



INAOE

Representación, reconocimiento y
clasificación de la actividad de
imaginación motora en el
dominio tiempo-frecuencia

por

Obed Carrera León

Tesis presentada en el Instituto Nacional
de Astrofísica, Óptica y Electrónica
para obtener el grado de Maestro en Ciencias
en el departamento de Electrónica.

Asesores:

Dr. Juan Manuel Ramírez Cortes
Dr. Vicente Alarcón Aquino

Sta. Ma. Tonantzintla, Pue.
Agosto, 2011

©INAOE, 2011

Derechos Reservados

El autor otorga al INAOE el permiso de
reproducir y distribuir copias en su totalidad
o parcial de esta tesis.



Instituto Nacional de Astrofísica Óptica y Electrónica

Tonanzintla, Puebla, México



Obed Carrera León

***Representación, Reconocimiento y clasificación de la actividad de
imaginación motora en el dominio tiempo-frecuencia***

Asesores:

Dr. Juan Manuel Ramírez Cortes

Dr. Vicente Alarcón Aquino.

Agosto de 2011

Resumen

Una interface cerebro computadora es un sistema que permite el control de dispositivos electrónicos mediante el análisis de los patrones presentes en las señales cerebrales. Esta puede ser de gran ayuda a personas con discapacidades motoras. En el trabajo reportado en esta tesis se presentan dos algoritmos de extracción de características para la actividad de imaginación motora.

El primer algoritmo está basado en la transformada Hilbert, la cual permite establecer una representación tiempo frecuencia, que permite visualizar el fenómeno ERD/ERS presente en la señal y que a diferencia de algunos autores, no se realiza ningún tipo de transformación adicional a la información extraída. Esta representación permite identificar la frecuencia reactiva por sujeto y seleccionar la banda de frecuencias adecuada para poder generar un vector de características.

El segundo es un algoritmo de extracción de características basado en wavelets. Se descompone la señal en tres niveles utilizando tres diferentes wavelets, cada una con tres diferentes momentos de desvanecimiento diferentes. Luego se seleccionan las correspondientes bandas de frecuencias y el vector de características es construido con tres diferentes características estadísticas extraídas de los niveles seleccionados.

La clasificación de los patrones extraídos se hace por los métodos de: discriminantes lineales por Fisher (FLDA), discriminantes cuadráticos (QDA) y una máquina de soporte vectorial (SVM). El desempeño del algoritmo es evaluado por una validación cruzada por diez sobre tres bases de datos descargados de la página del BCI Competition y una base de datos obtenida mediante un experimento propio. En total los algoritmos se probaron sobre nueve sujetos diferentes.

En general los resultados son favorables de acuerdo a los métodos propuestos. Se plantea una extracción de características basada en la transformada Hilbert, la cual es sencilla y rápida para su aplicación on-line y no depende de transformaciones adicionales. Para el método basado en wavelet se demuestra que al usar tres características estadísticas se obtiene más información de la señal que a su vez mejora los resultados, obteniendo tasas de error de hasta 0.16 y 0.8

con cada método; siendo el análisis wavelet el de menor error, en comparación con el trabajo presentado por otros autores.

ÍNDICE

ÍNDICE DE FIGURAS	VI
ÍNDICE DE TABLAS.....	IX
1. Introducción	1
1.1. Elección de la actividad cerebral.....	3
1.2. Event Related Synchronization/Desynchronization (ERS/ERD).....	3
1.3. Objetivos	4
1.3.1. General.....	4
1.3.2. Específicos.....	4
1.4. Organización de la tesis	4
2. Bases Tiempo-Frecuencia.....	6
2.1. Introducción	6
2.2. Transformada de Fourier por ventanas (STFT)	6
2.3. Transformada Wavelet	7
2.3.1. Parámetros de escala-tiempo.....	9
2.3.2. Análisis multiresolución	10
2.3.3. Funciones de escalamiento y funciones wavelet	11
2.4. Transformada discreta wavelet	13
2.5. Bancos de filtros y la transformada wavelet	13
2.6. Transformada Hilbert.....	16
2.6.1. Señal analítica	16
2.7. Representación Tiempo-frecuencia	17
2.8. Sumario	20
3. Clasificadores.....	21
3.1. Análisis de Discriminantes Lineales por Fisher (LDA).....	21
3.2. Análisis de Discriminantes Cuadráticos (QDA).....	22
3.3. Máquina de Soporte Vectorial (SVM)	23
3.3.1. SVM y Kernels	25
3.4. Sumario	26

4. Método Propuesto: Extracción y Clasificación de Patrones	27
4.1. Grupos de datos (<i>Data sets</i>).....	27
4.1.1. Grupos de datos de BCI competition.....	27
4.1.2. Grupo de datos propios	28
4.1.2.1. Experimento.....	28
4.1.2.2. Procesamiento de la señal	29
4.2. Extracción de características por medio de <i>Patrones Espaciales</i>	30
4.2.1. Vector de características.....	33
4.3. Extracción de características usando <i>wavelets</i>	33
4.3.1. Vector de características.....	35
4.4. Clasificación	35
4.4.1. Parámetros.....	35
4.4.2. Clasificación: patrones espaciales.....	36
4.4.3. Clasificación: Características Wavelet	36
4.5. Sumario	37
5. Pruebas y Resultados.....	38
5.1. <i>Patrones Espaciales</i>	38
5.2. <i>Características wavelet</i>	39
5.3. Sumario	42
6. Conclusiones y Trabajo a Futuro	43
Referencias	44
A. Apéndice A: Figuras	48
A1: <i>Patrones espaciales</i>	48
A2: <i>Wavelet Features</i>	53

ÍNDICE DE FIGURAS

Fig. 2.1. Regiones de Heisenberg para la STFT [22].	7
Fig. 2.2. Funciones de escalamiento y espacios vectoriales wavelets [33].	12
Fig. 2.3. Banco de análisis de dos bandas.	14
Fig. 2.4. Árbol de análisis de dos etapas y dos bandas.	15
Fig. 2.5. Filtrado para tres etapas.	15
Fig. 2.6. Banco de filtros butterworth. Bw = 3 Hz. Traslape de 50 %.	18
Fig. 2.7. Proceso de extracción de envolventes de la señal para la obtención de la TFR y el fenómeno ERD – ERS. Bw = 4 Hz.	18
Fig. 2.8. (a) Espectrograma calculado con STFT. (b) Representación tiempo-frecuencia por medio de patrones espaciales. (c) ERD-ERS extraídos de (b).	19
Fig. 2.9. Patrones espaciales promediados de las primeras 20 muestras de S1. Solo se toman las frecuencias de 7 a 14 Hz debido a que se encontró que las frecuencias reactivas se concentran en ese rango. Las líneas punteadas marcan el momento donde se muestran el estímulo del movimiento a imaginar. A partir del segundo tres se observan variaciones en la energía de la señal.	19
Fig. 3.1. Función de decisión cuadrática. Se forma una región separación para las dos clases, la cual no está directamente conectada [15].	23
Fig. 3.2. Hiperplano de separación óptimo para un espacio bi-dimensional.	24
Fig. 4.1. Línea de tiempo del estímulo diseñado para la actividad de imaginación motora.	29
Fig. 4.2. Diagrama de ubicación de los sensores en la red EGI con una distribución 10 – 10.	29
Fig. 4.3. Representación tiempo-frecuencia para S1. Los mapas son calculados para cada prueba y electrodo y al final se obtiene la media sobre todas las pruebas para obtener los mapas de referencia para cada electrodo. La línea punteada indica el tiempo en que se indica el movimiento.	31
Fig. 4.4. Vectores de características ERD/ERS. Se muestra el comportamiento para cada clase y ambos electrodos por clase.	32

Fig. 4.5. Vectores de características obtenidos al restar los electrodos C3 y C4 para cada clase. Se observa la diferencia de patrones ERD/ERS entre clases.	32
Fig. 4.6. Vectores de características antes y después de la reducción.	33
Fig. 5.1. Promedio de error de todos los sujetos para los diferentes clasificadores sobre todas la wavelets.	41
Fig. A.1. Vector de características extraído por medio de patrones espaciales para S1.....	48
Fig. A.2. Vector de características extraído por medio de patrones espaciales para S2.....	49
Fig. A.3. Vector de características extraído por medio de patrones espaciales para S3.....	49
Fig. A.4. Vector de características extraído por medio de patrones espaciales para S4.....	50
Fig. A.5. Vector de características extraído por medio de patrones espaciales para S5.....	50
Fig. A.6. Vector de características extraído por medio de patrones espaciales para S6.....	51
Fig. A.7. Vector de características extraído por medio de patrones espaciales para S7.....	51
Fig. A.8. Vector de características extraído por medio de patrones espaciales para S8.....	52
Fig. A.9. Vector de características extraído por medio de patrones espaciales para S9.....	52
Fig. A.10. Error obtenido en el tiempo para S1 con coif2.....	53
Fig. A.11. Error obtenido en el tiempo para S1 con coif3.....	54
Fig. A.12. Error obtenido en el tiempo para S1 con coif4.....	54
Fig. A.13. Error obtenido en el tiempo para S1 con db3.	55
Fig. A.14. Error obtenido en el tiempo para S1 con db4.	55
Fig. A.15. Error obtenido en el tiempo para S1 con db5.	56
Fig. A.16. Error obtenido en el tiempo para S1 con sym4.	56
Fig. A.17. Error obtenido en el tiempo para S1 con sym5.	57
Fig. A.18. Error obtenido en el tiempo para S1 con sym6.	57
Fig. A.19. Error obtenido en el tiempo para S2 con coif2.....	58
Fig. A.20. Error obtenido en el tiempo para S2 con coif3.....	58
Fig. A.21. Error obtenido en el tiempo para S2 con coif4.....	59
Fig. A.22. Error obtenido en el tiempo para S2 con db3.	59
Fig. A.23. Error obtenido en el tiempo para S2 con db4.	60
Fig. A.24. Error obtenido en el tiempo para S2 con db5.	60

Fig. A.25. Error obtenido en el tiempo para S2 con sym4.	61
Fig. A.26. Error obtenido en el tiempo para S2 con sym5.	61
Fig. A.27. Error obtenido en el tiempo para S2 con sym6.	62
Fig. A.28. Error obtenido en el tiempo para S5 con coif2.	62
Fig. A.29. Error obtenido en el tiempo para S5 con coif3.	63
Fig. A.30. Error obtenido en el tiempo para S5 con coif4.	63
Fig. A.31. Error obtenido en el tiempo para S5 con db3.	64
Fig. A.32. Error obtenido en el tiempo para S5 con db4.	64
Fig. A.33. Error obtenido en el tiempo para S5 con db5.	65
Fig. A.34. Error obtenido en el tiempo para S5 con sym4.	65
Fig. A.35. Error obtenido en el tiempo para S5 con sym5.	66
Fig. A.36. Error obtenido en el tiempo para S5 con sym6.	66
Fig. A.37. Error obtenido en el tiempo para S7 con coif2.	67
Fig. A.38. Error obtenido en el tiempo para S7 con coif3.	67
Fig. A.39. Error obtenido en el tiempo para S7 con coif4.	68
Fig. A.40. Error obtenido en el tiempo para S7 con db3.	68
Fig. A.41. Error obtenido en el tiempo para S7 con db4.	69
Fig. A.42. Error obtenido en el tiempo para S7 con db5.	69
Fig. A.43. Error obtenido en el tiempo para S7 con sym4.	70
Fig. A.44. Error obtenido en el tiempo para S7 con sym5.	70
Fig. A.45. Error obtenido en el tiempo para S7 con sym6.	71

ÍNDICE DE TABLAS

Tabla 1-1. Clasificación de los mecanismos o proceso neurológicos empleados para BCI.....	2
Tabla 3-1. Descripción de tres de las kernels más utilizadas para la transformación de espacios de entrada [35].	26
Tabla 4-1. Descripción de los Grupos de datos obtenidos del BCI competition.	27
Tabla 4-2. Grupo de datos propios extraídos en Texas Tech University.	30
Tabla 4-3. Frecuencias correspondientes a los primeros cuatro niveles de descomposición para una señal de 128 Hz y 250 Hz.....	34
Tabla 5-1. Errores de clasificación de las características de patrones espaciales.	38
Tabla 5-2. Tasa de error obtenida para la base de datos 1. Para diferentes tipos de wavelet y clasificadores.	39
Tabla 5-3. Tasa de error obtenida para la base de datos 2. Para diferentes tipos de wavelet y clasificadores.	40
Tabla 5-4. Tasa de error obtenida para la base de datos 3. Para diferentes tipos de wavelet y clasificadores.	40
Tabla 5-5. Tasa de error obtenida para la base de datos propia. Para diferentes tipos de wavelet y clasificadores.....	41

1. INTRODUCCIÓN

El cerebro humano es considerado como la estructura biológica más compleja que existe. En años recientes, avances tecnológicos han permitido la medición de los campos magnéticos y eléctricos producidos por las neuronas que son las que procesan y transmiten la información dentro del cerebro [1]. La actividad cerebral puede observarse por medio de un electroencefalograma (EEG), el cual es generado por inhibición y excitación de potenciales pos-sinápticos de las células nerviosas corticales. Estos potenciales se suman en la corteza y se extienden hacia la superficie del cráneo donde son grabados como el EEG [2].

Los sistemas BCI permiten a una persona enviar comandos a un dispositivo electrónico por medio de la actividad cerebral, en consecuencia, no exigen ninguna actividad muscular por parte del usuario [1]. El propósito de un sistema BCI es permitir a un individuo con discapacidades motoras controlar de forma efectiva dispositivos como computadoras, procesadores de lenguaje, prótesis, silla de ruedas, etc. Un sistema BCI detecta la presencia de patrones específicos en la actividad proveniente del cerebro y que están relacionados con la intención de la persona de realizar alguna actividad [3].

Es posible clasificar las actividades cerebrales que se usan en el diseño de los sistemas BCI. En los sistemas BCI, las fuentes electrofisiológicas se refieren a los mecanismos neurológicos o procesos empleados por un usuario para generar las fuentes de control. Los sistemas BCI actuales caen dentro de siete categorías, basadas en los neuro-mecanismos y las técnicas empleadas para la obtención de estas señales. La tabla 1 muestra estas categorías.

Para poder controlar al sistema el individuo debe producir diferentes patrones de actividad cerebral que serán identificados y traducidos en comandos de control. En la mayoría de las BCI, esta identificación se basa en algoritmos de clasificación [4], por ejemplo, un algoritmo que tenga como objetivo la estimación automática de la clase de un dato representado por un vector de características. Estos algoritmos se usan para identificar “patrones” de la actividad cerebral. El desempeño del sistema de reconocimiento de patrones depende tanto de las características como del algoritmo de clasificación empleado [3].

Tabla 1-1. Clasificación de los mecanismos o proceso neurológicos empleados para BCI.

Neuromecanismo	Descripción
Actividad sensorial motora	<p>Cambios en los ritmos del cerebro (μ, beta y gamma)</p> <ul style="list-style-type: none"> • Cuando se da un movimiento voluntario se presentan cambios en las frecuencias μ y beta. Desincronización relacionada a eventos (ERD) y Sincronización relacionada a eventos (ERS) [5], [16], [23]-[26]. <p>Potenciales relacionados al movimiento (MRP)</p> <ul style="list-style-type: none"> • Son movimientos de baja frecuencia que se presentan 1 – 1.5 s antes del movimiento.
Potenciales corticales lentos	Cambios en el potencial provocados por el individuo cuando no hay movimientos.
P300	Pico que se presenta 300 ms después de que se responde a un estímulo visual, auditivo o somatosensorial cuando se intercala con alguna actividad de rutina (parietal) [6].
Potenciales evocados visuales (VEPs)	Provocado por estímulos visuales (occipital).
Respuesta a tareas mentales	Tareas mentales que no implican movimiento.
Actividad de las células neuronales (ANC)	Movimientos ejecutados en la dirección preferida de las neuronas.
Neuromecanismos múltiples (MNs)	BCIs basados en la combinación de dos o más de los anteriores.

Es esencial conocer y entender que tipos características se extraen para poder seleccionar el clasificador más apropiado para un sistema BCI dado. Se ha probado con una gran variedad de características para el diseño de una BCI; como valores de amplitud de los EEG [7], bandas de potencia (BP) [8], densidad espectral de potencia (PSD) [9], [10], parámetros autoregresivos (AR) y autoregresivos adaptables (AAR) [11], características de tiempo-frecuencia [12], correlación cruzada, etc. Algunas propiedades críticas de estas características que deben ser consideradas son:

- Ruido y valores atípicos
- Alta dimensionalidad
- Información en el tiempo
- No estacionarios
- Conjuntos de entrenamiento pequeños

Estas propiedades son válidas para la mayoría de las características actualmente usadas en la investigación de BCI. Sin embargo, pueden ser diferentes para BCIs fuera de los laboratorios [1].

1.1. Elección de la actividad cerebral

Al observar la literatura relacionada con sistemas BCI en los cuales se aplican técnicas de tiempo-frecuencia para la extracción de características y clasificación de las señales, la actividad cerebral sensorial motora es la que más se ha estudiado. Dentro de la actividad sensorial motora se encuentra la llamada “motor imagery” (Imaginación Motora) que es la relacionada con la ejecución de movimientos sin mover la extremidad físicamente (brazos, pies, dedos, etc.).

La importancia de analizar esta actividad recae en el hecho de que las personas con deficiencias motrices serias no cuentan con la posibilidad de moverse, y lo único que pueden realizar es pensar en mover una mano o un pie. Al poder caracterizar estas señales relacionadas con este tipo de actividad sería posible controlar una prótesis, una silla de ruedas, o cualquier dispositivo que sirva de ayuda a personas con discapacidades. Por lo anterior, se ha decidido trabajar con este tipo de señales.

1.2. Event Related Synchronization/Desynchronization (ERS/ERD)

Existen varios fenómenos presentes en las señales EEG. En la Imaginación motora se puede observar un fenómeno en particular que resulta en la alteración de los componentes rítmicos del EEG. Este está relacionado con los cambios en la actividad de poblaciones de neuronas en el lóbulo parietal, estos fenómenos son llamados potenciales relacionados al evento (ERP). Estos cambios en la actividad están fijos en tiempo pero no en fase, por lo que se requiere de métodos en tiempo-frecuencia para detectarlos; esto debido a que los ERPs varían con respecto al tiempo de evolución de la actividad [16], [23]-[26].

Este fenómeno representa cambios en frecuencia específicos en la señal de EEG y consisten, en términos generales, en decrementos o incrementos de potencia a ciertas bandas de frecuencia.

El primer caso es llamado “*Event Related Desynchronization*” o ERD, y el segundo “*Event Related Synchronization*” o ERS [16], [19].

1.3. Objetivos

1.3.1. General

Extraer y clasificar características de señales de EEG para la actividad de imaginación motora empleando técnicas de tiempo-frecuencia para dos clases.

1.3.2. Específicos

- Diseñar e implementar un experimento para la obtención de señales EEG y obtener una base de datos propia de señales de imaginación motora.
- Extraer características usando diferentes técnicas de tiempo-frecuencia: Transformada discreta wavelet, patrones espaciales.
- Con las características obtenidas, implementar los algoritmos de clasificación: LDA, QDA y SVM.
- Evaluar el desempeño del algoritmo diseñado por medio validación cruzada.

1.4. Organización de la tesis

El trabajo reportado en esta tesis está organizado en seis capítulos y un apéndice. En el Capítulo 2 se hace una revisión de técnicas de tiempo-frecuencia empezando por la transformada de Fourier para establecer un punto de comparación con las técnicas posteriores. Se presenta la transformada wavelet partiendo de la wavelet continua y sus características en frecuencia y describe su forma discreta. A continuación se plantea el concepto de expansión en sub-espacios. Se definen las funciones de escalamiento y de detalle y de manera conjunta se muestra la representación de estas funciones como bancos de filtros para señales de una dimensión. Así mismo se hace una breve descripción de la transformada Hilbert y se describe una representación tiempo frecuencia partiendo del concepto anterior, con la cual, se puede observar el efecto ERD/ERS.

En el Capítulo 3 se resume la teoría relacionada a las diferentes técnicas de clasificación utilizadas. Se parte del análisis de discriminantes lineales (LDA) para seguir con su forma cuadrática. Por último se aborda la técnica de máquina de vectores de soporte (SVM) desde la función discriminante básica hasta su extensión al uso de kernels para la transformación del espacio de entrada.

