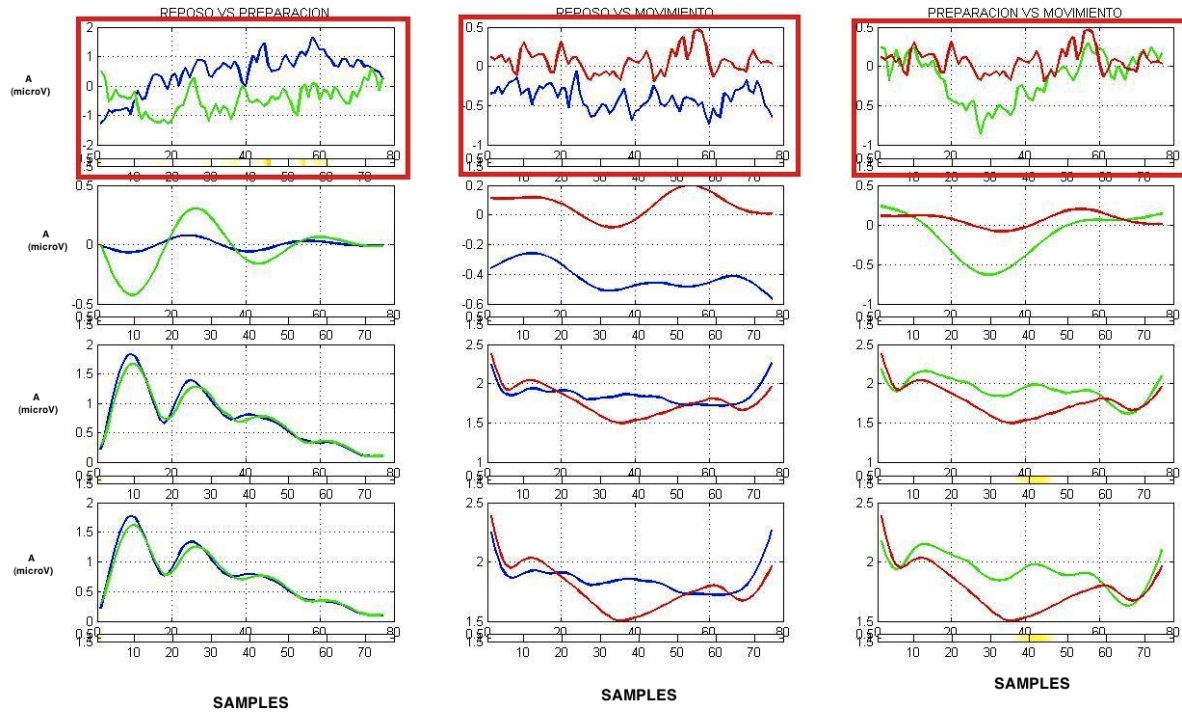


Figura B.18: Filtros bipolares seleccionados para el sujeto 6

Apéndice C

Resultados detallados con las tres clases simultáneas

C.1. Sujeto 1

C.1.1. LDA

El promedio de clasificación para este sujeto ronda el 68'19 %

REAL VS CLASIFICADO	REPOSO	PREP	MOVIMIENTO
REPOSO	62'75	25	12'25
PREPARACIÓN	19'39	64'80	15'81
MOVIMIENTO	8'16	14'80	77'04

Tabla C.1: Promedio de clasificación para los tres estados usando "*LDA*"

C.1.2. Shrinkage

El promedio de aciertos es del 72'08 %

REAL VS CLASIFICADO	REPOSO	PREP	MOVIMIENTO
REPOSO	68'88	21'43	9'69
PREPARACIÓN	16'84	69'39	13'77
MOVIMIENTO	9'69	12'24	78'66

Tabla C.2: Clasificación para los tres estados usando "*shrinkage*"

C.2. Sujeto 2

C.2.1. LDA

Ahora el número de aciertos con respecto al total se mueve en torno al 58'13 %

REAL VS CLASIFICADO	REPOSO	PREP	MOVIMIENTO
REPOSO	57'14	25'51	17'35
PREPARACIÓN	24'49	54'08	21'43
MOVIMIENTO	16'33	20'41	63'26

Tabla C.3: Promedio de clasificación para los tres estados usando "*LDA*"

C.2.2. Shrinkage

La tasa de aciertos se sitúa para este sujeto en un 61'84 %

REAL VS CLASIFICADO	REPOSO	PREP	MOVIMIENTO
REPOSO	61'22	23'47	15'31
PREPARACIÓN	22'45	60'20	17'34
MOVIMIENTO	15'30	20'41	64'29

Tabla C.4: Clasificación para los tres estados usando "*shrinkage*"

C.3. Sujeto 3

C.3.1. LDA

La tasa de aciertos repunta en este caso hasta el 62'21 %

REAL VS CLASIFICADO	REPOSO	PREP	MOVIMIENTO
REPOSO	61'78	15'92	22'29
PREPARACIÓN	15'92	65'61	18'47
MOVIMIENTO	22'93	17'83	59'24

Tabla C.5: Promedio de clasificación para los tres estados usando "*LDA*"

C.3.2. Shrinkage

El porcentaje de aciertos para el individuo número 3 alcanzó el 54'33 %

REAL VS CLASIFICADO	REPOSO	PREP	MOVIMIENTO
REPOSO	59'24	17'20	23'57
PREPARACIÓN	18'47	54'14	27'39
MOVIMIENTO	21'66	28'66	49'68

Tabla C.6: Clasificación para los tres estados usando "*shrinkage*"

C.4. Sujeto 4

C.4.1. LDA

Ahora alcanza su valor mínimo en LDA con un 41'33 % de aciertos.