En el Capítulo 4 se describen los dos algoritmos de extracción y se describe cómo y bajo qué condiciones son desarrolladas las pruebas realizadas con el algoritmo y la métrica utilizada para medir la calidad de los resultados, así como los parámetros utilizados para la clasificación los patrones extraídos.

En el Capítulo 5 se analizan discuten los resultados obtenidos.

Por último, se hacen las conclusiones pertinentes partiendo del análisis de los resultados y se presentan ideas para trabajos a futuro.

En el Apéndice A se presentan una serie de graficas de salida de cada algoritmo de clasificación. Para el algoritmo de patrones espaciales muestran la distribución del vector referencia obtenido. Para el algoritmo basado en wavelets, las gráficas obtenidas como salida de la clasificación.

2. BASES TIEMPO-FRECUENCIA

2.1. Introducción

Cuando se habla de señales no estacionarias (EEG) se entiende que son fenómenos cuyas propiedades espectrales varían con respecto al tiempo. Estas propiedades pueden ser reveladas por transformadas que descomponen la señal sobre un conjunto de funciones base o elementales que están bien delimitadas en tiempo y frecuencia. La transformada de tiempo corto de Fourier (STFT, por sus siglas en ingles) y la transformada Wavelet son dos clases importantes de descomposiciones en tiempo-frecuencia [22]. También se presenta una forma alternativa de extracción de patrones basada en la metodología de Wang *et. Al* [12] y Wang y He [17].

2.2. Transformada de Fourier por ventanas (STFT)

Una transformada lineal busca la de componentes de la señal con un conjunto de formas de ondas concentradas en tiempo y frecuencia supongamos que ese conjunto de funciones es representado por $\phi \in \mathbf{L}^2(\mathbb{R})$. En el caso de la transformada de Fourier ϕ se define como,

$$\phi = g_{u,\xi}(t) = e^{i\xi t} g(t - u). \quad (2-1)$$

donde una ventana real y simétrica $g(t) = g(-t)$ es trasladada por u y modulada por la frecuencia ξ , por lo que la transformada de $f \in \mathbf{L}^2(\mathbb{R})$ es

$$Sf(u, \xi) = \langle f, g_{u,\xi} \rangle = \int_{-\infty}^{\infty} f(t) g(t - u) e^{-i\xi t} dt. \quad (2-2)$$

Con la STFT se puede definir la densidad espectral de potencia llamada espectrograma, descrito por P_s :

$$P_s f(u, \xi) = |Sf(u, \xi)|^2 = \left| \int_{-\infty}^{\infty} f(t) g(t - u) e^{-i\xi t} dt \right|^2. \quad (2-3)$$

El espectrograma mide la energía de f en el área tiempo-frecuencia delimitada por (u, ξ) delimitada por el principio de incertidumbre de Heisenberg o regiones de Heisenberg $g_{u,\xi}$, que se detalla en [22], En la Fig. 2.1 se muestran dos regiones tiempo-frecuencia; como la ventana que se utiliza es de un tamaño definido, se tiene la misma resolución para el plano tiempo-frecuencia. Lo anterior es una de las desventajas de la STFT; la selección de una ventana que permita obtener la información necesaria de una señal altamente no estacionaria, podría no dar información útil para su análisis [22], [13].

La STFT discreta sigue la idea de la transformada rápida de Fourier (FFT) y es,

$$Sf(m, l) = \langle f, g_{m,l} \rangle = \sum_{n=0}^{N-1} f[n]g[n - m] \exp\left(-\frac{i2\pi ln}{N}\right). \quad (2-4)$$

donde $0 \leq m < N$ es el desplazamiento de la ventana utilizada, $0 \leq l < N$ la frecuencia a la cual es calculada el coeficiente n , y N es la frecuencia de muestreo o periodo de la señal.

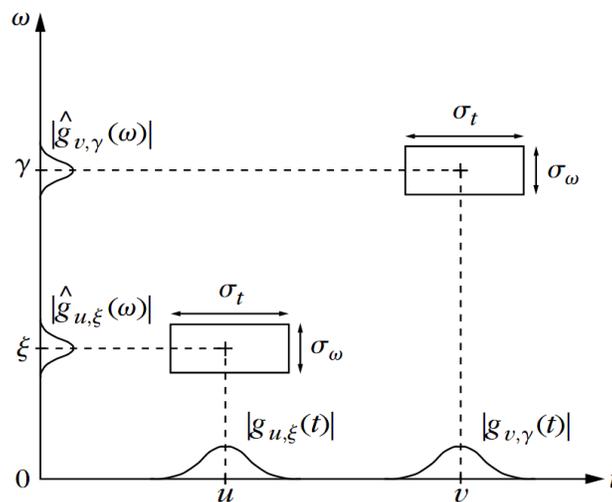


Fig. 2.1. Regiones de Heisenberg para la STFT [22].

2.3. Transformada Wavelet

Así como en la teoría de Fourier, las wavelets revelan características de la señal a través de la amplitud de sus coeficientes y su estructura permite el desarrollo de un algoritmo computacionalmente eficiente. Al contrario de Fourier, la transformada wavelets solo requiere

de unos cuantos coeficientes para representar estructuras transitorias. Para analizar estructuras a diferentes tamaños se requiere de funciones con soporte flexible en tiempo. La transformada wavelet descompone una señal sobre una base de wavelets escaladas y trasladadas. Una wavelet es una función $\psi \in \mathbf{L}^2(\mathbb{R})$ con media cero, está normalizada $\|\psi\| = 1$ y centrada alrededor de $t = 0$, el conjunto de funciones que se obtiene de escalar ψ por s y trasladarla por u se define a continuación [13], [22], [34]

$$\left\{ \psi_{u,s}(t) = \frac{1}{\sqrt{s}} \psi \left(\frac{t-u}{s} \right) \right\}_{u \in \mathbb{R}, s \in \mathbb{R}^+} \quad (2-5)$$

Por lo tanto la transformada wavelet de una señal $f \in \mathbf{L}^2(\mathbb{R})$ es,

$$Wf(u, s) = \langle f, \psi_{u,s} \rangle = \int_{-\infty}^{\infty} f(t) \frac{1}{\sqrt{s}} \psi^* \left(\frac{t-u}{s} \right) dt \quad (2-6)$$

Esta transformada puede interpretarse como un filtro lineal, si es representada como una convolución [22],

$$Wf(u, s) = \int_{-\infty}^{\infty} f(t) \frac{1}{\sqrt{s}} \psi^* \left(\frac{t-u}{s} \right) dt = f * \bar{\psi}_s(u) \quad (2-7)$$

donde,

$$\bar{\psi}_s = \frac{1}{\sqrt{s}} \psi^* \left(\frac{-t}{s} \right) \quad (2-8)$$

Al obtener la transformada de Fourier de (2-8) se tiene,

$$\hat{\psi}_s(\omega) = \sqrt{s} \psi^*(s\omega) \quad (2-9)$$

como $\psi(0) = \int_{-\infty}^{\infty} \psi(t) dt = 0$, se puede ver que $\hat{\psi}_s$ es la función de transferencia de un filtro pasa-banda. La convolución en (2-7) obtiene la transformada wavelet con filtros pasa-banda dilatados [22].

Los coeficientes $Wf(u, s)$ que se obtienen con la transformada wavelet describen la correlación entre la señal o forma de onda y la wavelet a diferentes escalas y traslaciones.

Debido a que el área de la wavelet es cero para señales constantes la relación será de cero, esta relación se hace más fuerte cuando la señal se asemeja a la wavelet utilizada a una escala dada [34].

2.3.1. Parámetros de escala-tiempo

Los sistemas wavelet son generados a partir de una simple función de escalamiento o wavelet por transformaciones de escalamiento y traslación en tiempo. Si hacemos un cambio de variable con $u = b$ y $s = a$, la familia de wavelets parte de la función wavelet madre definida como

$$\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right), \quad (2-10)$$

donde $a \in \mathbb{R}^+$, $b \in \mathbb{R}$, con $a \neq 0$. $a = f_0/f$ siendo el factor de escala, f_0 denota la frecuencia central de la wavelet y b denota el factor de traslación en tiempo. Como $a_0 < a_1$ corresponden a las dos frecuencias f_0/f_1 , la discretización natural de los parámetros de escala-tiempo se obtiene sub-muestreando los coeficientes de la escala a_1 a la tasa de los coeficientes de la escala a_0 de acuerdo al teorema de Nyquist. Para muestrear la wavelets los parámetros de escala (o dilatación) y traslación de la Ec. (2-10) se eligen de forma que al nivel j la wavelet $a_0^j \psi(a_0^{-j}t)$ es a_0^j veces el tamaño de $\psi(t)$. Entonces se tiene, el parámetro de escala $\{a = a_0^j : j \in \mathbb{Z}\}$ y el parámetro de traslación $\{b = kb_0 a_0^j : j, k \in \mathbb{Z}\}$ [22], [33], donde \mathbb{Z} es el grupo de todos los enteros. De este modo, la familia de wavelet es dada por

$$\psi_{j,k}(t) = a_0^{-j/2} \psi(a_0^{-j}t - kb_0) \quad (2-11)$$

Siendo $a_0 > 1$ y $b_0 > 0$, donde la escala en espacio o tiempo está dada por k y la escala de frecuencia está dada por j siendo. Dados a_0 y b_0 y partiendo de (2-6), la versión discreta de la transformada wavelet es

$$d_{j,k} = \langle x(t), \psi_{j,k}(t) \rangle = a_0^{-j/2} \int_{-\infty}^{\infty} x(t) \psi(a_0^{-j}t - kb_0) dt, \quad (2-12)$$

donde $\langle \cdot, \cdot \rangle$ es el producto interno para L^2 .

Los parámetros de escalamiento a_0 , traslación b_0 y la función madre wavelet $\psi(t)$ se escogen tales que $\psi_{j,k}(t)$ constituya una base ortogonal de $L^2(\mathbb{R})$. Para formar una base ortogonal con buenas propiedades de localización tiempo-frecuencia, los parámetros de escala-tiempo (a, b) son muestreados sobre la llamada *malla diádica* en el plano escala-tiempo, asignando, $a_0 = 2$ y $b_0 = 1$, [33]. Por lo que sustituyendo estos valores en (2-12) se obtiene una familia wavelet con bases ortogonales,

$$\psi_{j,k}(t) = 2^{-j/2} \psi(2^{-j} t - k), \quad (2-13)$$

donde ortogonalidad se puede definir como,

$$\langle \psi_{j,k}(t), \psi_{j',k'}(t) \rangle = \int_{-\infty}^{\infty} \psi_{j,k}(t) \psi_{j',k'}(t) dt = \begin{cases} 1 & \text{si } j = j' \text{ y } k = k' \\ 0 & \text{otro caso} \end{cases} \quad (2-14)$$

Usando (2-12), la transformada wavelet ortogonal está dada por

$$d_{j,k} = \langle x(t), \psi_{j,k}(t) \rangle = 2^{-j/2} \int_{-\infty}^{\infty} x(t) \psi(2^{-j} t - k) dt \quad (2-15)$$

donde $d_{j,k}$ son los coeficientes de detalle.

Una aproximación formal para la construcción de bases ortogonales está dada por el análisis multiresolución (MRA por sus siglas en ingles). La idea del MRA es escribir una función $x(t)$ como un límite de aproximaciones sucesivas, siendo cada una, una versión más suave de $x(t)$. Por lo tanto, las aproximaciones sucesivas corresponden a diferentes resoluciones [33].

2.3.2. Análisis multiresolución

El análisis multiresolución está definido como un grupo de secuencias de sub-espacios cerrados $\{V_j \subset L^2(\mathbb{R}) : j \in \mathbb{Z}\}$ con las siguientes propiedades:

- (i) ... $V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \subset \dots \subset L^2(\mathbb{R})$ (ej., $V_j \subset V_{j-1}$);
- (ii) $\bigcap_{j \in \mathbb{Z}} V_j = \{0\}$ y $\bigcup_{j \in \mathbb{Z}} V_j = L^2(\mathbb{R})$;

(iii) $\forall j \in \mathbb{Z}, x(t) \in V_j \Leftrightarrow x(2t) \in V_{j-1}$;

(iv) $\forall k \in \mathbb{Z}, x(t) \in V_0 \Leftrightarrow x(t - k) \in V_0$;

(v) Existe una función $\varphi(t) \in V_0$ tal que $\{\varphi_{j,k}(t) = 2^{-j/2}\psi(2^{-j}t - k) : j, k \in \mathbb{Z}\}$ satisface la Ec. (2-13) y forma un espacio base ortogonal de V_0 .

La propiedad (i) denota los espacios sucesivos usados para representar a las diferentes resoluciones o escalas, mientras que la propiedad (ii) garantiza que los sub-espacios estén completos y asegura que el $\lim_{j \rightarrow -\infty} x_j(t) = x(t)$. La propiedad (iii) dice que V_{j-1} consiste de todas las versiones re-escaladas de V_j , mientras que la propiedad (iv) significa que cualquier versión de una función trasladada pertenece al mismo espacio que la función original. Finalmente, la propiedad (v), la función $\varphi(\cdot)$ es llamada la *función de escalamiento* en el análisis multiresolución [22], [33], [34].

2.3.3. Funciones de escalamiento y funciones wavelet

Como la función $\varphi(t) \in V_0 \subset V_{-1}$ y $\varphi_{-1,l}(t)$ es una base ortogonal en V_{-1} , $\varphi(t)$ se puede escribir como sigue:

$$\varphi(t) = \sum_{l \in \mathbb{Z}} \bar{g}_l \varphi_{-1,l}(t), \quad (2-16)$$

donde $\bar{g}_l = \langle \varphi(t), \varphi_{-1,l}(t) \rangle$. Aplicando las propiedades (iii) y (v) de MRA resulta que $\varphi_{-1,l}(t) = \sqrt{2} \varphi(2t - l)$. Entonces la Ec. (2-16) puede expresarse como:

$$\varphi(t) = \sqrt{2} \sum_{l \in \mathbb{Z}} \bar{g}_l \varphi(2t - l), \quad (2-17)$$

donde los coeficientes \bar{g}_l son una secuencia de números reales o tal vez complejos llamados coeficientes de escalamiento, filtros de escalamiento o elementos de un vector de escalamiento. Esta ecuación recursiva también es llamada como *ecuación de dos escalas* o *ecuación de dilatación*.

La respuesta en frecuencia de la función de escalamiento $\varphi(t)$ es la de un filtro pasa bajas y la secuencia $\{\bar{g}_l\}$ es conocida como coeficientes del filtro pasa bajas [22], [33], [34].

Ahora si dejamos que W_j sea el complemento ortogonal de V_j en V_{j-1} , por lo satisface la condición,

$$V_{j-1} = V_j \oplus W_j \quad y \quad V_j \perp W_j \quad (2-18)$$

donde el símbolo \oplus significa una suma directa, y \perp significa ortogonalidad y W_j es extendido por el grupo $\{\psi_{j,k} : k \in \mathbb{Z}\}$. La suma directa significa que cualquier elemento en V_{j-1} puede ser expresado como la suma de dos elementos ortogonales $\{V_j \text{ y } W_j\}$. La Ec. (2-18) dice que el sub-espacio W_j contiene la información detallada que se necesita para ir de una aproximación en el nivel de resolución j , a una aproximación en el nivel de resolución $j - 1$. Aplicando la propiedad (ii) se puede ver que la suma directa de sub-espacios infinitos W_j es todo el $L^2(\mathbb{R})$. Esto se puede ver en la Fig. 2.2 [33].

Una función $\psi(\cdot)$ es una wavelet si la colección de funciones $\{\psi(t - k) : k \in \mathbb{Z}\}$ es una base completa ortogonal de W_0 . Entonces la función wavelet puede ser representada como la suma ponderada de la función de escalamiento desplazada $\varphi(2t)$ por

$$\psi(t) = \sqrt{2} \sum_{l \in \mathbb{Z}} \bar{h}_l \varphi(2t - l). \quad (2-19)$$

La secuencia $\{\bar{h}_l\}$ se conoce como un filtro pasa banda, dado que la respuesta en frecuencia de la wavelet es un filtro pasa banda [22], [33].

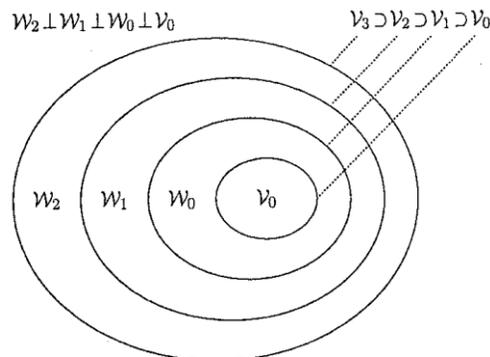


Fig. 2.2. Funciones de escalamiento y espacios vectoriales wavelets [33].

2.4. Transformada discreta wavelet

Como la idea del análisis de multiresolución es determinar una señal $x(t)$ como un límite de aproximaciones sucesivas, las diferencias entre las dos aproximaciones sucesivas en la resolución 2^{j-1} y 2^j dan los detalles de la señal en la resolución 2^j . En otras palabras, después de elegir una resolución inicial J , cualquier señal $x(t) \in L^2(\mathbb{R})$ puede ser expresada como [33]:

$$x(t) = \sum_{k \in \mathbb{Z}} c_{J,k} \varphi_{J,k}(t) + \sum_{j=J}^{\infty} \sum_{k \in \mathbb{Z}} d_{j,k} \psi_{j,k}(t), \quad (2-20)$$

donde los coeficientes de detalle o *coeficientes wavelet* $\{d_{j,k}\}$ están definidos por

$$d_{j,k} = 2^{-j/2} \int_{-\infty}^{\infty} x(t) \psi_{j,k}(2^{-j}t - k) dt \quad (2-21)$$

mientras que los coeficientes de aproximación o *coeficientes de escalamiento* $\{c_{j,k}\}$ están definidos por

$$c_{j,k} = 2^{-j/2} \int_{-\infty}^{\infty} x(t) \varphi_{j,k}(2^{-j}t - k) dt. \quad (2-22)$$

Las ecuaciones (2-21) y (2-22) expresan que una señal $x(t)$ se descompone en detalles $\{d_{j,k}\}$ y aproximaciones $\{c_{j,k}\}$ para crear un análisis multiresolución de la señal original [33].

2.5. Bancos de filtros y la transformada wavelet

Una propiedad atractiva de las wavelets es que existe una relación recursiva entre los coeficientes de escalamiento $\{c_{j,k}\}$ y wavelets $\{d_{j,k}\}$ a niveles sucesivos de resolución. Usando la Ec. (2-22) y la ecuación de dilatación (2-17) se tiene que

$$\begin{aligned} c_{j,k} &= \sum_{l \in \mathbb{Z}} \bar{g}_l \int_{-\infty}^{\infty} x(t) 2^{\frac{-j+1}{2}} \varphi(2^{-j+1}t - 2k - l) dt \\ &= \sum_{l \in \mathbb{Z}} \bar{g}_{m-2k} \int_{-\infty}^{\infty} x(t) 2^{\frac{-j+1}{2}} \varphi(2^{-j+1}t - m) dt, \end{aligned} \quad (2-23)$$

donde $m = 2k + l$. Sustituyendo la ecuación (2-22) en (2-23) se obtiene la relación iterativa [33]:

$$c_{j,k} = \sum_m \bar{g}_{m-2k} c_{j-1,m}. \quad (2-24)$$

Ahora empezando con (2-21) y (2-19), empleando el mismo procedimiento, se obtiene

$$d_{j,k} = \sum_m \bar{h}_{m-2k} c_{j-1,m}. \quad (2-25)$$

Haciendo un cambio de variable, dejando $g_l = \bar{g}_{-l}$, $h_l = \bar{h}_{-l}$ y $l = 2k - m$, se obtiene las siguientes ecuaciones [22], [28], [33], [34]:

$$c_{j,k} = \sum_m g_l c_{j-1,2k-l} \quad (2-26)$$

$$d_{j,k} = \sum_m h_l c_{j-1,2k-l}. \quad (2-27)$$

En las Ecuaciones (2-26) y (2-27) se realiza un filtrado y un sub-muestreo de la señal de entrada y muestran que los coeficientes de escalamiento y wavelet a diferentes niveles de escala pueden ser obtenidos haciendo la convolución de los coeficientes de expansión a escala j con los coeficientes recursivos invertidos en tiempo g_l y h_l luego sobre-muestreando para dar los coeficientes de expansión para el siguiente nivel de $j - 1$. La implementación de (2-26) y (2-27) se ilustra en la Fig. 2.3 en donde $\downarrow 2$ denotan el sub-muestreo y los recuadros denotan el filtrado FIR (Respuesta al impulso finita) o la convolución por g_l y h_l [28].

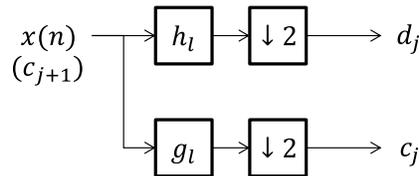


Fig. 2.3. Banco de análisis de dos bandas.

El filtrado que hace g_l es pasa bajas y h_l hace un filtrado pasa altas. Para los coeficientes de escalamiento, estos no dan mucho detalle de la señal por lo que dan información de las bajas

frecuencias. Los coeficientes wavelet o de detalle, están relacionados con las altas frecuencias debido a que la información que nos dan es de las componentes de alta resolución de la señal. En la Fig. 2.4 se puede observar un arreglo de dos etapas y de dos bandas [28].

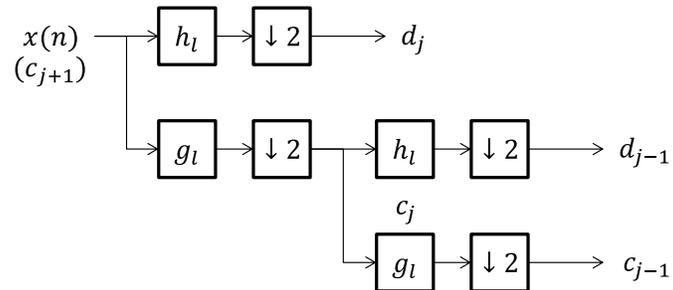


Fig. 2.4. Árbol de análisis de dos etapas y dos bandas.

Este arreglo se obtiene aplicando el filtrado y el submuestreo a los coeficientes de escalamiento. La repetición de este proceso es llamada iteración de los bancos de filtros. De acuerdo a lo anterior, se puede realizar la descomposición partiendo de niveles de resolución de menor detalle a mayor detalle como se muestra en la Fig. 2.5 [22], [28].

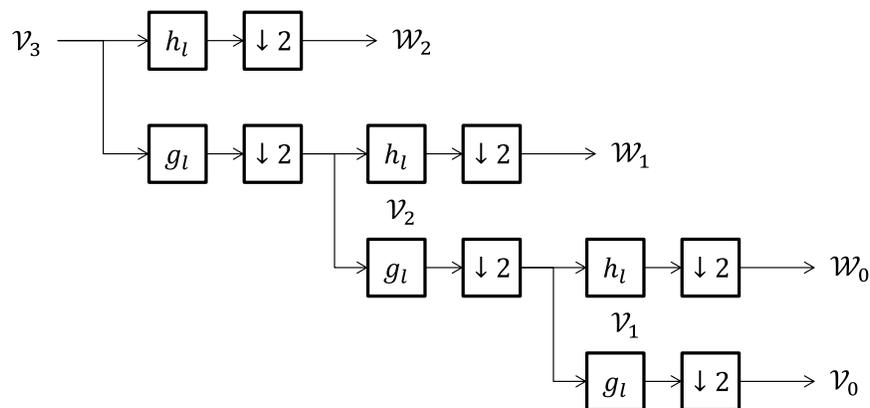


Fig. 2.5. Filtrado para tres etapas.

La reconstrucción se realiza de la escala de menor detalle a la de mayor detalle. Primero se realiza el proceso inverso al submuestreo, denominado sobremuestreo. Esto quiere decir que a la entrada del filtro se insertan ceros entre cada muestra original de la señal [28].

2.6. Transformada Hilbert

Con la transformada Hilbert se puede obtener la potencia relativa de la señal a lo largo del tiempo; si ésta es filtrada a cierta frecuencia con un ancho de banda angosto. A continuación se describe de forma general la metodología que se sigue para obtener una representación para un intervalo de frecuencias dado. La transformada Hilbert de una señal $g(t)$ se denota como $\hat{g}(t)$ y se muestra en la siguiente ecuación,

$$\mathcal{H}[g(t)] = \hat{g}(t) = g(t) * \frac{1}{\pi t} \quad (2-28)$$

Al calcular la transformada de Fourier de $1/\pi t$ se obtiene la siguiente expresión,

$$-j \operatorname{sgn}(f) = \begin{cases} -j, & \text{si } f > 0 \\ 0, & \text{si } f = 0 \\ j, & \text{si } f < 0 \end{cases} \quad (2-29)$$

Si la transformada de Fourier de $g(t)$ es $G(f)$ por las propiedades de la convolucion de la transformada de Fourier se tiene que $\hat{g}(t)$ tiene la transformada de Fourier,

$$G(f) = -j \operatorname{sgn}(f)G(f) \quad (2-30)$$

por lo anterior se puede ver que la transformada Hilbert es más fácil de comprender en el dominio de la frecuencia donde se observa que no afecta a la magnitud de la señal al hacer solamente un cambio de fase de $\pm \frac{\pi}{2}$ dependiendo del signo de la transformada de Fourier de la señal. También es importante mencionar que la transformada Hilbert no afecta a la densidad espectral de potencia de la señal. Supongamos que $g(t)$ es una señal de potencia, entonces, $|\hat{G}(f)| = |G(f)|$ [21], [13].

2.6.1. Señal analítica

Una señal filtrada por un filtro pasa banda con una banda de paso pequeña se puede observar como una señal modulada en amplitud (ver Fig. 2.7) la cual se puede describir de la siguiente forma,

$$g(t) = a(t) \cos(2\pi f_0 t + \theta) \quad (2-31)$$

donde $a(t)$ es una señal de baja frecuencia que contiene información relacionada con ERD-ERS que está ligada al tiempo en que se presenta el estímulo, f_0 es la frecuencia de la señal modulada y θ es la fase inicial de la señal. La transformada Hilbert de (2-31) puede ser obtenida como [21], [17], [18],

$$\hat{g}(t) = a(t) \sin(2\pi f_0 t + \theta) \quad (2-32)$$

Con la cual se puede formar la señal analítica,

$$z(t) = g(t) + j\hat{g}(t) = a(t)e^{2\pi f_0 t + \theta}, \quad (2-33)$$

por lo tanto la envolvente puede ser estimada de la siguiente manera,

$$w(t) = |z(t)| = \sqrt{g^2(t) + j\hat{g}^2(t)}. \quad (2-34)$$

El fenómeno ERD–ERS puede ser observado cuando se hace el promedio de todas las muestras (Fig. 2.7). Para una sola muestra la señal filtrada contiene información correspondiente a su potencia instantánea; una de las propiedades de la transformada Hilbert.

2.7. Representación Tiempo-frecuencia

La representación tiempo frecuencia (TFR) pretende identificar los fenómenos presentes en la señal (ERD, ERS) los cuales fueron explicados anteriormente. Como herramienta de cálculo se utiliza la transformada Hilbert con la cual se obtiene la envolvente de la señal. Wang *et. Al* [12] y Wang y He [17] emplearon la transformada Hilbert para generar una representación tiempo-frecuencia; a diferencia de estos trabajos, en este estudio no se realizaron transformaciones posteriores de las envolventes para formar un vector de características. En la Fig. 2.7 se observa la señal correspondiente a una prueba, la señal filtrada en para una banda en particular y su envolvente, y el promediado de las envolventes para todas las pruebas.

Partiendo de ese concepto se puede formar una representación en tiempo-frecuencia. Se hace un barrido en la señal para cada banda de frecuencia con filtros Butterworth de orden 5, de

ancho de banda de 2 Hz y traslapados en un 50 % (ver Fig. 2.6). En la Fig. 2.9 se observa una TFR para las primeras 20 muestras de S1 (ver Tabla 4-1 y Tabla 4-2) partiendo de este principio.

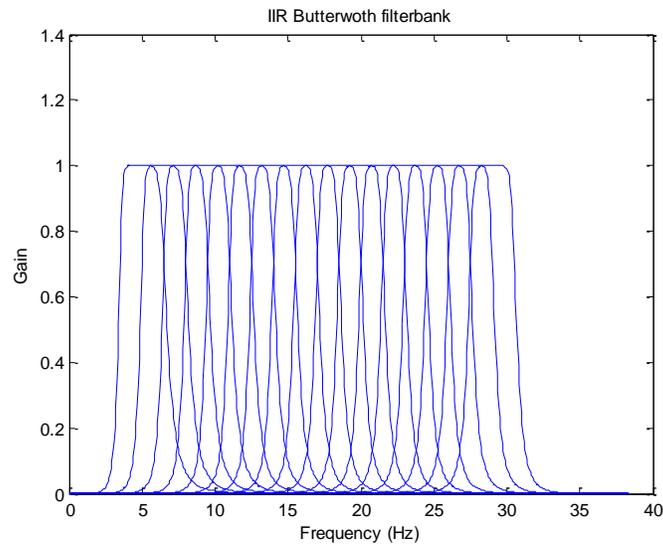


Fig. 2.6. Banco de filtros butterworth. Bw = 3 Hz. Traslape de 50 %.

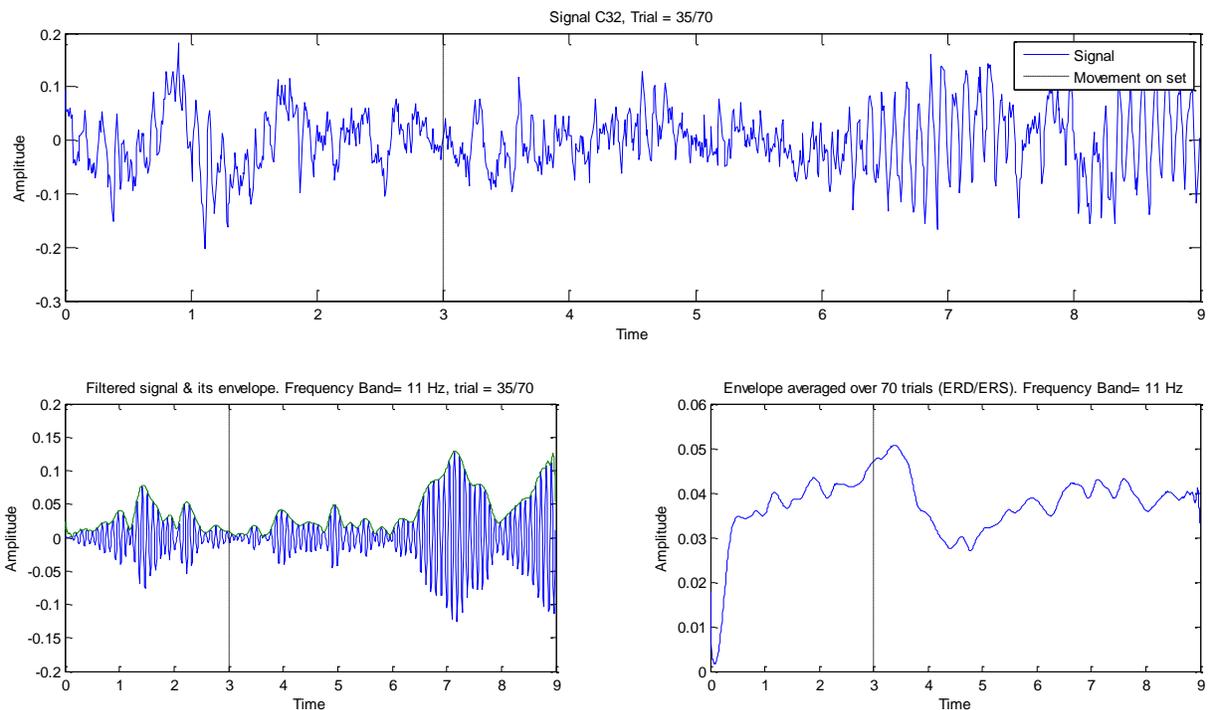


Fig. 2.7. Proceso de extracción de envolventes de la señal para la obtención de la TFR y el fenómeno ERD – ERS. Bw = 4 Hz.

En la Fig. 2.8(a) se muestra el espectrograma de una señal simulada, si lo comparamos con la Fig. 2.8(b), representación tiempo-frecuencia con patrones espaciales, observamos que la información es similar con un poco más de resolución en frecuencia. Si solo se extrae la información en frecuencia de las bandas simuladas obtenemos la gráfica de la Fig. 2.8(c) la cual muestra los eventos ERD-ERS en cada banda.

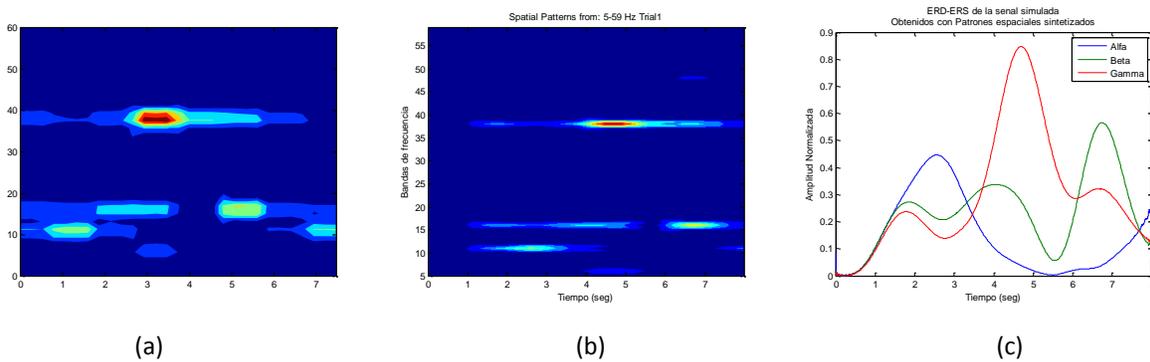


Fig. 2.8. (a) Espectrograma calculado con STFT. (b) Representación tiempo-frecuencia por medio de patrones espaciales. (c) ERD-ERS extraídos de (b).

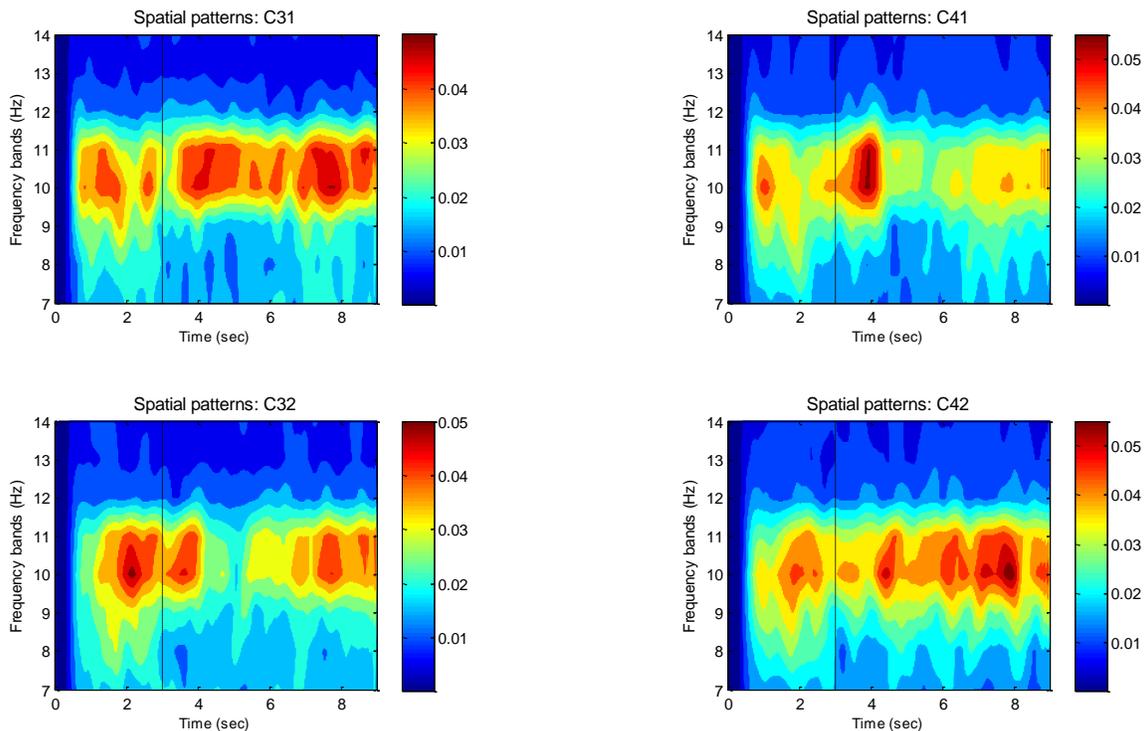


Fig. 2.9. Patrones espaciales promediados de las primeras 20 muestras de S1. Solo se toman las frecuencias de 7 a 14 Hz debido a que se encontró que las frecuencias reactivas se concentran en ese rango. Las líneas punteadas marcan el momento donde se muestran el estímulo del movimiento a imaginar. A partir del segundo tres se observan variaciones en la energía de la señal.

2.8. Sumario

En este capítulo se establecieron las características principales de las wavelets iniciando la discusión con las diferencias, similitudes y ventajas que existen con el análisis de Fourier. Se explicó la wavelet continua como punto de partida para la wavelet discreta, para lo cual, se describió el análisis multiresolución. De lo anterior, se estableció su relación con los bancos de filtros y su representación como un árbol de descomposición a diferentes niveles. Por otro lado se mencionó la transformada Hilbert como herramienta de cálculo de los fenómenos ERD/ERS, con lo que se pudo construir una representación tiempo-frecuencia que mejora la visualización del fenómeno para este tipo de señales en comparación con la STFT.

3. CLASIFICADORES

3.1. Análisis de Discriminantes Lineales por Fisher (LDA)

Para el caso general de dos clases se tiene que una función discriminante es una combinación lineal de los componentes en \mathbf{x} y puede ser escrita como,

$$g(\mathbf{x}) = \mathbf{w}^t \mathbf{x} + w_0 \quad (3-1)$$

donde \mathbf{w} es el vector de pesos o vector de transformación y w_0 es el umbral que mueve al vector de pesos del origen. Un clasificador lineal de dos clases implementa la siguiente regla de decisión: ω_1 si $g(\mathbf{x}) > 0$ y ω_2 si $g(\mathbf{x}) < 0$ [15]. Lo que significa que \mathbf{x} es asignada a ω_1 si el producto interno o proyección $\mathbf{w}^t \mathbf{x}$ excede el umbral $-w_0$, o a ω_2 si sucede lo contrario. Lo importante no es la magnitud sino la dirección de \mathbf{w} . La clave es encontrar una \mathbf{w} que mejor separe ambas clases. El análisis por discriminantes lineales por Fisher considera que las matrices de covarianza de cada clase son iguales por lo que se calculan matrices de dispersión en donde agrupan las varianzas de ambas clases. El mejor \mathbf{w} es el que maximiza la función de criterio (3-1) la cual esta expresada en términos de la distancias entre las medias y de las varianzas agrupadas proyectadas

$$J(\mathbf{w}) = \frac{|\tilde{m}_1 - \tilde{m}_2|^2}{\tilde{s}_1^2 + \tilde{s}_2^2} \quad (3-2)$$

Para obtener $J(\cdot)$ como función de \mathbf{w} se definen las matrices de dispersión \mathbf{S}_i y \mathbf{S}_W como,

$$\mathbf{S}_i = \sum_{\mathbf{x} \in \mathcal{D}_i} (\mathbf{x} - \mathbf{m}_i)(\mathbf{x} - \mathbf{m}_i)^t \quad (3-3)$$

$$\mathbf{S}_W = \mathbf{S}_1 + \mathbf{S}_2 \quad (3-4)$$

por lo que,

$$\tilde{s}_1^2 + \tilde{s}_2^2 = \mathbf{w}^t \mathbf{S}_W \mathbf{w} \quad (3-5)$$

$$\begin{aligned}
(\tilde{m}_1 - \tilde{m}_2)^2 &= (\mathbf{w}^t \mathbf{m}_1 - \mathbf{w}^t \mathbf{m}_2)^2 \\
&= \mathbf{w}^t (\mathbf{m}_1 - \mathbf{m}_2) (\mathbf{m}_1 - \mathbf{m}_2)^t \mathbf{w} \\
&= \mathbf{w}^t \mathbf{S}_B \mathbf{w}
\end{aligned} \tag{3-6}$$

donde,

$$\mathbf{S}_B = (\mathbf{m}_1 - \mathbf{m}_2)(\mathbf{m}_1 - \mathbf{m}_2)^t \tag{3-7}$$

donde \mathbf{m}_i , para $i = 1, 2$ es la media de cada clase en \mathbf{x} , \mathbf{S}_W es la matriz de dispersión dentro de la clase y \mathbf{S}_B la matriz de dispersión entre clases. Por lo tanto,

$$J(\mathbf{w}) = \frac{\mathbf{w}^t \mathbf{S}_B \mathbf{w}}{\mathbf{w}^t \mathbf{S}_W \mathbf{w}} \tag{3-8}$$

Haciendo de simplificaciones, la solución para el vector \mathbf{w} que maximiza $J(\cdot)$ es,

$$\mathbf{w} = \mathbf{S}_W^{-1} (\mathbf{m}_1 - \mathbf{m}_2) \tag{3-9}$$

Y el umbral w_0 queda en función de la media total de las clases $\mathbf{m} = \mathbf{m}_1 + \mathbf{m}_2$ y de \mathbf{w} como,

$$w_0 = \mathbf{w} \mathbf{m}^t \tag{3-10}$$

Una vez obtenidos estos parámetros con el conjunto de entrenamiento, se puede aplicar al clasificador una muestra de prueba.