REAL VS CLASIFICADO	REPOSO	PREP	MOVIMIENTO
REPOSO	36	29'33	30'67
PREPARACIÓN	34'67	41'33	25'33
MOVIMIENTO	32	21'33	46'67

Tabla C.7: Promedio de clasificación para los tres estados usando "*LDA*"

C.4.2. Shrinkage

En este caso el promedio decayó al 39'93 %

REAL VS CLASIFICADO	REPOSO	PREP	MOVIMIENTO
REPOSO	32	37'33	30'66
PREPARACIÓN	36	41'33	22'66
MOVIMIENTO	20	20'48	46'67

Tabla C.8: Clasificación para los tres estados usando "*shrinkage*"

C.5. Sujeto 5

C.5.1. LDA

El quinto sujeto se sitúa en el 43'21 %

REAL VS CLASIFICADO	REPOSO	PREP	MOVIMIENTO
REPOSO	40	21'18	38'52
PREPARACIÓN	26'67	46'67	26'67
MOVIMIENTO	30'37	26'67	42'96

Tabla C.9: Promedio de clasificación para los tres estados usando "*LDA*"

C.5.2. Shrinkage

En el mejor escenario posible para el sujeto 5 sólo alcanzamos el 40'23 % de la clasificación correctamente.

REAL VS CLASIFICADO	REPOSO	PREP	MOVIMIENTO
REPOSO	34'82	26'67	38'52
PREPARACIÓN	24'44	46'67	28'85
MOVIMIENTO	36'81	27'41	37'78

Tabla C.10: Clasificación para los tres estados usando "*shrinkage*"

C.6. Sujeto 6

C.6.1. LDA

Por último, alcanza un 47'18 %

REAL VS CLASIFICADO	REPOSO	PREP	MOVIMIENTO
REPOSO	47'69	21'54	30'77
PREPARACIÓN	22'31	50'77	26'92
MOVIMIENTO	34'61	22'31	43'08

Tabla C.11: Clasificación para los tres estados usando "*LDA*"

C.6.2. Shrinkage

Para el último sujeto, obtuvimos un 48 % bien clasificado

REAL VS CLASIFICADO	REPOSO	PREP	MOVIMIENTO
REPOSO	46'92	20'77	32'31
PREPARACIÓN	20	54'61	25'38
MOVIMIENTO	31'54	25'38	43'08

Tabla C.12: Clasificación para los tres estados usando "*shrinkage*"

Apéndice D

Código de Matlab

D.1. Introducción

En este apéndice se muestran cuáles son los archivos que han sido programados para realizar todo el proyecto: procesamiento de la señal y sus funciones auxiliares y por otro lado los clasificadores con los algoritmos que requerían.

D.2. Procesado de la señal

Para el procesamiento de la señal, se han codificado seis archivos, uno por cada uno de los sujetos, puesto que cada uno de ellos al final tenía unas características distintas. Todos ellos necesitaban como datos de entrada la señal de *EEG en crudo* suministrada por el Grupo de Robótica de la Universidad de Zaragoza. Estos datos tenían una longitud variable y un número de trials por cada sujeto, mientras que el número de canales siempre es el mismo: 28.

Lo primero que había que hacer es segmentar la señal en los tres estados relevantes: reposo, preparación y movimiento. Para ello, con las etiquetas temporales, se buscaba el primer punto desde donde se grababa y se cogían los 300 ms primeros, después se buscaba la etiqueta temporal marcada como "0" y se contaban los 300 ms previos como anticipación y los 300 milisegundos posteriores, incluyendo al 0, como ejecución del movimiento.

Ahora hay que elegir qué tipo de filtro queremos aplicar, esto se hace con la variable **tipo_filtro**: con valor "1" indicaremos que queremos aplicar el CAR, con valor "2" laplaciano y, por lo general, con un "3" seleccionamos un filtro bipolar. Cuando se aplica el bipolar necesitamos dos parámetros más: **canal_filtro_1** y **canal_filtro_2** que serán los

dos filtros bipolares que queremos aplicar a la señal.

La ejecución continuará mostrándonos la diferencia estadística r^2 del canal *C3* para cada pareja de señales y con la señal *raw* y una vez filtrada, pero sólo muestra el valor obtenido del **canal_filtro_1**. Después mostrará todos los canales y el coeficiente de determinación para todos ellos, con la señal pre-filtrado y post-filtrado.

Lo último que hacen las funciones de procesado de cada sujeto es salvar los datos de la señal filtrada en el disco, para así poder ser utilizado por las funciones de clasificación o por cualquier usuario futuro.

Durante el proyecto se debieron de probar uno a uno todos los filtros y comprobar en cuál de ellos se producía la diferencia significativa más grande en *C3*.

Los nombres de los procedimientos son: **PROCESADO_SUJETO_N**, donde *N* es el número del Sujeto.

D.2.1. Funciones Auxiliares

- **calcular_r2:** Esta función a partir de la señal de entrada y con el número de *trials* que forman la señal, calcula el coeficiente r^2 . Devuelve el valor de r^2 para toda la señal y además el valor máximo de cada canal, esto será muy útil para comparar el valor del coeficiente de determinación antes y después de aplicar los bipolares.
- **dibujar_r2:** Esta función muestra el valor del r^2 para un canal dado, además tiene en cuenta el valor máximo dado, por lo que sirve para comparar si ha habido mejora o no de las características en ese canal dado.
- **dibujar_canales:** Este procedimiento muestra todos los canales teniendo, en el caso de que la señal esté haya sido filtrada, muestra el valor del coeficiente de determinación con la escala de la señal en crudo.