3.2. Análisis de Discriminantes Cuadráticos (QDA)

El análisis por discriminante cuadrático, a diferencia del lineal, considera que las matrices de covarianzas por clase son diferentes por lo que la función de decisión es cuadrática en \mathbf{x} . Lo anterior nos permite tener superficies o hiperplanos de decisión cuadráticos a como se muestra en la Fig. 3.1 [15].

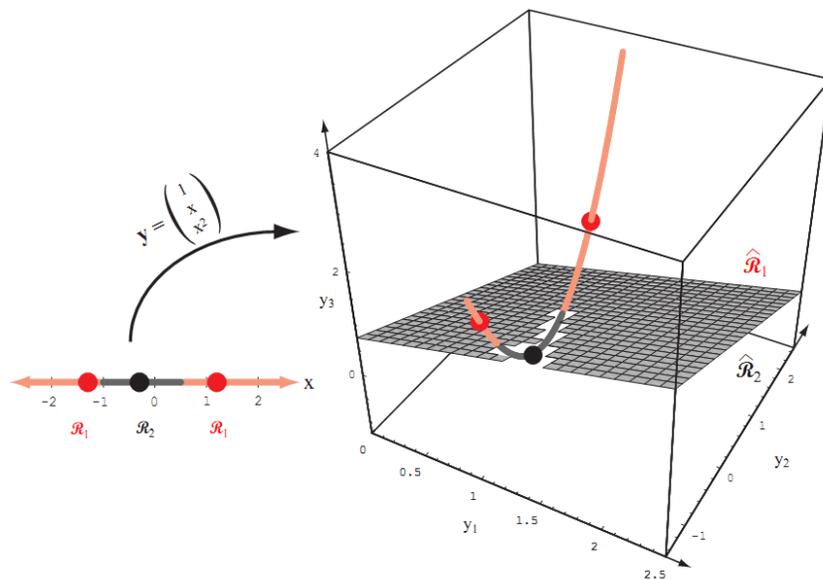


Fig. 3.1. Función de decisión cuadrática. Se forma una región separación para las dos clases, la cual no está directamente conectada [15].

3.3. Máquina de Soporte Vectorial (SVM)

La máquina de soporte vectorial surge como una alternativa para encontrar la solución al sobre ajuste; que es cuando un clasificador y su plano de decisión quedan muy ajustados al conjunto de datos, de modo que no hay ningún margen de error que permita que datos de prueba puedan ser clasificados. Una SVM es entrenada con el fin de que la habilidad de generalización de la función de decisión se maximice. En otras palabras el espacio de entrada \mathbf{x} m -dimensional es mapeado a un espacio de características l -dimensional ($l \geq m$) \mathbf{z} . En este espacio se resuelve el problema por programación cuadrática para separar ambas clases por el hiperplano óptimo de separación.

La función discriminante de la cual se parte es la misma que en (3-1) pero en lugar de igualarse a cero se iguala a una constante $-1 < c < 1$. Esta es la región de generalización de la función de decisión. En la Fig. 3.2 se muestra la región de separación óptima para cuando $D(\mathbf{x}) = 0$ y el margen máximo de separación se da para $D(\mathbf{x}) = 1$ y $D(\mathbf{x}) = -1$. Como en el caso del FLDA, la clave es encontrar la dirección de \mathbf{w} para la cual el margen de separación es el máximo. Por lo tanto, el hiperplano de separación óptimo puede ser obtenido minimizando,

$$Q(\mathbf{w}) = \frac{1}{2} \|\mathbf{w}\|^2, \tag{3-11}$$

con respecto a \mathbf{w} y b sujeto a la condición $y_i(\mathbf{w}^t \mathbf{x}_i + b) \geq 1$. El cuadrado de la norma euclidiana en (3-11) es para que la solución del problema de optimización sea de programación cuadrática. Para ello se hace una transformación que pasa de un problema condicionado a uno no condicionado de la siguiente manera,

$$Q(\mathbf{w}, b, \alpha) = \frac{1}{2} \mathbf{w}^t \mathbf{w} - \sum_{i=1}^M \alpha_i \{y_i(\mathbf{w}^t \mathbf{x}_i + b) - 1\}, \tag{3-12}$$

donde M es el tamaño del vector de entrada \mathbf{x} y y_i es la salida (clase: 1 o -1) correspondiente al vector de entrada. $\alpha = (\alpha_1, \dots, \alpha_M)^t$ son los multiplicadores de Lagrange no negativos. La solución óptima está dada por las derivadas con respecto a \mathbf{w} y b ; condiciones establecidas conocidas como Karush-Kuhn-Tucker (KKT) [15], [35], [36].

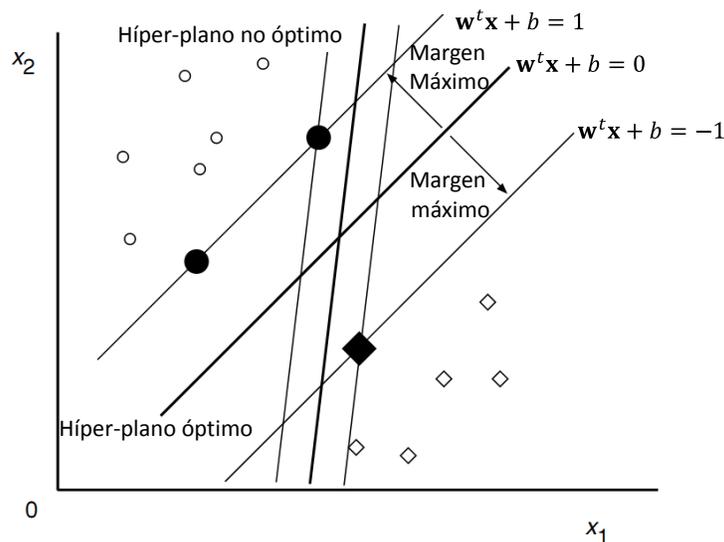


Fig. 3.2. Hiperplano de separación óptimo para un espacio bi-dimensional.

Los datos asociados a los valores positivos de α_i son los vectores de soporte para la clase 1 y la clase 2. Después de resolver (3-11) de acuerdo a las condiciones KKT se obtiene \mathbf{w} como,

$$\mathbf{w} = \sum_{i=1}^M \alpha_i y_i \mathbf{x}_i. \tag{3-13}$$

De (3–13) la función de decisión está dada por,

$$D(\mathbf{x}) = \sum_{i \in S}^M \alpha_i y_i (\mathbf{x}_i^t \mathbf{x}) + b, \quad (3-14)$$

donde S es el grupo de índices de los vectores de soporte (ver Fig. 3.2), y de las condiciones KKT se tiene que b esta dada por,

$$b = y_i - \mathbf{w}^t \mathbf{x}_i, \quad (3-15)$$

donde \mathbf{x}_i es un vector de soporte. Desde el punto de vista de la precisión de cálculo es mejor tomar la media de los vectores de soporte como sigue:

$$b = \frac{1}{|S|} \sum_{i \in S} (y_i - \mathbf{w}^t \mathbf{x}_i). \quad (3-16)$$

3.3.1. SVM y Kernels

Si los datos de entrenamiento de la SVM no son linealmente separables, el clasificador obtenido no tendrá la mejor generalización del problema aunque los hiperplanos encontrados sean óptimos. Por lo tanto, para mejorar la separabilidad de los datos el espacio de entrada original es mapeado a un espacio de igual o mayor dimensión llamado *espacio de características*; puede ser visto como una transformación no lineal de las características de entrada. La función de decisión de (3–14) puede ser reescrita como

$$D(\mathbf{x}) = \sum_{i \in S}^M \alpha_i y_i H(\mathbf{x}_i, \mathbf{x}) + b \quad (3-17)$$

Donde $H(\mathbf{x}_i, \mathbf{x})$ es una función simétrica o kernel, positiva semidefinida, que cumple con las condiciones establecidas por la teoría de Hilbert-Schmidt [35]–[37]. Y b esta dada por,

$$b = y_j - \sum_{i \in S} \alpha_i y_i H(\mathbf{x}_i, \mathbf{x}_j) \quad (3-18)$$

Para asegurar la precisión del cálculo se toma la media de (3–14):

$$b = \frac{1}{U} \sum_{j \in U} \left(y_j - \sum_{i \in S} \alpha_i y_i H(\mathbf{x}_i, \mathbf{x}_j) \right) \quad (3-19)$$

Existen diferentes tipos de kernels que pueden ser utilizadas. En la TABLA N se describen tres de las funciones más comunes de las cuales solo se tomó en cuenta la función de base radial (RBF).

Tabla 3-1. Descripción de tres de las kernels más utilizadas para la transformación de espacios de entrada [35].

Kernels	Función
<p>Lineal:</p> <p>Si el problema es linealmente separable no es necesario hacer un mapeo.</p>	$H(\mathbf{x}, \mathbf{x}') = \mathbf{x}^t \mathbf{x}'$
<p>Polinomial:</p> <p>La kernel polinomial de grado d, donde d es un número natural, esta dada por:</p>	$H(\mathbf{x}, \mathbf{x}') = (\mathbf{x}^t \mathbf{x}' + 1)^d$
<p>Función de base radial:</p> <p>Está basada en una gaussiana donde γ es un parámetro que controla el radio.</p>	$H(\mathbf{x}, \mathbf{x}') = e^{(-\gamma \ \mathbf{x} - \mathbf{x}'\ ^2)}$

3.4. Sumario

En este capítulo se esbozaron los conceptos relacionados con los diferentes clasificadores empleados en este trabajo. Primero se introdujo el Análisis de Discriminantes Lineales por Fisher de manera general. Este a su vez nos permitió introducir los conceptos necesarios para presentar el Análisis de Discriminantes Cuadráticos. Por último, se presentaron las nociones básicas de una Máquina de Vectores de Soporte iniciando con la analogía de un clasificador lineal hasta llegar a las transformaciones que se hacen a los vectores de entrada al clasificador por medio del uso de Kernels. También se mencionaron las kernels más comunes y una breve descripción de las mismas.

4. MÉTODO PROPUESTO: EXTRACCIÓN Y CLASIFICACIÓN DE PATRONES

4.1. Grupos de datos (Data sets)

Para evaluar el desempeño del algoritmo de extracción y clasificación desarrollado se utilizaron dos diferentes bloques de grupos de datos. El primer bloque consta de siete sujetos y fue descargado de la página web del grupo de BCI competition, sitio donde se concentra gran cantidad de los datos relacionados a sistemas BCI, extraídos en base a diferentes neuromecanismos, provenientes de diferentes universidades y que están basados en cierta metodología ya establecida en muchos de los casos (ver Tabla 4-1). El segundo bloque son datos propios basados en la metodología que se considera en los datos del bloque 1.

4.1.1. Grupos de datos de BCI competition

En la Tabla 4-1 se describe el primer grupo de datos. La base de datos (1) contiene información de C3, Cz y C4 por lo que se restó Cz de C3 y C4 para incrementar la relación señal a ruido. Para la base de datos (2) se proporciona la información de los 64 electrodos por lo cual se aplicó un filtrado espacial laplaciano tomando en cuenta los 4 electrodos adyacentes a C3 y C4 de acuerdo a (4-1) y con el mismo fin de (1). Por último, la (3) contiene información de C3, y C4 por lo que se trabajó de forma directa.

$$V_j^{lap} = V_j - \sum_{k \in S_j} V_k \quad (4-1)$$

donde V_j es la señal del canal j , y S_j es el índice correspondiente a los cuatro de los canales adyacentes a V_j .

Tabla 4-1. Descripción de los Grupos de datos obtenidos del BCI competition.

Grupos (referencia)	Descripción
(1). Dataset III: 2 classes. Department of Medical Informatics, Institute for Biomedical Engineering, Graz University of Technology, (Gert Pfurtscheller, Alois Schlögl)	S1: 70 muestras p/c Fs: 128 Hz Indicación de movimiento: 3 s. Electrodos: C3, Cz, C4

	Duración de la muestra: 9 s.
(2). Dataset IIIa: 4-class EEG data. BCI Competition 2003. Graz University of Technology, (Gert Pfurtscheller, Alois Schlögl)	S2: 45 muestras p/c S3: 30 muestras p/c S4: 30 muestras p/c Fs: 250 Hz Indicación de movimiento: 3 s. Electrodos: 64 (10 – 20) Duración de la muestra: 7 s.
(3). Dataset IIIb: Non-stationary 2-class BCI data. BCI Competition III. Graz University.	S5: 150 muestras p/c S6: 260 muestras p/c S7: 270 muestras p/c Fs: 125 Hz Indicación de movimiento: 3 s. Electrodos: C3, C4 Duración de la muestra: 8 s.

4.1.2. Grupo de datos propios

Basados en la actividad elegida se hizo la adquisición de los datos en el laboratorio NICE de la universidad de Texas Tech en colaboración con la Dra. Mary Baker. A continuación se describe la metodología y equipo utilizados.

4.1.2.1. Experimento

El experimento se diseñó basado en la metodología estudiada por otros autores (ver Tabla 4-1). Cada prueba consta de 8 segundos. En los primeros dos segundos se presenta una pantalla en blanco, en el segundo 2 se reproduce un sonido (beep) de 0.125 segundos y se muestra una cruz en el centro de la pantalla con una duración de un segundo. A partir del segundo 3 se superpone a la cruz una flecha apuntando hacia la derecha o izquierda, con una duración de 1.125 segundos, indicando el movimiento a imaginarse (mano izquierda o derecha). Al segundo 4.125 la flecha desaparece y solo queda la cruz en el centro de la pantalla hasta el segundo 7 en el que desaparece la cruz y queda la pantalla en blanco por un segundo (ver Fig. 4.1). El experimento fue programado en E-PRIME, sistema de adquisición que viene acompañado con el equipo de EEG utilizado. El experimento fue repetido de forma consecutiva N veces en donde se le indicó al sistema cuantas pruebas se quieren de cada clase para ser presentadas de forma aleatoria.

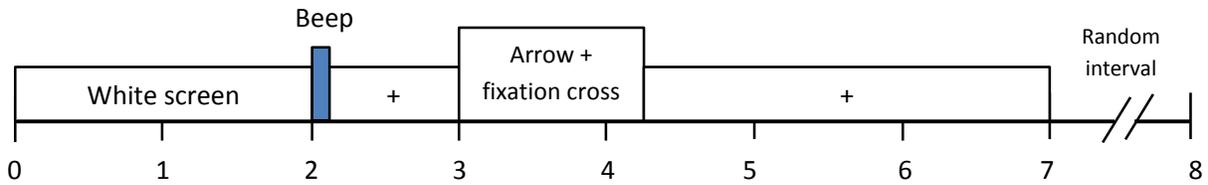


Fig. 4.1. Línea de tiempo del estímulo diseñado para la actividad de imaginación motora.

4.1.2.2. Procesamiento de la señal

El experimento fue grabado con una red de 64 electrodos HCGSN de EGI con una distribución 10 – 10 que muestra en la Fig. 4.2. El muestreo fue de 125 Hz. Después de que la señal fue grabada se realizó un filtrado pasa banda de 0.3 – 30 Hz, con lo que se limita el ancho de banda a el rango de frecuencias donde se presenta el fenómeno ERD/ERS. Una vez filtrada, se realizó la segmentación por clase y se obtuvieron arreglos que puede ser leídos en Matlab con las variables: Class1 y Class2. Cada una es un arreglo tridimensional de $(ExMxN)$ donde E es el número de electrodos, M el numero de muestras en tiempo (125 muestras por segundo) y N es el número de pruebas.

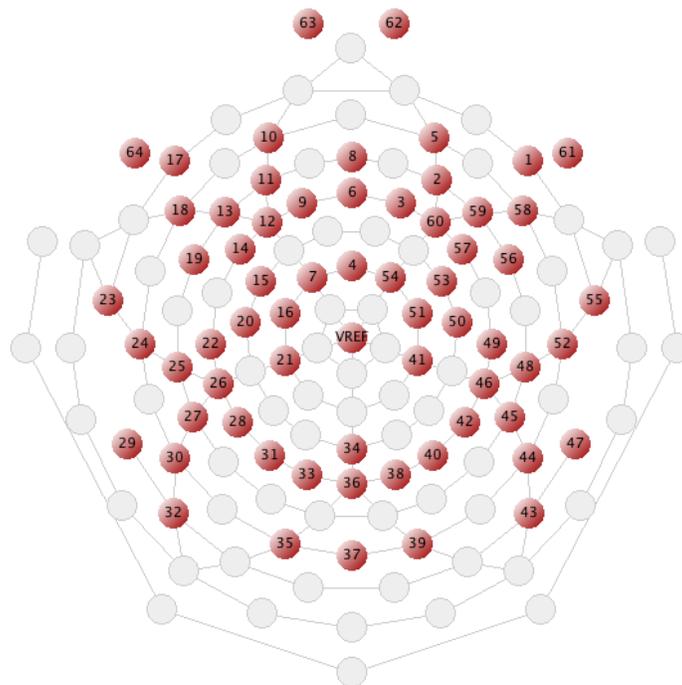


Fig. 4.2. Diagrama de ubicación de los sensores en la red EGI con una distribución 10 – 10 (Imagen tomada del software EGI).

Para fines de análisis solo se toma la información de los electrodos C3 y C4. Como se tiene la información de los 64 electrodos se realiza un filtrado espacial laplaciano para mejorar la relación señal a ruido de la señal (Ec. 4–1). De la Fig. 4.2 se observa que los electrodos 20 y 50 corresponden a C3 y C4 respectivamente. Los electrodos adyacentes a C3 son 15, 16, 21 y 22 y para C4 son 41, 49, 51 y 53 con los cuales se realiza el filtrado. Se grabaron varias sesiones con diferentes sujetos, de las cuales solo se seleccionaron las mejores. En la tabla 3 se describe la base de datos después de la selección y procesamiento de las señales.

Tabla 4-2. Grupo de datos propios extraídos en Texas Tech University.

Grupo (referencia)	Descripción
(1). Motor Imagery, 2 classes. Department of Electrical Engineering, Texas Tech University, NICE Lab, (Carrera-Leon, Obed; Dacroz-Baron, David; Baker, Mary; Ramirez-Cortes, J. Manuel)	S8: 112 muestras p/c S9: 196 muestras p/c Fs: 125 Hz Indicación de movimiento: 3 s. Electrodos: 64 (10 – 10) Duración de la muestra: 8 s.

4.2. Extracción de características por medio de Patrones Espaciales

Como se describió anteriormente, mediante la técnica de patrones espaciales se puede observar el fenómeno ERD/ERS, este resultado se puede utilizar como vector de características para la clasificación de ambas clases. La única desventaja es que el resultado es de la misma longitud de la señal. Por lo anterior se requiere de una reducción en la dimensión de este para poder tener una clasificación más eficiente. A continuación se describe la manera en que los patrones espaciales fueron manejados para tener un vector de características de dimensión menor.

Se tiene una prueba x_i donde $i = 1, 2, \dots, N$ con M muestras, a una frecuencia de muestreo F_s , cada prueba tiene una duración promedio de 7 s (dependiendo de la base de datos y la F_s , ver Tabla 4-1 y Tabla 4-2). La señal x_i es filtrada en una banda específica usando un filtro butterworth pasa-banda de orden 5. La banda de paso se eligió de 9 a 12 Hz que corresponde a la parte alta de la banda alpha y la parte baja de la banda beta [17]-[19], [23]-[26], esta elección

se puede justificar al ver la representación tiempo frecuencia, en la Fig. 4.3 se muestra la representación tiempo-frecuencia para S1 en la cual se puede confirmar que la actividad ERD/ERS de imaginación motora se presenta alrededor de dicha banda. Una vez filtrada la señal se obtiene la envolvente de acuerdo a (1-7) y se utiliza como vector de características, como se mencionó anteriormente, este resultado muestra el fenómeno ERD/ERS de la señal.

Como el vector de características tiene el mismo número de muestras de la señal, para reducirlo, se toman las muestras correspondientes a un segundo de la señal y se calcula la media, de modo que si la prueba tiene una duración de 7 s, el vector se reducirá a 7 características. Este proceso se hace para cada electrodo (C3, C4) y para cada clase (1 = Izq., 2 = Der.). En la Fig. 4.4 se observan los vectores de características antes de reducir el número de muestras, se puede ver que el comportamiento ERD/ERS está definido para cada clase. Para $3.5 < s < 7$, para la mano izquierda se tiene que C3 tiene mayor amplitud y C4 menor y para la mano derecha se tiene lo contrario, C3 es menor y C4 mayor.

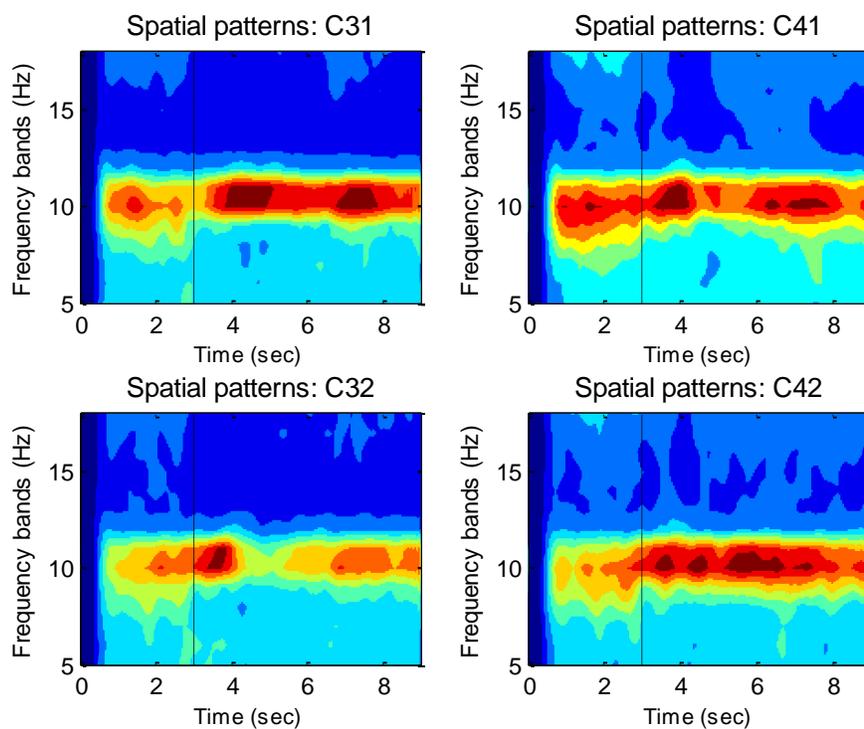


Fig. 4.3. Representación tiempo-frecuencia para S1. Los mapas son calculados para cada prueba y electrodo y al final se obtiene la media sobre todas las pruebas para obtener los mapas de referencia para cada electrodo. La línea punteada indica el tiempo en que se indica el movimiento.

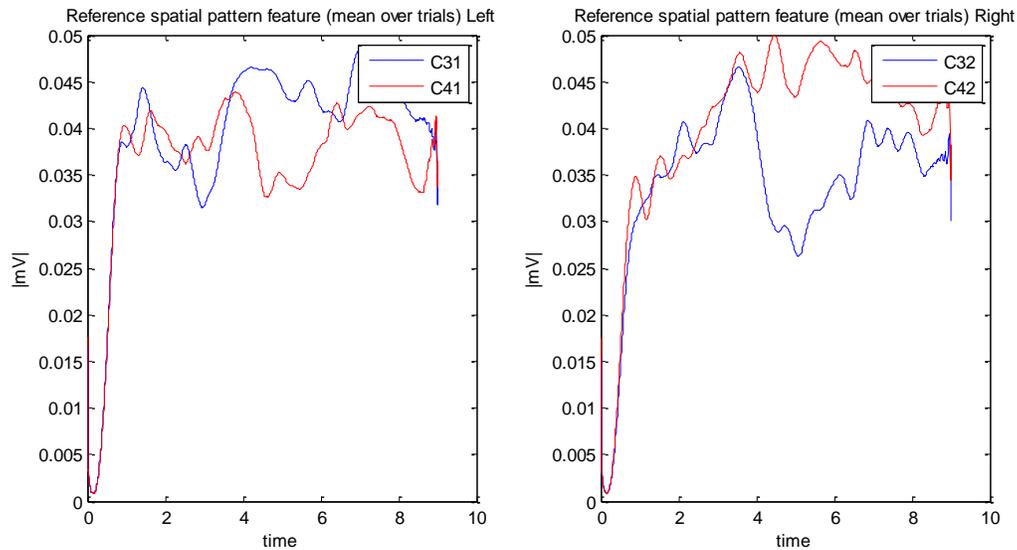


Fig. 4.4. Vectores de características ERD/ERS. Se muestra el comportamiento para cada clase y ambos electrodos por clase.

En la Fig. 4.5 se observa lo que se obtiene cuando se restan los electrodos por clase antes de reducir el número de características. A partir del $s = 3.5$ se observa la diferencia entre los patrones ERD/ERS de cada clase. Partiendo de esta observación se puede decir que es posible la clasificación.

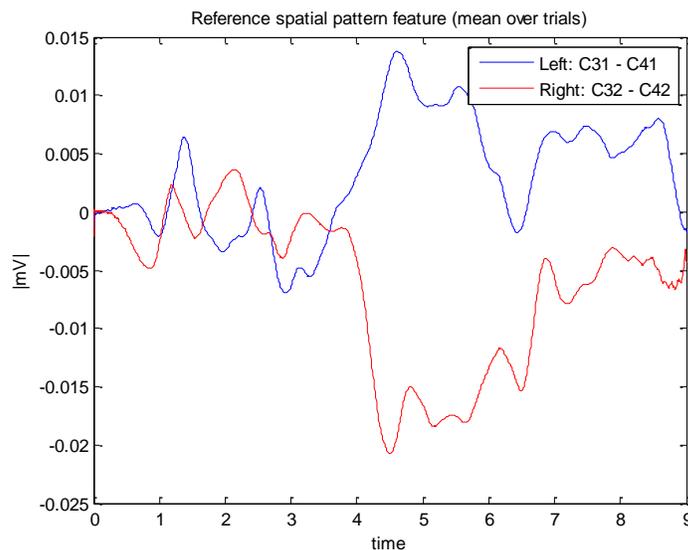


Fig. 4.5. Vectores de características obtenidos al restar los electrodos C3 y C4 para cada clase. Se observa la diferencia de patrones ERD/ERS entre clases.

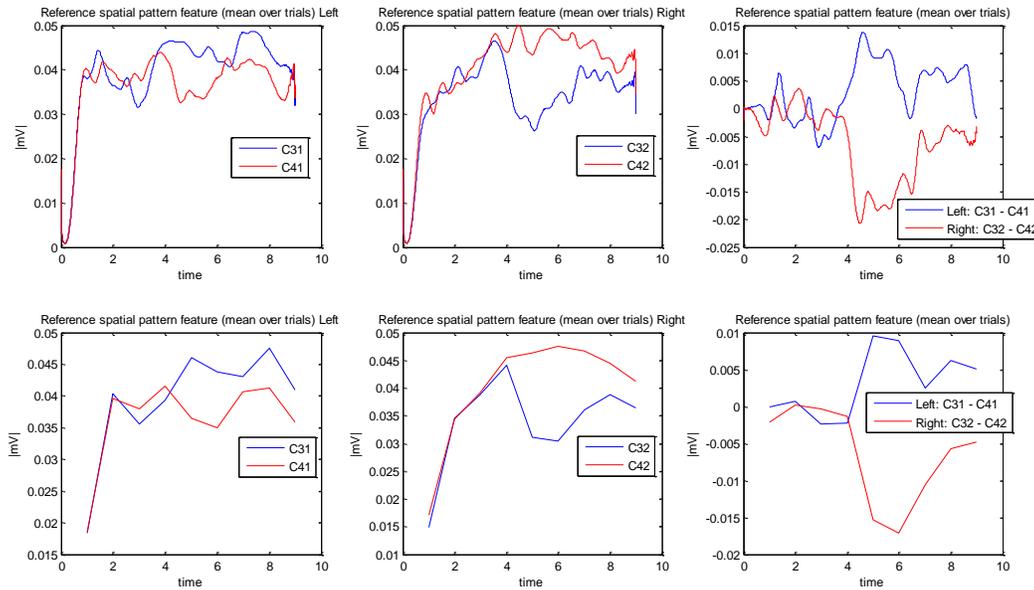


Fig. 4.6. Vectores de características antes y después de la reducción.

En la Fig. 4.6 se tiene la comparación entre los vectores de características antes y después de la reducción. Se observa que no se pierde el patrón ERD/ERS. Estos vectores se utilizan como entradas del clasificador.

4.2.1. Vector de características

El número de características que se tienen para cada muestra x_i de cada clase depende del tiempo de la prueba. Si se tienen pruebas de 7 s para cierto grupo de datos entonces se tendrán por cada x_i un vector con 7 elementos. Como se describió anteriormente, cada elemento es la media del número de muestras correspondientes a un segundo de la señal después de que se extrae el patrón espacial de la señal.

$$SF_i^{cec} = [m_{s1}, m_{s2}, m_{s2}, \dots, m_{sj}] \quad \text{Donde } j = 1, 2, \dots, s; \quad e = 3, 4; \quad c = 1, 2 \quad (4-2)$$

4.3. Extracción de características usando wavelets

Se desarrolló un algoritmo de extracción de características a partir de los coeficientes wavelets ya que estos proporcionan información sobre la distribución espacial y temporal de la señal. Mediante la DWT se obtiene una descomposición finita de la señal discreta en la que cada nivel

de descomposición está relacionado a una banda de frecuencias. El número de niveles o escala de descomposición se determinó basado en los componentes de la frecuencia dominante de la señal. Para una señal de 128 Hz se tiene una descomposición en frecuencia de acuerdo a la siguiente tabla [32].

Tabla 4-3. Frecuencias correspondientes a los primeros cuatro niveles de descomposición para una señal de 128 Hz y 250 Hz

<i>Niveles de descomposicion (125 Hz)</i>	<i>Frequency range (Hz)</i>	<i>Niveles de descomposicion (250 Hz)</i>	<i>Frequency range (Hz)</i>
D1	32 – 64	D1	62.5 – 125
D2	16 – 32	D2	31.3 – 62.5
D3	8 – 16	D3	15.6 – 31.3
A3	0 – 8	D4	7.8 – 15.6
–	–	A4	3.9 – 7.8

La extracción de características se realizó con tres diferentes wavelets: Coiflet, Daubechies y Symlet, con diferentes momentos de desvanecimiento. La elección de la wavelet fue en base a la experimentación y por algunas de las características que comparten con las daubechies que son las más utilizadas para aplicaciones de EEG de acuerdo a la literatura estudiada. Para la coiflet se eligieron las de 2, 3 y 4 momentos de desvanecimiento. Para la Daubechies las de 3, 4 y 5 y para la symlet las de 4, 5 y 6.

Los fenómenos relacionados al movimiento pueden ser vistos en las bandas de frecuencias alpha y beta, de lo anterior, para una señal de 128 Hz se puede descartar el nivel A3. Como las señales son filtradas entre 0.1 y 30 Hz el nivel D1 también se descarta, de modo que solo la extracción de características está basada en los niveles D3 y D2 para una señal de 128 Hz y en D4 y D3 para una señal de 250 Hz (Tabla 4-3) que reflejan la actividad μ y β por lo que solo se requiere de una descomposición de 3 niveles para frecuencias de muestreo de 125 y 128 Hz y de 4 niveles para una señal de 250 Hz. Es posible utilizar los coeficientes como características de la señal, sin embargo, extraer información estadística de los coeficientes ayuda a reducir la dimensión del vector de características. A diferencia de Qi *et. al* en [32], que usaron 2 características estadísticas para reducir la dimensión de los vectores, el vector de características está formado por las siguiente información estadística:

- 1) Media de los valores absolutos en cada sub-banda.
- 2) Potencia media de los coeficientes en cada sub-banda.
- 3) Desviación estándar de los coeficientes en cada sub-banda.

Las características 1) y 2) representan la distribución en frecuencia de la señal y la característica 3) refleja la cuanto cambia la distribución de frecuencia [27], [29] – [31].

4.3.1. Vector de características

Después de la descomposición wavelet se extraen las características descritas en la sección anterior. Como se tienen dos electrodos y se toman dos sub-bandas se forma el siguiente vector de características.

$$WF_i^{Cec} = \left[\frac{1}{m} \sum_{n=1}^m |D_s(n)|, \frac{1}{m} \sum_{n=1}^m (D_s(n))^2, std(D_s) \right] \quad (4-3)$$

Donde m es el número de coeficientes wavelets por sub-banda, n es el coeficiente y s es la sub-banda que se toma para el cálculo correspondiente, dependiendo si la señal es de 128 Hz o 250 Hz (Ver **Tabla 4-3**)

4.4. Clasificación

En esta sección se explica cómo se realizó el proceso de clasificación para ambos métodos de extracción de características. Los métodos o algoritmos implementados se explican en la sección 3.

4.4.1. Parámetros

El rendimiento o porcentajes de clasificación fueron estimados por ten-fold-cross-validation; tomando como parámetro el error de clasificación en cada validación. En este método se generan del total de las muestras ($2N$) de ambas clases 10 grupos de entrenamiento y 10 grupos para validación y el desempeño total es el promedio de los resultados independientes de cada corrida. Los arreglos que se describen a continuación son los alimentados a los

diferentes clasificadores. La Clase 1 es etiquetada como 0 y la Clase 2 como 1 para indicar las salidas requeridas para el entrenamiento en ambos casos. Para la SVM se utilizó la una RBF con un radio $\gamma = 1$.

4.4.2. Clasificación: patrones espaciales.

En la sección 4.2.1 Ec. (4-2) se tiene el vector de características que resulta del proceso de extracción. Para el caso de los patrones espaciales se obtiene una matriz (N, M) por electrodo y por clase, donde N es el número de pruebas y M es el número de características.

Se probaron dos arreglos de entradas al clasificador, el primero es concatenando los electrodos para cada clase y el segundo restando los electrodos [14] para cada clase, de la (4-2) se tiene lo siguiente, concatenando los vectores,

$$\begin{aligned} \text{Clase 1} &= [SF_i^{C31}, SF_i^{C41}] \\ \text{Clase 2} &= [SF_i^{C32}, SF_i^{C42}], \end{aligned} \quad (4-4)$$

restando los vectores,

$$\begin{aligned} \text{Clase 1} &= [SF_i^{C31} - SF_i^{C41}] \\ \text{Clase 2} &= [SF_i^{C32} - SF_i^{C42}] \end{aligned} \quad (4-5)$$

4.4.3. Clasificación: Características Wavelet

El vector de características descrito en la sección 4.3.2, generado a partir de los coeficientes wavelets (Ec. 4-3), y el porcentaje de clasificación fueron calculados a cada muestra con una ventana deslizante rectangular de un ancho de $2s$ con lo cual se captura suficiente información en frecuencia de la señal. Es decir, de la señal $x_i(n)$; $i = 1, \dots, N$, se extraen secciones de dos segundos con la ventana deslizante, cada vez que la ventana se desplaza una muestra, se extrae el vector de características para las N pruebas y estas son alimentadas a los clasificadores. La ventana se mueve entre $3 \leq s < (\frac{M}{Fs} - 2 * Fs)$. Los vectores de características se formaron al igual que las ecuación (4-1) pero con las características wavelets, y quedan de la siguiente forma,

$$\text{Clase 1} = [WF_i^{C31}, WF_i^{C41}]$$

$$\text{Clase 2} = [WF_i^{C32}, WF_i^{C42}] \quad (4-6)$$

4.5. Sumario

En este capítulo se han presentado los criterios tomados para el desarrollo de los algoritmos y se establecieron las formulaciones necesarias para la creación de un algoritmo fundamentado y factible, así como, la descripción de las bases de datos utilizadas para su evaluación. Se mencionó la metodología seguida para extracción de los vectores de características de ambos algoritmos. Por último se describió la manera en la cual estos vectores son clasificados con los diferentes métodos de clasificación y la métrica de evaluación del error. Es importante mencionar que la lógica del algoritmo no es complicada, por lo que permite un fundamento flexible que puede ser programado con diferentes lenguajes como: C++, Java, etc...

5. PRUEBAS Y RESULTADOS

5.1. Patrones Espaciales

Como se mencionó anteriormente, los vectores de características extraídos por patrones espaciales fueron extraídos para todo el intervalo de la prueba. Debido a que la información que se extrae está relacionada al cálculo del ERD/ERS, este vector representa la información de la prueba para el tiempo total correspondiente a cada prueba. En la Tabla 5-1 se muestra una comparativa de los errores de clasificación para los dos tipos de vectores de características. Se puede observar que el restar los vectores de características ($C3 - C4$) mejora el resultado de la clasificación en la mayoría de los sujetos. En la Fig. 4.6 se tienen los patrones y características extraídas para S1, y se puede observar que al restar los electrodos para una clase se logra maximizar el umbral de separación entre las clases. En el apéndice A1 se pueden observar las gráficas de referencia para cuando se restan los vectores de características para los nueve diferentes sujetos. Se puede notar que el error incrementa mientras menos separabilidad se vea en las gráficas.

Tabla 5-1. Errores de clasificación de las características de patrones espaciales.

<i>Sujeto</i>	Concatenando			Restando		
	<i>FLDA</i>	<i>QDA</i>	<i>SVM</i>	<i>FLDA</i>	<i>QDA</i>	<i>SVM</i>
<i>S1</i>	0.1714	0.2357	0.3929	0.1643	0.2357	0.2786
<i>S2</i>	0.2857	0.3500	0.5000	0.2357	0.4143	0.4643
<i>S3</i>	0.3929	0.5000	0.4857	0.3643	0.4571	0.5214
<i>S4</i>	0.4071	0.5286	0.4286	0.3643	0.5143	0.5286
<i>S5</i>	0.2567	0.2133	0.3233	0.2467	0.2400	0.2500
<i>S6</i>	0.4611	0.4344	0.4488	0.5143	0.4754	0.4795
<i>S7</i>	0.4627	0.5056	0.4701	0.4403	0.4925	0.4571
<i>S8</i>	0.5019	0.4888	0.4478	0.4813	0.4944	0.4552
<i>S9</i>	0.5187	0.4757	0.4981	0.5112	0.5056	0.5336
Promedio	0.38424	0.41468	0.44392	0.36916	0.42548	0.44092

Dentro de los algoritmos de clasificación el que mejor clasifica es el FLDA aunque sea el método menos complejo, lo que hace que el algoritmo sea rápido para el entrenamiento y la

clasificación de nuevas entradas al sistema. También se puede ver que para cuando las características son menos linealmente separables; que es el caso de S3, S4, S6-S9, si se requiere de la implementación de un SVM, ya que esta mapea las entradas a un espacio de mayor dimensión donde sean más separables (ver Tabla 5-1).