D.3. Clasificadores

Para cada sujeto, se ha diseñado una función de clasificación distinta, pues cada uno de ellos tenía una selección de características totalmente distinto al resto. Una vez que

ya tenemos la señal filtrada, lo primero que hay que hacer es recuperar esos datos. Una vez que se tienen las señales filtradas, lo primero que se hace es mostrar la señal filtrada espacialmente, filtrada frecuencialmente, rectificada y suaviza, y por cada una de estos cuatro tipos de señales muestra el valor de r^2 comparado siempre con el valor de la señal filtrada espacialmente.

Una vez que hemos encontrado el tipo de señal que maximiza el valor del coeficiente de determinación en *C3* para cada pareja de estados, entonces nos quedamos con los canales que contienen la información más relevante para la clasificación, y después colocamos las características, tanto para los clasificadores por parejas, como para el clasificador de los tres estados simultáneamente.

Una vez que tenemos la señal ordenada de la manera idónea, separamos los datos en dos conjuntos: conjunto de "entrenamiento", con el 90 % de los datos, y con conjunto de "test", con el 10 % restante. Ahora calculamos la matriz de transformación de PCA con los datos de "entrenamiento" para reducir el número de dimensiones del problema, quedándonos con el 99 % de la variabilidad de la señal, aplicamos la matriz de transformación calculada a los datos de test.

Ahora procedemos a normalizar los datos de "entrenamiento" entre "0" y "1", y luego lo aplicamos al "test" teniendo en cuenta los valores del "entrenamiento". Con los datos normalizados, clasificaremos, para lo cual calcularemos la γ que necesita el shrinkage, como se puede ver en el **Anexo A.2.3** y entrenamos con los datos de "entrenamiento" los clasificadores y los probamos con los datos de "test". Al final del proceso se mostrarán las estadísticas de los dos tipos de clasificadores, tanto en porcentaje de aciertos en total, como por clases.

Los nombres de los procedimientos son siempre los mismos: **CLASIFICADOR_SUJETO_N**, dónde N es el número del sujeto.

D.3.1. Funciones Auxiliares

- **normalizacion**: Esta función con los datos de entrenamiento y prueba, los normaliza entre "0" y "1", tomando como máximo, el valor máximo del entrenamiento, y como valor mínimo, el mínimo del entrenamiento. Devuelve dos señales separadas **sentrenamiento** y **stest**.
- **calcular_GAMMA_PAPER**: Con el valor de la señal calcula el valor de γ como

se puede ver en la fórmula A.11.

- **clasificación_cuadratica**:: a partir del valor de la media de la clase dada, la covarianza de la señal, el punto del cuál queremos calcular el valor de su función discriminante y logaritmo neperiano del número de muestras de clase dada partido por el número total de elementos, esta función calcula el valor de la función discriminante para una clase.
- **shrinkage_general**:: esta función se encarga del grueso de la clasificación: tiene como entrada la señal de "entrenamiento" y sus etiquetas, la señal de "test" y sus etiquetas, la γ calculada anteriormente, el número de clases distintas que hay que clasificar, y una opción: "c" indica que queremos usar el "*shrinkage*" como aproximación entre LDA y QDA, "*Quadratic Discriminan Analysis*", o bien con "p", este es caso utilizado siempre, y que aplica "*shrinkage*" como compendio entre LDA o el uso de matrices de covarianza esféricas. Para clasificar con LDA, se pasa manualmente $\gamma = 0$.

Con los datos de "entrenamiento" calculamos la media de cada clase y la covarianza de todas. Para clasificar la función trabaja de cada manera, calcula para todos los puntos del "test", el valor de la función discriminante para cada clases, y asigna a ese punto la clase con el menor valor discriminante. Como tenemos los valores reales de cada "test", comparamos, y así obtenemos el porcentaje de aciertos globales. Esta función devuelve el porcentaje de aciertos y para cada dato del "test" devuelve la clase asignada por el clasificador.

- **desglose_estadisticas**: Esta función, con todos los datos de "test" clasificados y los reales, crea una matriz de dispersión con tantas filas y tantas columnas como clases distintas hay que clasificar.

Apéndice E

Artículo científico

Classification of Resting, Anticipation and Movement States in Self-Initiated Arm Movements for EEG Brain Computer Interfaces

Miguel Rodrigo, Luis Montesano, Javier Mínguez

Abstract—In the last years, there has been an increasing interest in using Brain Computer Interfaces (BCI) within motor rehabilitation therapies that use robotic devices or functional electro stimulation to help or guide the efforts of the patient to move her body. A crucial step of these therapies is to provide help to the user just when she is actually trying to accomplish a certain motion or task. One of the most promising applications of BCI systems in this context is its ability to measure the user intentions and actions to trigger the rehabilitation devices accordingly. This paper studies the single-trial classification based on EEG measurements of three basic states during the execution of self-initiated motion: rest, motion preparation (or anticipation) and motion. We conducted an experiment where the participants had to reach at their will eight different locations from a fixed starting position. Results for six different healthy subjects show that it is possible to achieve good classification rates given that features are carefully selected for each subject and for each pair of states.