5.2. Características wavelet

Para las características wavelets solo se trabajó con el vector de que resulta de la resta de los vectores de características extraídos para cada electrodo. Esta decisión se tomó en base al resultado obtenido anteriormente con los patrones espaciales. De la Tabla 5-2 a la Tabla 5-4 se tienen los resultados para los nueve diferentes sujetos. Se puede ver que la SVM y el QDA no obtuvieron los mejores resultados en la clasificación. En general FLDA obtuvo los mejores resultados de clasificación.

Tabla 5-2. Tasa de error obtenida para la base de datos 1. Para diferentes tipos de wavelet y clasificadores.

Wavelet	S1		
	LDA	QDA	SVM
<i>coif2</i>	0.1286	0.1500	0.2332
<i>coif3</i>	0.1286	0.1571	0.2424
<i>coif4</i>	0.1214	0.1286	0.2464
<i>db3</i>	0.1286	0.1643	0.2183
<i>db4</i>	0.1214	0.1357	0.2361
<i>db5</i>	0.1286	0.1286	0.2218
<i>sym3</i>	0.1286	0.1643	0.2258
<i>sym4</i>	0.1357	0.1500	0.2251
<i>sym5</i>	0.1357	0.1500	0.2319
<i>sym6</i>	0.1357	0.1571	0.2278
<i>sym7</i>	0.1286	0.1571	0.2335

Para el caso de las wavelets se observa que la *coif4*, *db4* y *db5* son las que mejor se asemejan a la señal y por lo tanto logran obtener mejor información de la señal para los diferentes sujetos. De los sujetos con los que mejores resultados se obtienen son S1, S2, S4, S5. Con un promedio de error sobre todas las wavelets con LDA para S1 de 0.13, S2 de 0.92, S4 de 0.19 y S5 de 0.22.

Tabla 5-3. Tasa de error obtenida para la base de datos 2. Para diferentes tipos de wavelet y clasificadores.

Wavelet	S2			S3			S4		
	LDA	QDA	SVM	LDA	QDA	SVM	LDA	QDA	SVM
<i>coif2</i>	0.0909	0.0909	0.1941	0.25	0.3	0.48	0.2167	0.25	0.3464
<i>coif3</i>	0.0909	0.125	0.1924	0.2833	0.25	0.45	0.15	0.2333	0.348
<i>coif4</i>	0.0909	0.1023	0.191	0.2833	0.3	0.4521	0.1833	0.2333	0.3591
<i>db3</i>	0.1023	0.1023	0.1892	0.2833	0.2667	0.3833	0.2	0.2333	0.3382
<i>db4</i>	0.1023	0.1364	0.1807	0.2833	0.2833	0.4859	0.2	0.2	0.3167
<i>db5</i>	0.0795	0.1023	0.1847	0.2833	0.2667	0.4649	0.2	0.2667	0.3447
<i>sym4</i>	0.1136	0.1023	0.1785	0.2833	0.3167	0.45	0.1833	0.2167	0.3377
<i>sym5</i>	0.0795	0.1023	0.1817	0.2833	0.3167	0.4667	0.2	0.25	0.3361
<i>sym6</i>	0.0795	0.125	0.1854	0.2833	0.2833	0.4801	0.2	0.2333	0.3313

Tabla 5-4. Tasa de error obtenida para la base de datos 3. Para diferentes tipos de wavelet y clasificadores.

Wavelet	S5			S6			S7		
	LDA	QDA	SVM	LDA	QDA	SVM	LDA	QDA	SVM
<i>coif2</i>	0.2267	0.2500	0.2567	0.2971	0.3217	0.3981	0.2201	0.2444	0.3781
<i>coif3</i>	0.2200	0.2567	0.2667	0.2930	0.3176	0.3879	0.2164	0.2369	0.3764
<i>coif4</i>	0.2200	0.2700	0.2667	0.2930	0.3217	0.3913	0.2220	0.2332	0.3787
<i>db3</i>	0.2333	0.2700	0.2900	0.2971	0.3156	0.4099	0.2257	0.2444	0.3777
<i>db4</i>	0.2267	0.2533	0.2633	0.3074	0.3094	0.4010	0.2313	0.2500	0.3786
<i>db5</i>	0.2167	0.2567	0.2500	0.2725	0.3074	0.3927	0.2295	0.2369	0.3753
<i>sym4</i>	0.2400	0.2600	0.2633	0.3053	0.3135	0.3995	0.2295	0.2463	0.3769
<i>sym5</i>	0.2200	0.2400	0.2200	0.2889	0.3074	0.4037	0.2220	0.2425	0.3815
<i>sym6</i>	0.2467	0.2800	0.2400	0.3012	0.3135	0.3914	0.2127	0.2351	0.3728

En la Fig. 5.1 se graficó el promedio para cada sujeto sobre todas las wavelets que se usaron para la extracción de características. Se puede observar que el resultado de la clasificación depende mucho de que tan entrenado esté el sujeto para realizar la actividad requerida. Tal fue el caso de los sujetos antes mencionados. Para las bases de datos propias, a pesar de que no se contó con el entrenamiento necesario se obtuvieron tasas de hasta 0.3 de error para S8. En el Apéndice A2 se pueden observar las gráficas de error de clasificación en el tiempo de los sujetos S1, S2, S4 y S5, para cada tipo de wavelet en las que se observan las diferencias que existen entre los diferentes clasificadores para un procesamiento on-line. Se observa que el

comportamiento a partir de $s = 3$ es el que se espera, ya que el estímulo se presenta a partir de ese momento.

Tabla 5-5. Tasa de error obtenida para la base de datos propia. Para diferentes tipos de wavelet y clasificadores

Wavelet	S8			S9		
	LDA	QDA	SVM	LDA	QDA	SVM
<i>coif2</i>	0.2946	0.3080	0.4327	0.4235	0.4388	0.4566
<i>coif3</i>	0.2991	0.3170	0.4307	0.4184	0.4209	0.4971
<i>coif4</i>	0.2991	0.3036	0.4267	0.4286	0.4337	0.4388
<i>db3</i>	0.2902	0.3125	0.4401	0.4362	0.4388	0.4864
<i>db4</i>	0.3036	0.3348	0.4388	0.4439	0.449	0.4566
<i>db5</i>	0.2946	0.3170	0.4332	0.4388	0.4413	0.4796
<i>sym4</i>	0.2991	0.3080	0.4328	0.4184	0.4439	0.4816
<i>sym5</i>	0.2991	0.3125	0.4336	0.4158	0.4388	0.4962
<i>sym6</i>	0.3080	0.3036	0.4279	0.4235	0.4311	0.4643
<i>promedio</i>	0.2986	0.3130	0.4329	0.4275	0.4374	0.4730

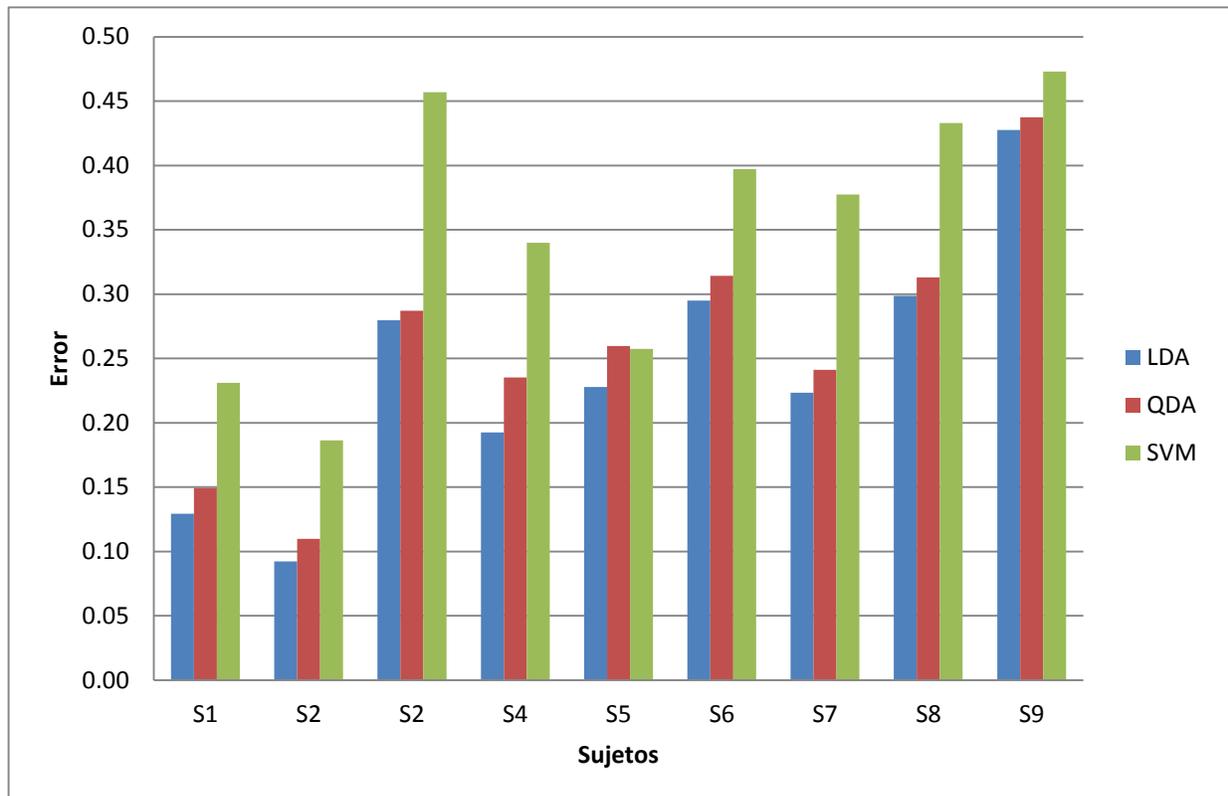


Fig. 5.1. Promedio de error de todos los sujetos para los diferentes clasificadores sobre todas la wavelets.

5.3. Sumario

Los resultados anteriores mostraron una clara tendencia entre ellos. Para el algoritmo basado en los patrones espaciales el clasificador con el que mejores resultados se obtuvo, como promedio de todos los sujetos, fue el FLDA con las características formadas de la resta de los vectores extraídos de cada electrodo. En el caso de las características wavelets, del resultado anterior, se decidió usar esa misma estructura del vector de características. De nuevo se observó que el clasificador FLDA es el mejor. Por otro lado de entre las wavelets utilizadas la db5 fue la que en general, dio el mejor resultado junto con la coif2 lográndose errores de 0.12 y 0.7 respectivamente.

6. CONCLUSIONES Y TRABAJO A FUTURO

La representación tiempo frecuencia por patrones espaciales sintetizados resulta ser la más adecuada para la extracción de los patrones en la señal, ya que refleja el comportamiento ERD/ERS de manera sencilla y sin mucho costo computacional.

Partiendo de esta representación; la extracción de características propuesta, basada en la reducción de dimensión del patrón espacial en la banda de interés, refleja tasas de error de hasta 0.12 con el algoritmo simplificado descrito.

La diferencia entre los métodos de clasificación es notable con S1 donde FLDA es mejor que un QDA o inclusive una SVM. Este es un resultado alentador ya que FLDA requiere de menos costo computacional, de modo que puede ser utilizado para una clasificación en línea.

El conjunto de las tres características estadísticas aplicadas a los coeficientes wavelets extraídos con la wavelet Coiflet ha demostrado, al igual que la Daubechies, que también son adecuadas para este análisis y en algunos casos se ha mejorado el porcentaje de clasificación.

Un algoritmo de selección de mejores características podría mejorar el desempeño del algoritmo, pero requeriría un costo computacional mayor. Sin embargo, en este caso no es necesario ya que los vectores de características son pequeños y la dimensión no es un problema de acuerdo a lo experimentado.

Como trabajo a futuro se propone aplicar las técnicas de extracción utilizadas para un procesamiento on-line en tiempo real. En el caso del algoritmo basado en wavelets, tomar como el clasificador, el cual obtuvo un mejor entrenamiento de acuerdo a las características a un cierto instante o periodo de tiempo, así como proponer una métrica de medición del error para toma de decisiones a la salida del clasificador.

REFERENCIAS

- [1]. Lotte F., Congedo M., L'écuyer A., Lamarche F., Arnaldi B. A review of classification algorithms for EEG-based brain–computer interfaces. *J. Neural Eng.* 4 (2007) R1–R13.
- [2] Fisch J. B., Spehlmann R. *Fisch and Spehlmann's EEG primer: basic principles of digital and analog EEG.* 1999. Elsevier.
- [3] Bashashati A., et Al. A survey of signal processing algorithms in brain–computer interfaces based on electrical brain signals. *J. Neural Eng.* 4 (2007) R32–R57.
- [4] McFarland D J, Anderson C W, Muller K-R, Schlogl A and Krusienski D J. (2006). BCI meeting 2005. Workshop on BCI signal processing: feature extraction and translation *IEEE Trans. Neural Syst. Rehabil. Eng.*
- [5] Pfurtscheller G., Lopes da Silva F H. 1999. Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clin. Neurophysiol* 110 1842–57.
- [6] Allison B. Z., Pineda J A. (2003). ERPs evoked by different matrix sizes: implications for a brain–computer interface (BCI) system *IEEE Trans. Neural Syst. Rehabil. Eng.* 11 110-3.
- [7] Kaper M, Meinicke P, Grossekhoefer U, Lingner T and Ritter H. 2004. BCI competition 2003–data set IIIb: support vector machines for the p300 speller paradigm *IEEE Trans. Biomed. Eng.* 51 1073–6.
- [8] Pfurtscheller G, Neuper C, Flotzinger D., Pregenzer M. 1997. EEG-based discrimination between imagination of right and left hand movement. *Electroencephalogr. Clin. Neurophysiol.* 103 642–51.
- [9] Chiappa S., Bengio S. 2004. HMM and IOHMM modeling of EEG rhythms for asynchronous BCI systems. *European Symposium on Artificial Neural Networks ESANN.*
- [10] Millan J. R., Mouriño J. 2003. Asynchronous BCI and local neural classifiers: an overview of the adaptive brain interface project. *IEEE Trans. Neural Syst. Rehabil. Eng.* 11 159–61.

- [11] Pfurtscheller G, Neuper C, Schlogl A. and Lugger K. 1998. Separability of EEG signals recorded during right and left motor imagery using adaptive autoregressive parameters. *IEEE Trans. Rehabil. Eng.* 6 316–25.
- [12] Wang T, Deng J, and He B. (2004). Classifying EEG-based motor imagery tasks by means of time-frequency synthesized spatial patterns. *Clin. Neurophysiol.* 115 2744–53
- [13] Semmlow L. J, *Biosignal and Biomedical Image Processing MATLAB-Based Applications.* (2004). Marcel Dekker, Inc.
- [14] Huaiyu X, Jian L, Ruidan S, Erpeng Z. (2009). Feature Extraction and Classification of EEG for Imaging Left-right Hands Movement. *Integrated Circuit Applied Software Lab.* Shenyang, China.
- [15] Duda R. *Pattern classification, 2nd E.* Wiley Interscience. 2000.
- [16] Pfurtscheller G., Neuper C., Flotzinger D., Pregenzer M. 1997 EEG-based discrimination between imagination of right and left hand movement *Electroenceph. Clin. Neurophysiol.* 103 642–51
- [17]. Wang T., He B. An efficient rhythmic component expression and weighting synthesis strategy for classifying motor imagery EEG in a brain–computer interface. *J. Neural Eng.* 1 (2004) 1–7
- [18]. Wang T., Deng J., He B. Classifying EEG-based motor imagery tasks by means of time–frequency synthesized spatial patterns. *Clinical Neurophysiology* 115 (2004) 2744–2753
- [19]. Durka, P. *Matching Pursuit and Unification in EEG Analysis.* 2007. Artech House, Inc. Norwood, MA.
- [20]. Scholkopf, B. Smola, A. J. *Learning with kernels: Support Vector Machines, Regularization, Optimization, and Beyond.* 2002. The MIT Press. London, England.
- [21]. Frank R. Kschischang. *The Hilbert Transform.* The Edward S. Rogers Sr. Department of Electrical and Computer Engineering. University of Toronto. October 22, 2006.

- [22]. Mallat, S. A wavelet tour of signal processing: the sparse way. Academic Press. 2009.
- [23]. Pfurtscheller G. (2000) Spatiotemporal ERD/ERS patterns during voluntary movement and motor imagery. *Clinical Neurophysiology at the Beginning of the 21st Century (Supplements to Clinical Neurophysiology Vol. 53).*, pp 196-198.
- [24]. Graimann B., Pfurtscheller G. (2006) Quantification and visualization of event-related changes in oscillatory brain activity in the time–frequency domain. *Progress in Brain Research*, Vol. 159., pp 79-97.
- [25]. Pfurtscheller G., Aranibar A. (1979). Evaluation of event-related desynchronization (erd) preceding and following voluntary self-paced movement. *Electroencephalography and clinical Neurophysiology*, 46 : 138--146.
- [26]. Pfurtscheller G., Lopes da Silva F.H., (1999) Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clinical Neurophysiology* 110 : 1842-1857.
- [27]. Subasi, A. EEG signal classification using wavelet feature extraction and a mixture of expert model. *Expert Systems with Applications* 32 (2007) 1084–1093.
- [28]. Alarcon Aquino, V., *Anomaly Detection and Prediction in Communication Networks Using Wavelet Transforms*, PhD Thesis, Imperial College London, London UK, 2003.
- [29]. Subasi, A. Application of adaptive neuro-fuzzy inference system for epileptic seizure detection using wavelet feature extraction. *Computers in Biology and Medicine* 37 (2007) 227–244.
- [30]. A. Kandaswamy, C. Sathish Kumar, Rm. Pl. Ramanathanc, S. Jayaramana. Neural classification of lung sounds using wavelet coefficients. *Computers in Biology and Medicine* 34 (2004) 523–537.
- [31]. Ubeyli, E., D. Probabilistic neural networks combined with wavelet coefficients for analysis of electroencephalogram signals. *Expert Systems*, (2009), 26-2 : 147-159.

- [32]. Qi Xu, Hui Zhou, Yongji Wang, Jian Huang. Fuzzy support vector machine for classification of EEG signals using wavelet-based features. *Medical Engineering & Physics* 31 (2009) 858–865.
- [33]. Burrus C., Gopinath A., Guo H. *Introduction to Wavelets and Wavelet Transforms: A Primer*. Prentice Hall. 1998.
- [34]. Blue T. y Lebron J. Linear time-frequency analysis II: Wavelet type representations, En “Time frequency analysis”. Ed. F. Hlawatsch y F. Auger. 93-109. Wiley. London, UK.
- [35] Abe, S. *Advances in pattern recognition: Support Vector Machines for pattern Classification*. Springerlink. 2005. Cap. 2: 15-50.
- [36] Cristianini N., Shawe-Taylor J. *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge University Press. 2000
- [37] Masujima, M. *Applied Mathematical Methods in Theoretical Physics*. Wiley-VCH. 2005

APÉNDICE A: FIGURAS

A1: Patrones espaciales

Se muestra la distribución de los vectores de características extraídos con los patrones espaciales. Por cada sujeto, en la figura de la izquierda se observa el vector de características formado de concatenar los vectores de cada electrodo para ambas clases. En la figura de la derecha se tiene la resta de los vectores de cada electrodo para ambas clases.