I. INTRODUCTION

One of the application contexts of brain-computer interfaces is in robotic related rehabilitation programs, where they could provide a medium to decode the human intentions to modulate the robot intervention. For instance, this is the case in motor rehabilitation programs in stroke patients where it has been suggested that the brain-computer interface could play a key role by providing human cognitive aspects such as motor intention, perception of the feedback, and higher level motivational states as the attention [1]. These are important factors with a key role in any seamless integration between the human and the rehabilitation robot with potential to favour a natural compliance and adherence of the patient in the therapy process.

Many motor therapies are based on the development of repetitive analytical movements of the targeted limb. Within this movement process there are at least three macro-states of interest: the rest state, the preparation of motion, and the motion execution. The on-line monitoring with a brain-computer interface of these states could provide a valuable information to design any robot-mediated therapy for control (triggering of events), for monitoring of the process, and as a measure of compliance and engagement in fully automated therapies. The decoding of these three macro-states simultaneously from EEG has not been addressed yet and it is the objective of this paper. An important observation

is that although the rest and the movement states could be decoded from other sensory modalities such as the EMG, the preparation of motion is the only state with anticipatory information of the intention of the subject and can only be decoded from brain activity. In addition to this, the on-line decoding of this state is gaining attention in brain-computer interfaces since it is also present during motor imagery [2] (which is one of the key mental tasks for operation).

Although there is not research in decoding the three states simultaneously, there is significant research in differentiating between the rest state from the movement (and imagination of the movement), and in between the rest state and the motor preparation. It is well known that during the movement execution there is an event-related desynchronization in motor areas (12-15Hz) [3] and also there is slow cortical activity involved (motor-related potentials) [4]. Usually the sensory-motor synchronization/desynchronization is the key feature used to differentiate between the rest and movement states in real or motor imagery [5]. In addition to this, several studies have demonstrated the appearance of EEG activity preceding human voluntary movement. These signals are associated to motor task preparation and are dissimilar from those during the actual execution. Similar to the movement, there is the event-related (de)synchronization (ERD/ERS) and a slow activity denoted the readiness potential. Several studies have focused in distinguishing between the anticipation of movement of limbs [6], [7], [8], [9] in time or frequency domains. An important point of these last works that report decoding of anticipatory motor rhythms is that they are not self-paced and the subject is always cued. The contribution of the present paper is a simultaneous decoding of the three states in a self-paced context (i.e. the subjects decides when to start the motion and choose among different types of motion), which is the ecological working setting.

II. METHODOLOGY

A. Experiment

Seven right-handed male students of the University participated voluntarily in the experiments (age range 25–32 years) after the protocol was approved by the Institutional Review Board of the University of Zaragoza. Subjects were seated in a comfortable chair in front of the reaching apparatus (Figure 1a). They were instructed to move the right arm from the homing location to a self-chosen location among eight possible locations and then come back to the homing position. This process was self-paced approximately every 7.5 secs on average (minimum 2.8 and maximum 9.7 secs). During the reaching operation, subjects were asked to maintain a natural,

Luis Montesano and Javier Mínguez are with the Instituto de Investigación en Ingeniería de Aragón (I3A) and Dpto. de Informática e Ingeniería de Sistemas (DIIS), Universidad de Zaragoza, Spain. E-mail: montesano,jminguez@unizar.es. M. Rodrigo is with Bit&Brain Technologies, Zaragoza, Spain. E-mail: mrodrigo@bitbrain.es. This work has been partially supported by projects HYPER-CSD2009-00067 and DPI2009-14732-C02-01 funded by the Spanish Government.

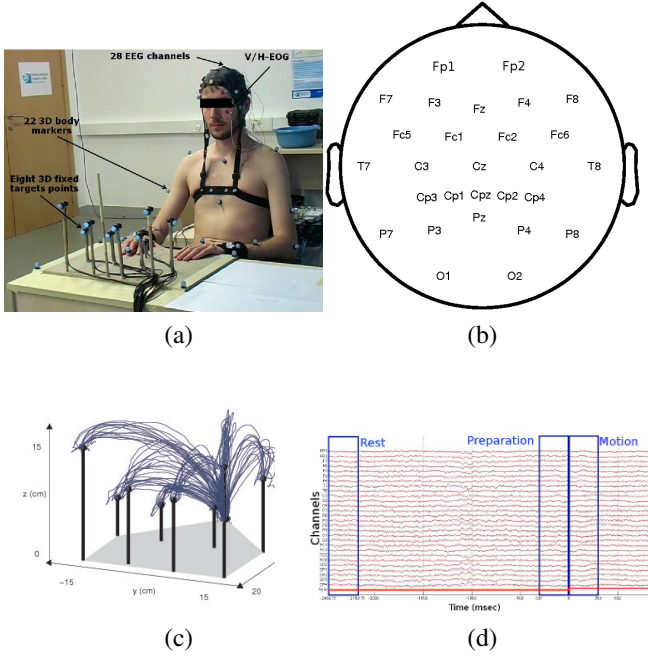


Fig. 1. (a) Experimental setup. The user sits in front of the reaching apparatus with a fixed start position and set of eight predefined positions. (b) Sensor placement. (c) Trajectories executed by the users. (d) Acquired signals including synchronization signal obtained from the buttons at the positions. The blue rectangles illustrate the extracted examples of each class (rest, preparation, motion) for each trial.

relaxed and constant posture; and to minimize blinking while maintaining the gaze fixed in the appropriate point of the apparatus. In the remaining time they were allowed to blink and relax. The first condition was executed in five blocks of 5 minutes each, where the subjects executed around 130 reaching operations to the eight locations (minimum after removing non valid examples 75 and maximum 196). Figure 1(c) shows the trajectories executed by subject 1.