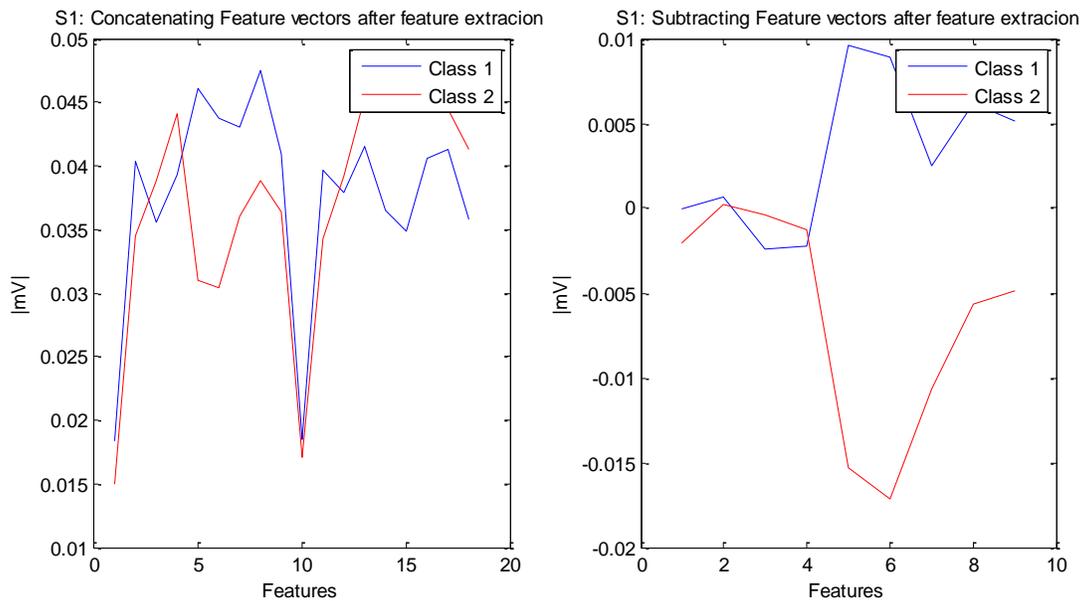


Fig. A.1. Vector de características extraído por medio de patrones espaciales para S1.

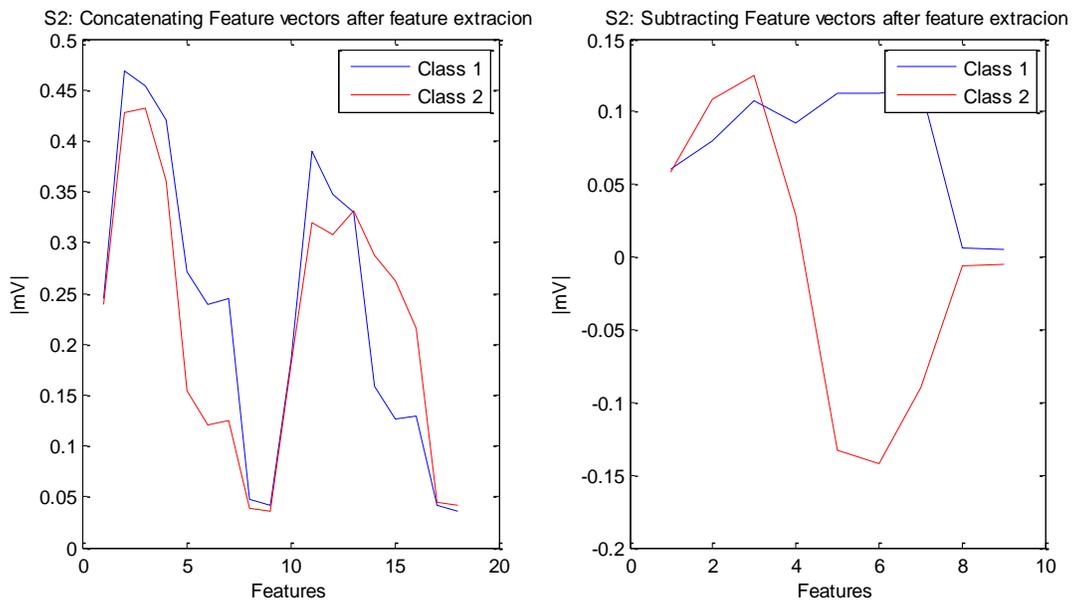


Fig. A.2. Vector de características extraído por medio de patrones espaciales para S2.

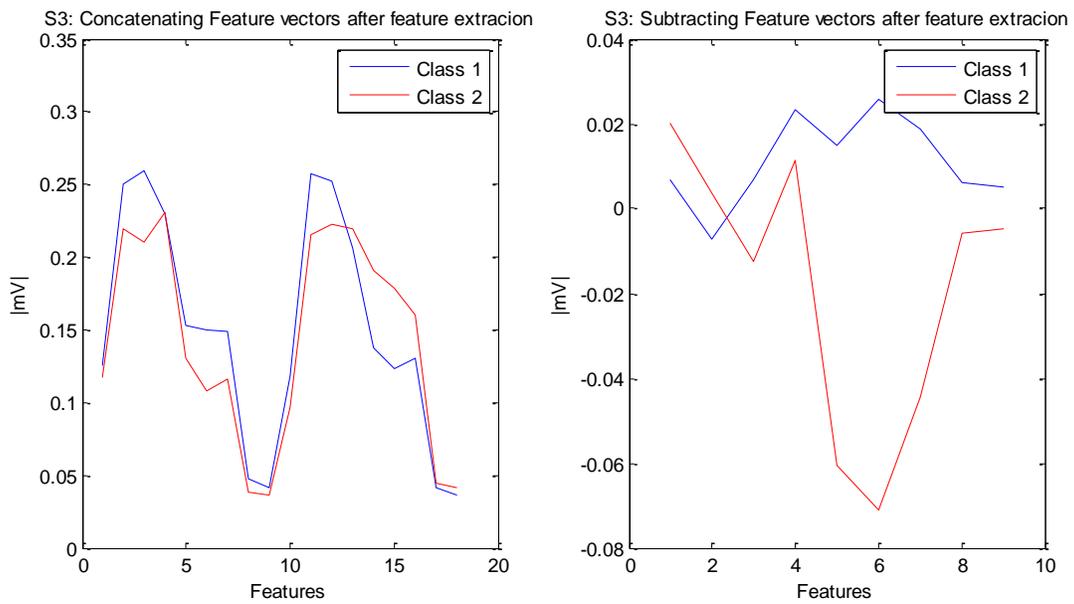


Fig. A.3. Vector de características extraído por medio de patrones espaciales para S3.

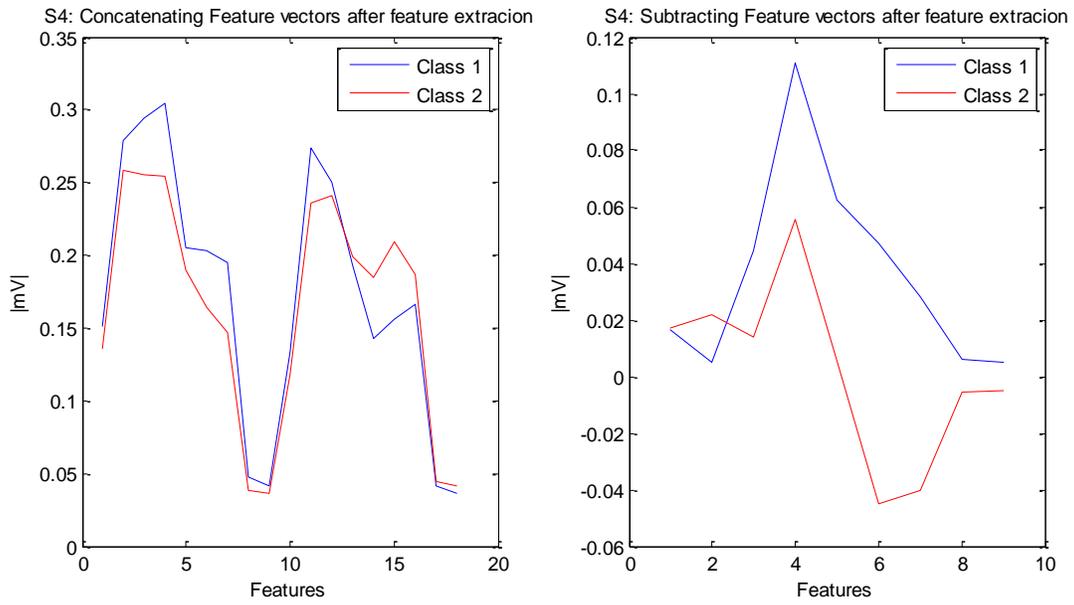


Fig. A.4. Vector de características extraído por medio de patrones espaciales para S4.

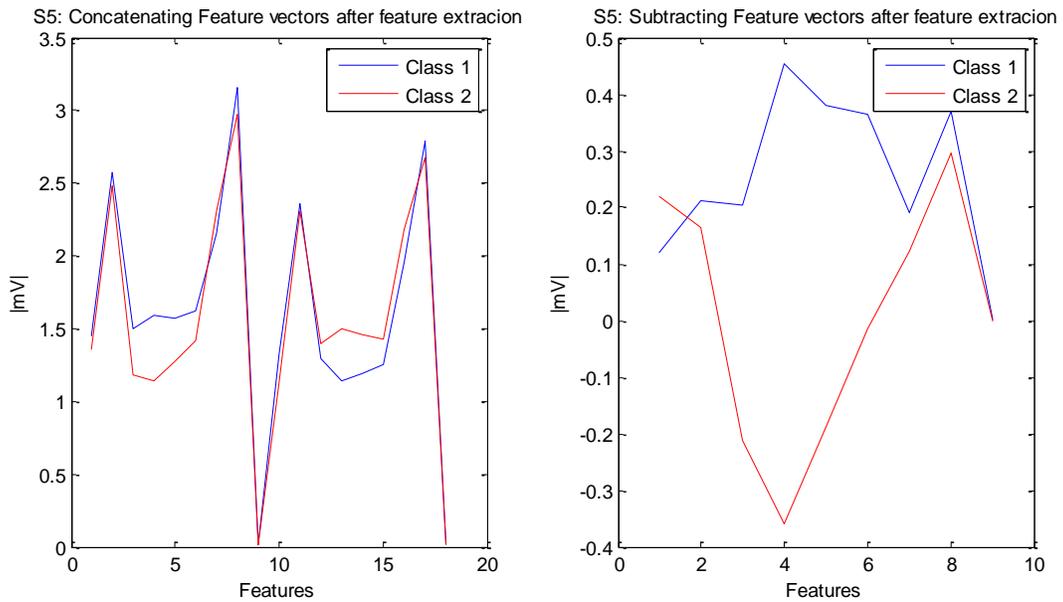


Fig. A.5. Vector de características extraído por medio de patrones espaciales para S5.

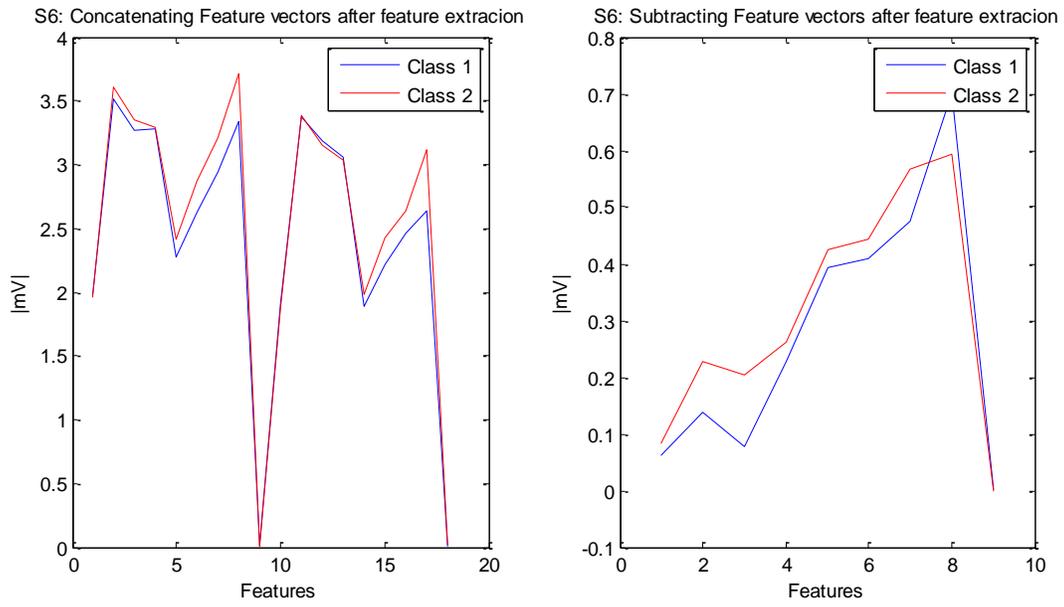


Fig. A.6. Vector de características extraído por medio de patrones espaciales para S6.

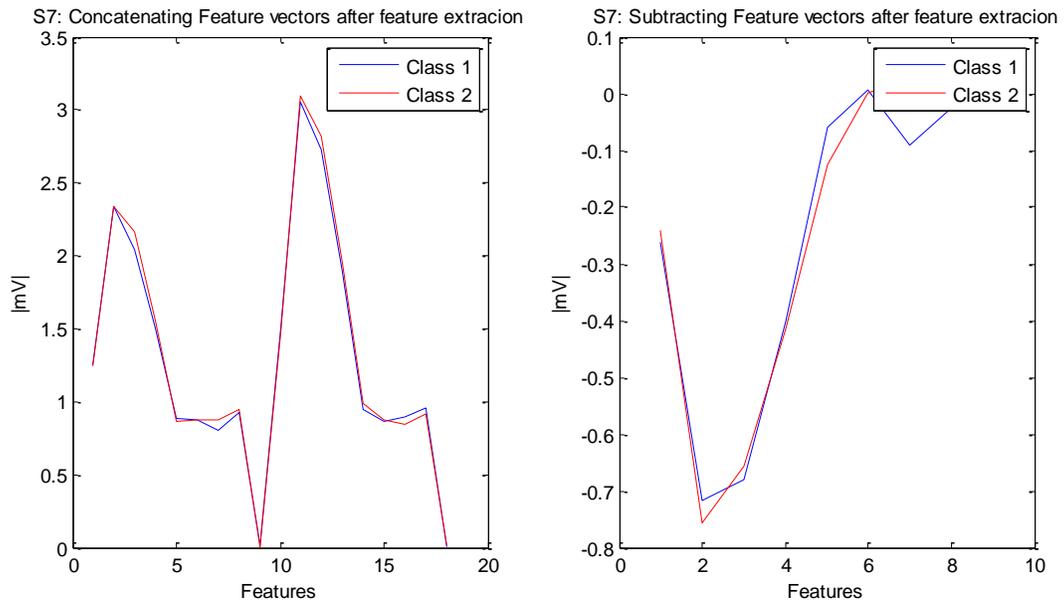


Fig. A.7. Vector de características extraído por medio de patrones espaciales para S7.

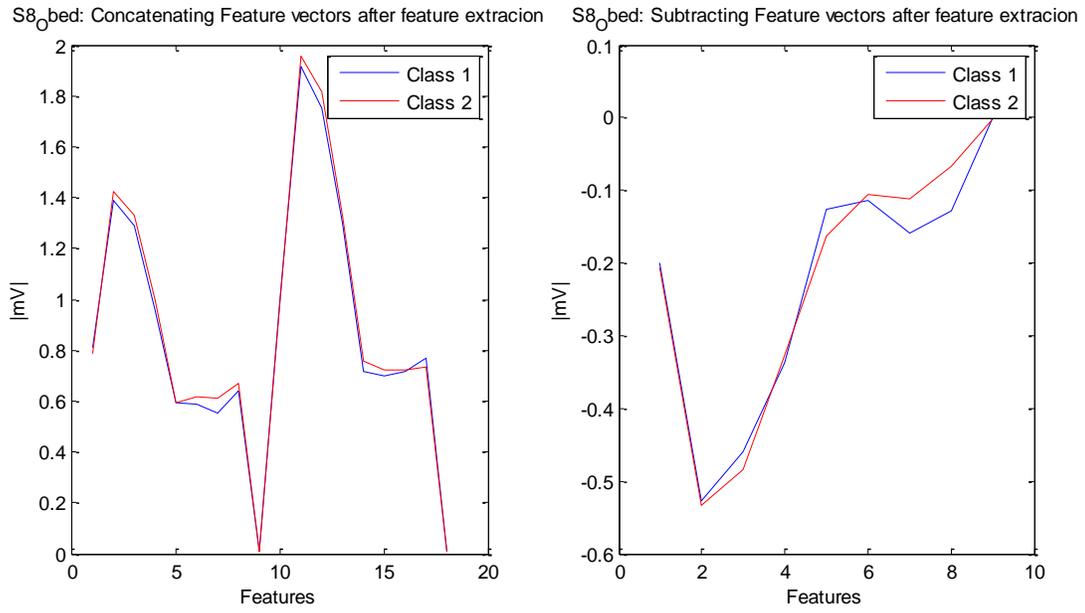


Fig. A.8. Vector de características extraído por medio de patrones espaciales para S8.

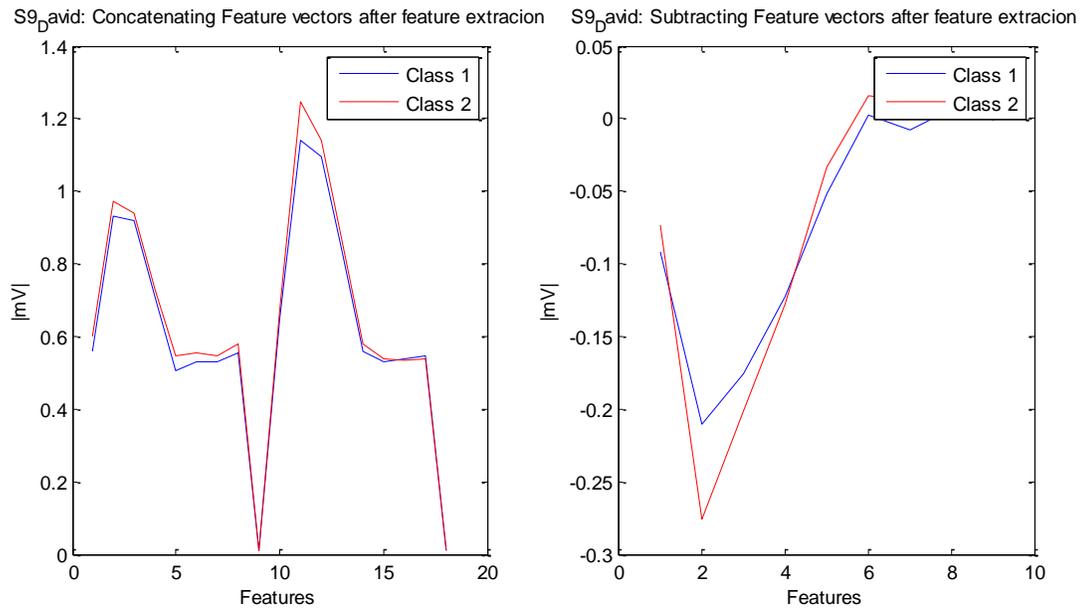


Fig. A.9. Vector de características extraído por medio de patrones espaciales para S9.

A2: Wavelet Features

En este apéndice se muestra el error de clasificación con respecto al tiempo de los patrones obtenidos de la extracción basada en wavelets. Se muestran las diferentes graficas obtenidas por sujeto para cada una de las diferentes wavelets.

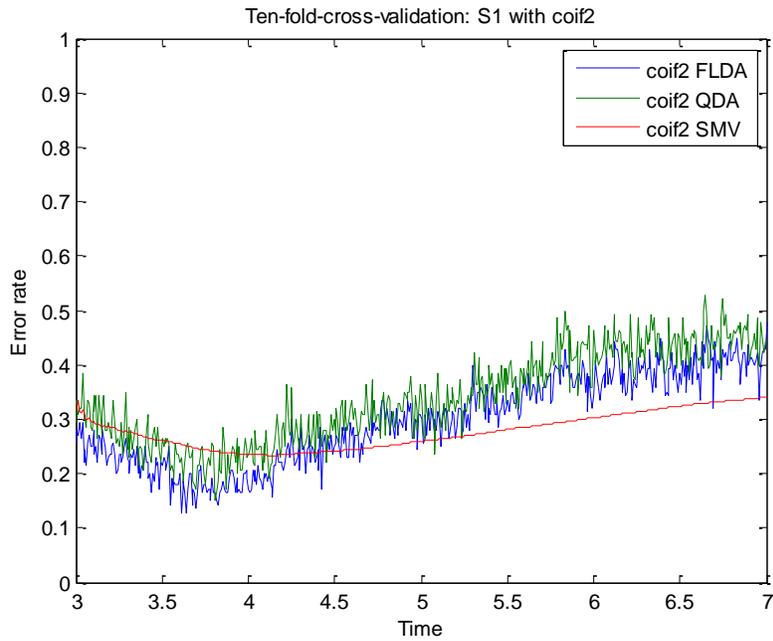


Fig. A.10. Error obtenido en el tiempo para S1 con coif2.