B. Data recording

EEG activity was recorded with a gTec system (2 gUS-Bamp amplifiers synchronized) with 28 electrodes distributed according to the 10/10 international system ((Figure 1b)), with the ground on FPz and the reference on the left earlobe. Vertical and horizontal EOG were also recorded. The EEG and EOG signals were digitalized with a sampling frequency of 256Hz, power-line notch-filtered and lowpass-filtered at 60Hz.

In addition to EEG measurements, the system also recorded time markers corresponding to the beginning and end of the arm reaching motion. This information was obtained using buttons both at the home and the target positions that were synchronized with the EEG measurements. Based on the motion markers, each trial was segmented in three different parts corresponding to rest, preparation and motion conditions with a fixed duration of 0.3 seconds each (see Figure 1(d)). Thus, each trial provides one example of each class.

C. Feature extraction and classification

Let $X \in \mathcal{R}^{C \times T}$ correspond to the acquired EEG signal of a class example where C is the number of channels and T the number of samples. One of the key aspects to distinguish rest, preparation and motion conditions is to select features from the EEG measurements that have the relevant information to distinguish between the classes. According to previous studies, this information is the event-related desynchronization/synchronization (ERD/ERS) in motor areas (12-15Hz) [3]. There is also important activity in lower frequencies corresponding to slow cortical potentials involved in the motor preparation (motor-related potentials) [4]. In principle, this type of information should be discriminative enough to distinguish between the three classes of interest (rest, preparation and motion). Firstly, as mentioned in the introduction ERD/ERS has been used to distinguish between rest and either motion or preparation. Secondly, the slow potentials are different for the preparation (or anticipation) of the motion and the motion execution.

In addition to the raw data X , the feature selection process applied two different filters to the raw data. This pre-processing step is common to improve the classification performance. The filters were:

- Re-referencing filter: The EEG measurements are re-referenced to one specific channel. By repeating this process for all the channels, we obtain a total of C filters that generate signals $X_{Bi} \in \mathcal{R}^{C \times T}$, with $i \in 1..C$.
- Laplacian filter: The EEG signal X was filtered using a 3×3 discrete Laplace operator to obtain $X_L \in \mathcal{R}^{C \times T}$.

Once X, X_{Bi}, X_L , The selection of the best filter for each subject was based on the r^2 test of the spectral density between pairs of classes (i.e. rest vs. preparation, rest vs. motion and preparation vs. motion). For each trial and channel, we estimated the spectral density of the different filtered signals using the periodogram. Then, for each pair of classes we selected just one filtered signal among X, X_{Bi}, X_L based on the r^2 maximum value of channel C3 (which is one of the most representative electrodes related to motion of the right arm). The final selection was done by visual inspection based on the periodograms (see Section III). Based on r^2 values, we also selected the range of frequencies f_r of the periodogram with the highest value for the filtered channel C3. In addition to this channel, all the channels of the selected filter that were above a certain threshold were also included in the feature vector.

The last step before training the classifier consisted in processing the selected channels in the time domain. First, a Butterworth band-pass filter was applied in the selected range f_r . This filtered temporal signal was rectified and smoothed (which empirically has shown to be a good way to improve the stability of the signal and to increase the statistical difference between conditions). As in the frequency domain, an r^2 test was used to select the best signal among the four different types resulting from applying sequentially to the filtered signal, the band-pass filter, the rectification and the smoothing.

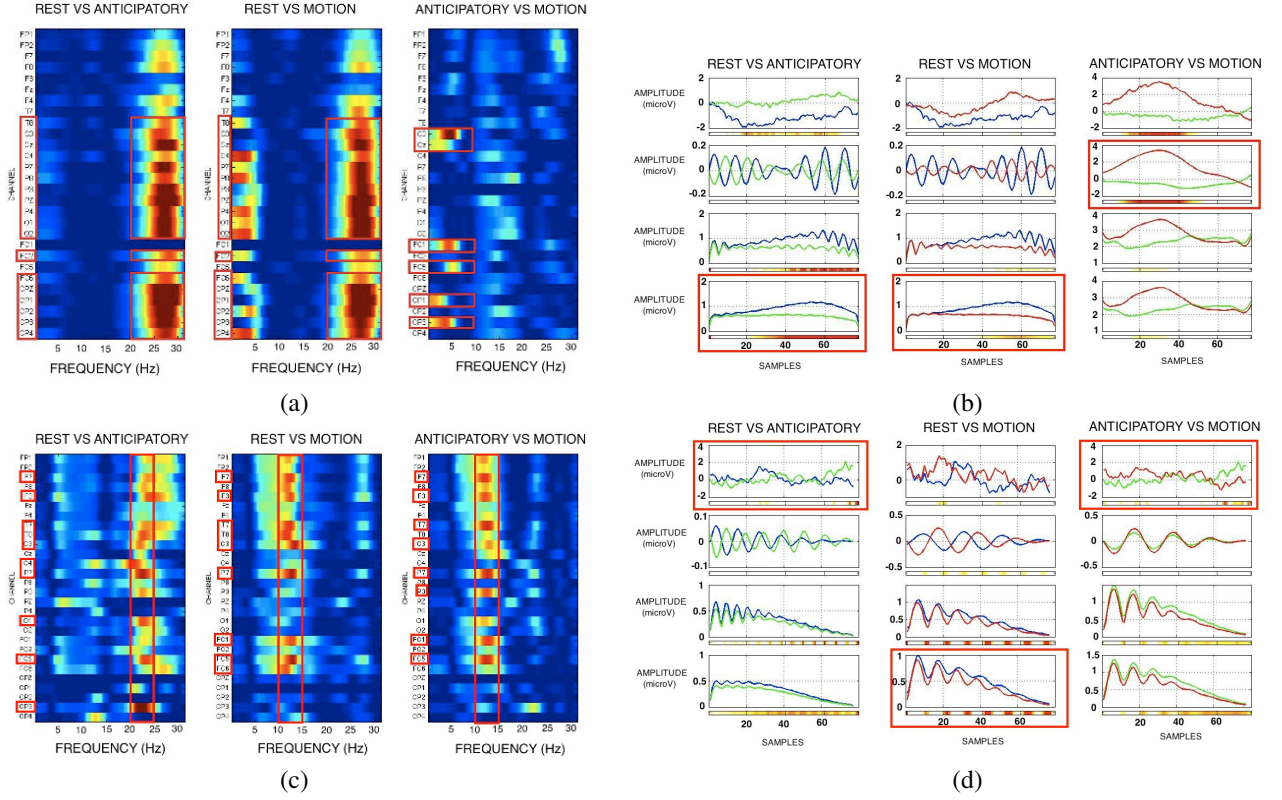


Fig. 2. Selected features. (a) r^2 -test values of the periodogram of the selected filters for each pair of classes (rest vs. motion, rest vs. preparation, preparation vs. motion) for subject one. The red boxes indicate the range of frequencies and channels used for classification. (b) Signals in the temporal domain for channel C3 of the filter selected for each pair of classes. Signals are averaged over all the examples of each class. From top to bottom, (i) selected features, (ii) filtered in the selected frequency band, (iii) rectified and (iv) smoothed. The corresponding r^2 -test is shown for each case. The red box indicates the final features fed to the classifier. (c) r^2 -test values of the periodogram for subject 4. (d) Final features fed to the classifier for subject 4.

The final selected features, subsampled to 64Hz, were used to train a (multi-class) LDA classifier. Since the feature vector dimension can be larger or close to the number of examples of each class, we used a shrinkage estimator for the covariance matrix [10]. When classifying between two classes, the feature vector was formed by the temporal signal of that specific pair (e.g. rest and motion). For the simultaneous classification of the three conditions, the feature vector was the concatenation of the filters selected for each pair of classes.

III. RESULTS

A. Selected Features

We first analysed the selected features for each user. Recall that filters were selected independently for each subject and for each pair of conditions. Table III-A shows the selected filters, where B_i indicates Bipolar with respect to channel i . The first remark is that in all cases (subjects and pairs of classes), bipolar filters were the ones with best r^2 compared to raw measurements X and to Laplacian filtered channels X_L . Second, for every pair of classes and every subject there were multiple filtered channels that contained discriminative information. The table shows that the best selected filters vary significantly among subjects as well as the corresponding frequency ranges.

Figures 2(a) and 2(c) show the r^2 -test values of the periodogram for the selected filters for the subjects one and four who had the largest and smallest statistical differences, respectively. Recall that they correspond to different filters as indicated in Table III-A. For the best subject, subject one, the information selected to separate rest and preparation corresponds to an ERD desynchronization in high beta bands (around 25 Hz) at the central and parietal zones. In the case of rest versus motion, the figure shows a similar pattern to the previous one plus some relevant statistical difference in low frequencies [0-5]Hz. The latter is probably due to slow potentials related to motion as described in [11]. Finally, the r^2 values between preparation and motion show a completely different structure where the most useful information appears in low frequencies around 5Hz. As in the previous case, these differences are probably caused by slow potentials such as the lateralized readiness potentials. Figure 2(b) show the averaged signal features in the time domain for channel C3.

The results for subject four shows that his best selected features have quite a different structure and indicated that the different conditions are not as different as those for subject one (see Fig. 2(c)). Furthermore, the channels and frequency bands are not as consistent with the literature as the ones from other subjects such as subject one. Consequently, the temporal signals shown in Fig. 2(d) are also less discriminative (as we will corroborate when analysing the classification

TABLE I

FILTERS AND FREQUENCY RANGE SELECTED FOR EACH SUBJECT AND PAIR OF CLASSES

Subj.	Rest VS Antici.	Rest VS Motion	Antici. VS Motion
1	B_{FC1} , [20-30]Hz	B_{FC1} , [20-30]Hz	B_{P3} , [0-10]Hz
2	B_{FZ} , [25-30]Hz	B_{FC2} , [10-20]Hz	B_{FC2} , [10-20]Hz
3	B_{CP3} , [20-30]Hz	B_{F7} , [0-5]Hz	B_{CP3} , [20-30]Hz
4	B_{CPz} , [20-25]Hz	B_{CP1} , [10-15]Hz	B_{CPz} , [10-15]Hz
5	B_{P4} , [10-15]Hz	B_{O1} , [15-25]Hz	B_{O1} , [17-20]Hz
6	B_{P8} , [5-10]Hz	B_{FP1} , [0-10]Hz	B_{FP1} , [0-10]Hz

TABLE II

CLASSIFICATION RESULTS FOR EACH PAIR OF CONDITIONS PER SUBJECT

Subject	Rest VS Antici.	Rest VS Motion	Antici. VS Motion
1	69.59	67.86	80.78
2	67.30	81.01	75.01
3	72.01	72.57	68.47
4	58.63	55.39	63.18
5	50.88	48.70	58.36
6	66.93	56.54	61.15
Mean	64.22	63.68	67.70

results).

B. Classification results

Here we present the classification results and discuss two different cases: (i) the classification of pairs of classes independently; and (ii) the joint classification of the three classes.

Table III-B summarizes the classification results of each pair of classes. The mean accuracy for each pair is around 64%. There is very little variation between the different pairs of classes. However, subjects behave differently with an average of 74% and 72% for the two best subjects (subject 2 and subject 1) and an average of 52% and 59% for the two worst cases (subject 5 and 4).

We now analyze the results of the LDA classifier in the multi-class case. Table III-B shows the confusion matrix averaged for the six subjects. The classification rates for each class are similar to the ones obtained in the previous case with a degradation of almost 10% on average from 65% in the two class case to 55% in the three class one. In the latter, the subject one is the best with a mean recognition rate of 72% (68%, 69%, 78% for rest, preparation and motion, respectively). The worse subject is subject 4 with a mean recognition rate of 40% (32%, 41%, 47% for each class). These results are as expected since the analysis of the features already revealed that there were important differences between the subjects in terms of statistical difference among the classes of the best features. Indeed, a deeper analysis of the EEG measurements of subjects 4 and 5 suggest that

TABLE III

MEAN CONFUSION MATRIX AVERAGED FOR ALL SUBJECTS

	Rest	Anticipation	Motion
Rest	54.9937	23.5145	21.4918
Anticipation	23.6410	54.7408	21.6182
Motion	22.6296	20.4804	56.8900

the ERD/ERS could not be clearly identified and there was no statistical difference in the frequency bands of the slow potentials in any of the filters. It is also worth to mention that those subjects that had the best classification rates (e.g. subjects 1 and 2) kept their performance in the three classes case, while those with the lower rates (e.g. subjects 4 and 5) degraded their performance considerably.

IV. CONCLUSIONS

This paper addresses the classification of the basic phases (rest, preparation and motion) involved in self-paced motion using EEG measurements obtained through a BCI system. The method explicitly builds and selects features for each pair of classes according to their statistical difference based on the χ^2 -test. The results show that it is possible to achieve reasonable classification rates around 70% by appropriately selecting the features for each of the six subjects.

Planned future work explores two different directions. First, we are currently working on more automatic feature selection processes that does not require any visual inspection by the therapist. Second, although motion is self-initiated by the user, classification still requires synchronized markers to segment the signal. We plan to investigate structured prediction techniques to label the full sequence as recorded from the BCI system.

REFERENCES

- [1] K. P. Tee, C. Guan, K. Keng Ang, K. Soon Phua, C. Wang, and H. Zhang, "Augmenting cognitive processes in robot-assisted motor rehabilitation," in *Proceedings of the 2nd Biennial IEEE/RAS-EMBS International Conference on Biomedical Robotics and Biomechatronics*, Scottsdale, USA, 2008.
- [2] Christa Neuper, Reinhold Scherer, Miriam Reiner, and Gert Pfurtscheller, "Imagery of motor actions: Differential effects of kinesthetic and visual-motor mode of imagery in single-trial eeg," *Cognitive Brain Research*, vol. 25, no. 3, pp. 668 – 677, 2005.
- [3] G. Pfurtscheller and F. Lopez da Silva (Eds.), "Event-related synchronization," in *Handbook of Electroencephalography and Clinical Neurophysiology*, 1999.
- [4] M. Kutas and E. Donchin, "Preparation to respond as manifested by movement-related brain potentials," in *Brain Research*, 1980, number 95.
- [5] C. Neuper and G. Pfurtscheller, "Event-related dynamics of cortical rhythms: frequency-specific features and functional correlates," in *International Journal of Psychophysiology*, 2001, pp. 41–58.
- [6] Valerie Morash, Ou Bai, Stephen Furlani, Peter Lin, and Mark Hallett, "Classifying eeg signals preceding right hand, left hand, tongue, and right foot movements and motor imagines," *Clinical Neurophysiology*, vol. 119, no. 11, pp. 2570 – 2578, 2008.
- [7] B.O. Peters, G. Pfurtscheller, and G. Edlinger, "Automatic differentiation of multichannel eeg signals," *IEEE Transactions on Biomedical Engineering*, vol. 48, no. 1, 2001.
- [8] B. Blankertz, G. Dornhege, C. Schafer, R. Krepki, J. Kohlmorgen, K.R. Muller, F. Losch V. Kunzmann, and G. Curio, "Boosting bit rates and error detection for the classification of fast-paced motor commands based on single-trial eeg analysis," *IEEE Transactions Neural Systems Rehabilitation Engineering*, vol. 11, no. 127, 2003.
- [9] M. Krauledat, G. Dornhege, B. Blankertz, F. Losch, G. Curio, and K.-R. Miller, "Improving speed and accuracy of brain-computer interfaces using readiness potential features," *Engineering in Medicine and Biology Society*, pp. 4511–4515, 2004.
- [10] J.B. Copas, "Regression, prediction and shrinkage," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 45, no. 3, pp. 311–354, 1983.
- [11] S. Waldert, H. Preissl, E. Demandt, C. Braun, N. Birbaumer, A. Aertsen, and C. Mehring, "Hand movement direction decoded from meg and eeg," *J. Neuroscience*, vol. 28, no. 4, pp. 1000–1008, 2008.

Bibliografía

- [1] K. P. Tee, C. Guan, K. Keng Ang, K. Soon Phua, C. Wang, H. Zhang. "*Augmenting cognitive processes in robot-assisted motor rehabilitation*", *Proceedings of the 2nd Biennial IEEE/RAS-EMBS International Conference on Biomedical Robotics and Biomechatronics*, Scottsdale, USA, 2008.
- [2] G. Pfurtscheller, F. Lopez da Silva. "*Event-related synchronization*", *Handbook of Electroencephalography and Clinical Neurophysiology*, 1999
- [3] M. Kutas, E. Donchin. "*Preparation to respond as manifested by movement-related brain potentials*", *Brain Research*, 1980, número 95
- [4] C. Nauper, G. Pfurtscheller. "*Event-related dynamics of cortical rhythms: frequency-specific features and functional correlates*", *International Journal of Psychophysiology*, 2001, págs. 41-58
- [5] Valerie Morash, Ou Bai, Stephen Furlani, Peter Lin, Mark Hallet. "*Classifying eeg signals preceding right hand, left hand, tongue and right foot movements and motor imageries*", *Clinical Neurophysiology*, vol. 119, no. 11, págs. 2570 - 2578, 2008
- [6] B. O. Peters, G. Pfurtscheller, G. Edlinger. "*Automatic differentiation of multichannel eeg signals*", *IEEE Transactions on Biomedical Engineering*, vol. 48, no. 1, 2001
- [7] B. Blankertz, G. Dornhege, C. Schafer, R. Krepi, J. Kohlmorgen, K. R. Muller, F. Losch, V. Kunzmann, G. Curio. "*Boosting bit rates and error detection for the classification of fast-paced motor commands based on single-trial eeg analysis*", *IEEE Transactions Neural Systems Rehabilitation Engineering*, vol. 11, no. 127, 2003
- [8] M. Krauledat, G. Dornhege, B. Blankertz, F. Losch, G. Curio, K. R. Muller. "*Improving speed and accuracy of brain-computer interfaces using readiness potential features*", *Engineering in Medicine and Biology Society*, págs. 4511 - 4515, 2004
- [9] Everitt, B. S. "*Cambridge Dictionary of Statistics*", *Cambridge University Press*, 2^a ed., 2002

-
- [10] Benjamin Blankertz, Ryota Tomioka, Steven Lemm, Motoaki Kawanabe, Klaus-Robert Müller. *"Optimizing Spatial Filters for Robust EEG Single-Trial Analysis"*, *IEEE signal processing magazine*, Vol. XX, 2008
- [11] Paul L. Nunez, Ramesh Srinivasan, Andrew F. Westdorp, Ranjith S. Wijesinghe, Don M. Tucker, Richard B. Silberstein, Peter J. Cadusch *"EEG coherency I: statistics, reference electrode, volume conduction, Laplacians, cortical imaging, and interpretation at multiple scales"*, *Electroencephalography and clinical Neurophysiology*, no. 103, págs. 499 - 515
- [12] Herbert Ramoser, Johannes Müller-Gerking, and Gert Pfurtscheller. *"Optimal Spatial Filtering of Single Trial EEG During Imagined Hand Movement"*, *IEEE Transactions on rehabilitation engineering*, vol. 8, no. 4, 2000
- [13] G. Pfurtscheller, C. Neuper *"Event-related synchronization of mu rhythm in the EEG over the cortical hand area in man"*, *Neuroscience Letters*, vol. 147, no. 1, págs. 93 - 96, 1994
- [14] S. Waldert, H. Preissl, E. Demandt, C. Braun, N. Birbaumer, A. Aertsen, C. Mehring. *"Hand movement direction decoded from meg and eeg"*, *J. Neuroscience*, vol. 28, no. 4, págs. 1000 - 1008, 2008
- [15] Carmen Vidaurre. *"Clasificación de componentes frecuenciales"*, Machine Learning Laboratory, Berlin Institute of Technology.
- [16] Benjamin Blankertz, Steven Lemm, Matthias Treder, Stefan Haufe, Klaus-Robert Müller. *"Single-trial analysis and classification of ERP components - A tutorial"*, *NeuroImage*. 2010
- [17] J. B. Copas *"Regression , prediction and shrinkage"*, *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 45, no. 3, págs. 311 - 354, 1983
- [18] Jerome H. Friedman. *"Regularized Discriminant Analysis"*, *Journal of the American Statistical Association*, vol. 84, no. 405, págs. 165 - 175, 1989.
- [19] Jia Li. *"Linear Discriminant Analysis"*, Department of Statistics, The Pennsylvania State University.
- [20] Jia Li. *"Regularized Discriminant Analysis and Reduced-Rank LDA"*, Department of Statistics, The Pennsylvania State University.
- [21] Christopher M. Bishop. *"Machine Learning and Pattern Recognition"*, Ed. Springer, 2ª ed., págs 101-135

-
- [22] Julianne Schäfer, Korbinian Strimmer. "*A Shrinkage Approach to Large-Scale Covariance Matrix Estimation and Implications for Functional Genomics*", *Statistical Applications in Genetics and Molecular Biology*, vol. 4, art. 32, 2005.
- [23] Eduardo López-Larraz, Iñaki Iturrate, Luis Montesano, Javier Mínguez "*Real-Time Recognition of Feedback Error-Related Potentials during a Time-Estimation Task*", *International Conference of the IEEE Engineering in Medicine and Biology Society*, 2010.