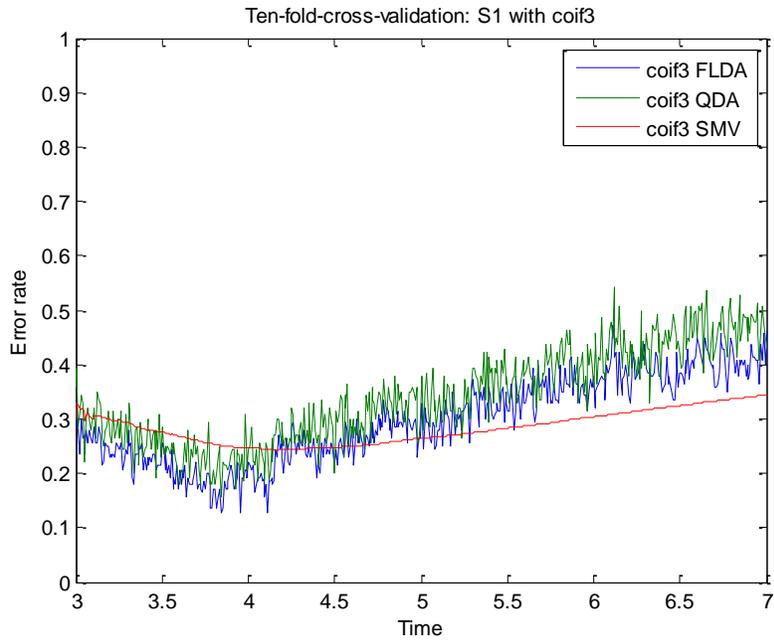


Fig. A.11. Error obtenido en el tiempo para S1 con coif3.

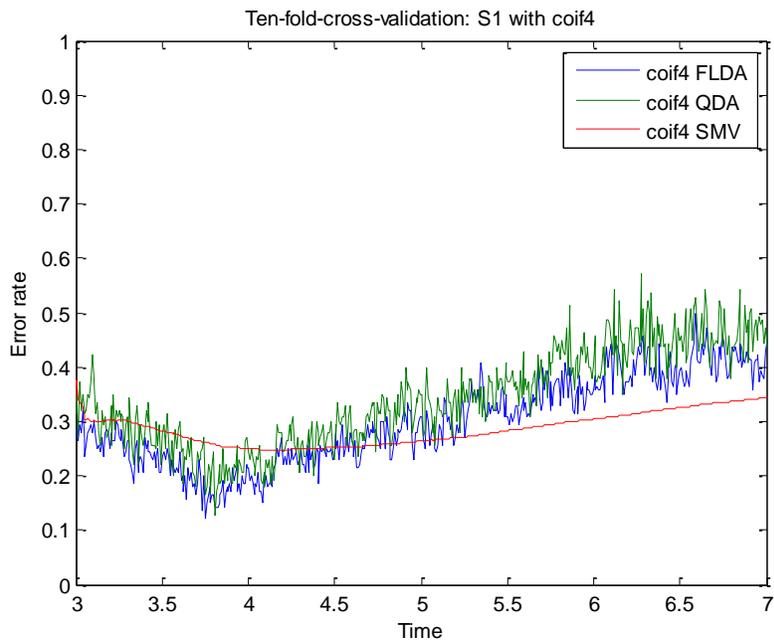


Fig. A.12. Error obtenido en el tiempo para S1 con coif4.

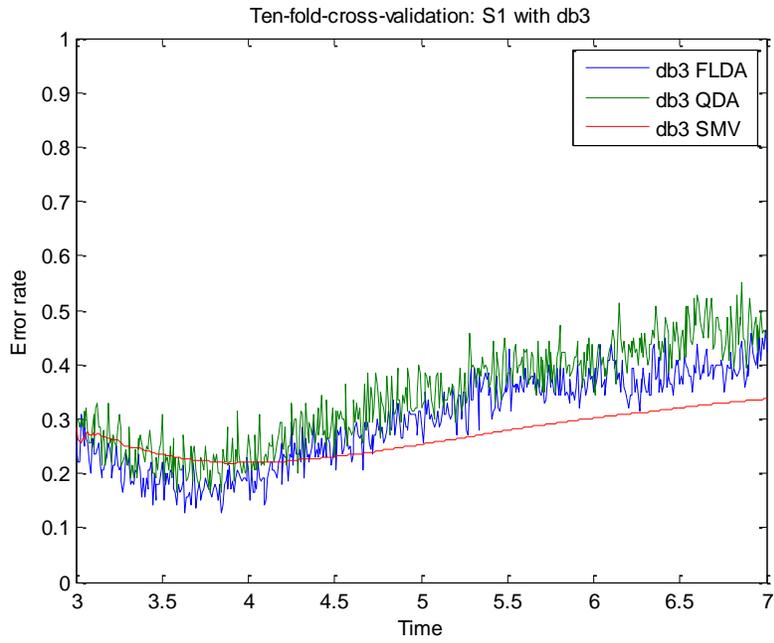


Fig. A.13. Error obtenido en el tiempo para S1 con db3.

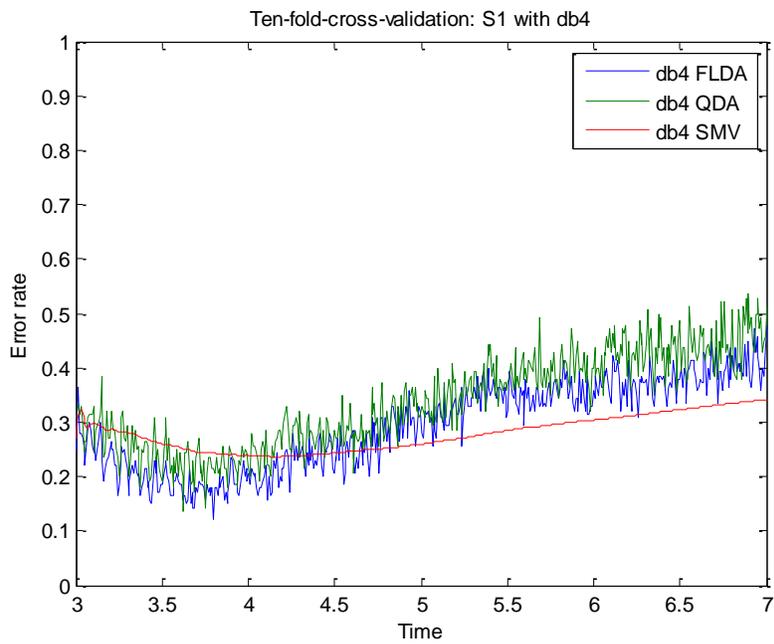


Fig. A.14. Error obtenido en el tiempo para S1 con db4.

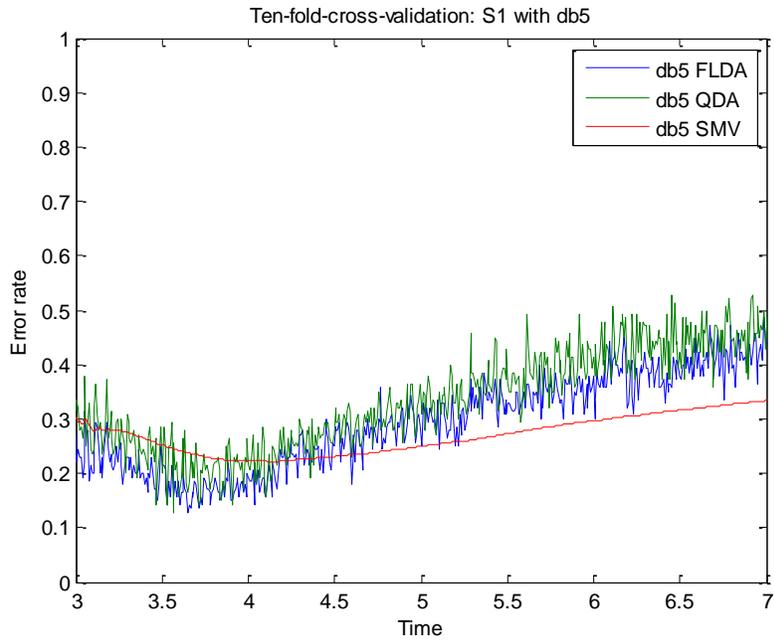


Fig. A.15. Error obtenido en el tiempo para S1 con db5.

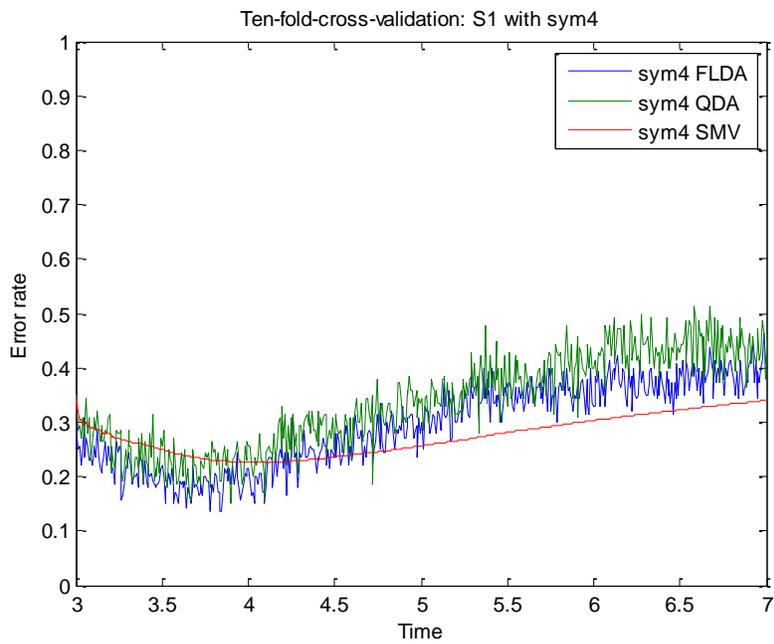


Fig. A.16. Error obtenido en el tiempo para S1 con sym4.

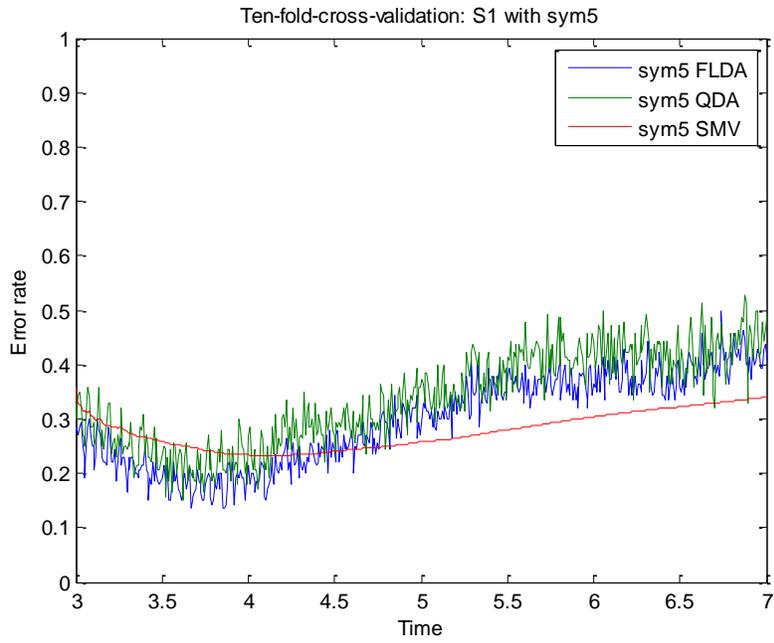


Fig. A.17. Error obtenido en el tiempo para S1 con sym5.

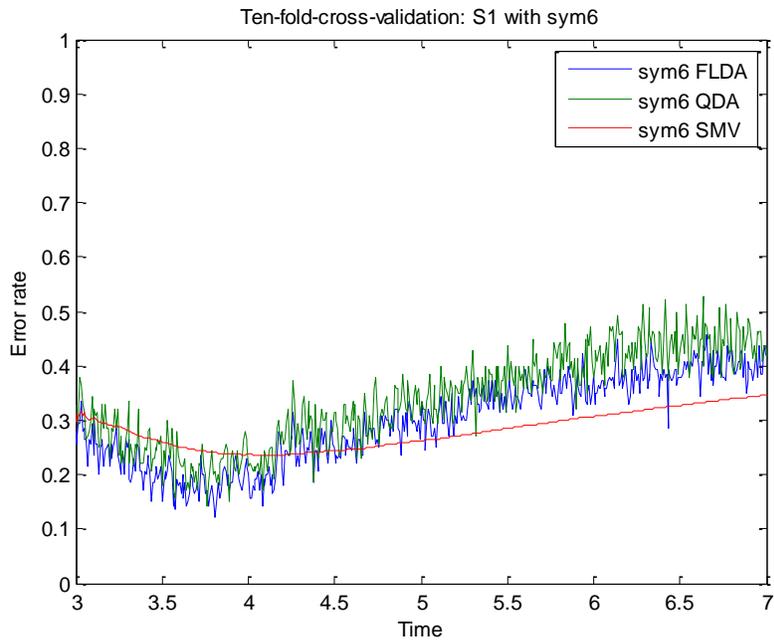


Fig. A.18. Error obtenido en el tiempo para S1 con sym6.

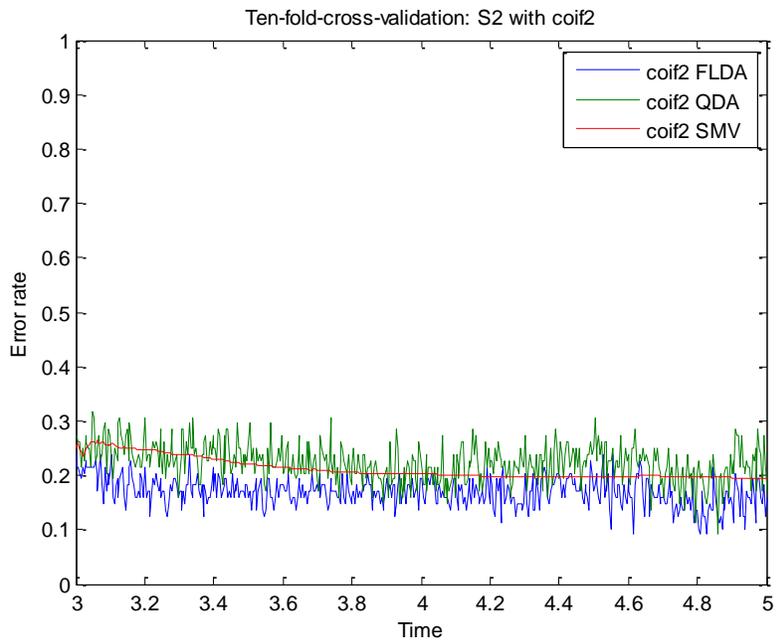


Fig. A.19. Error obtenido en el tiempo para S2 con coif2.

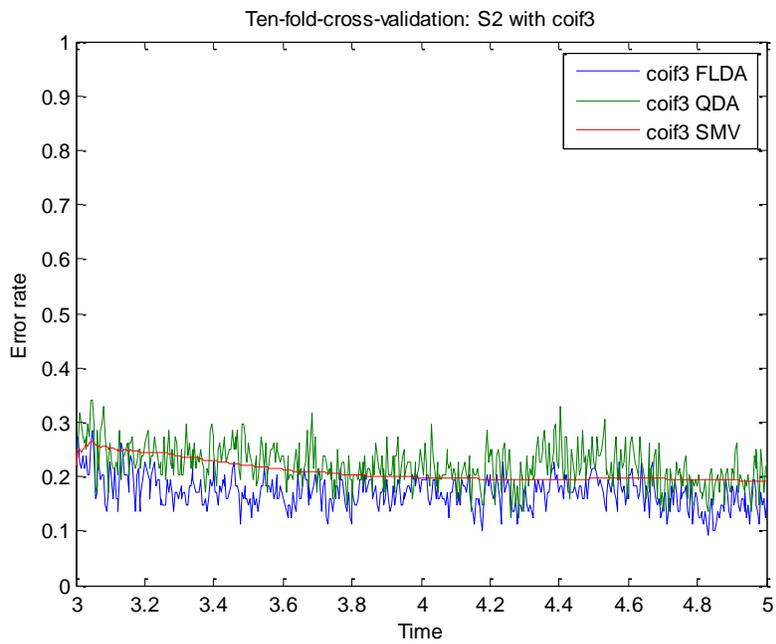


Fig. A.20. Error obtenido en el tiempo para S2 con coif3.

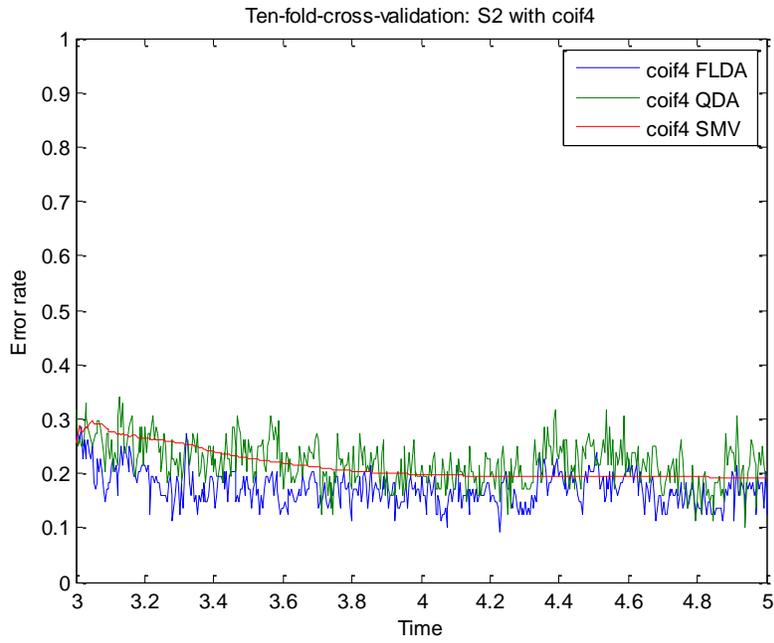


Fig. A.21. Error obtenido en el tiempo para S2 con coif4.

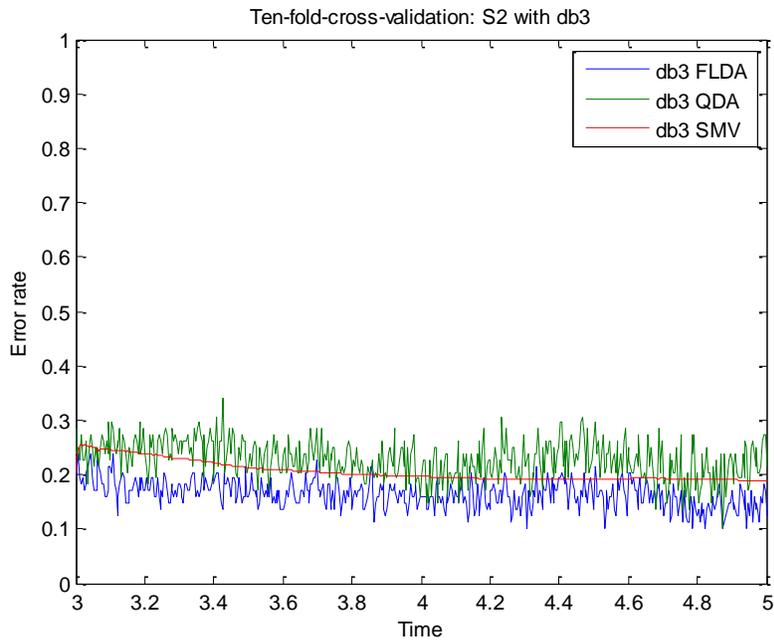


Fig. A.22. Error obtenido en el tiempo para S2 con db3.

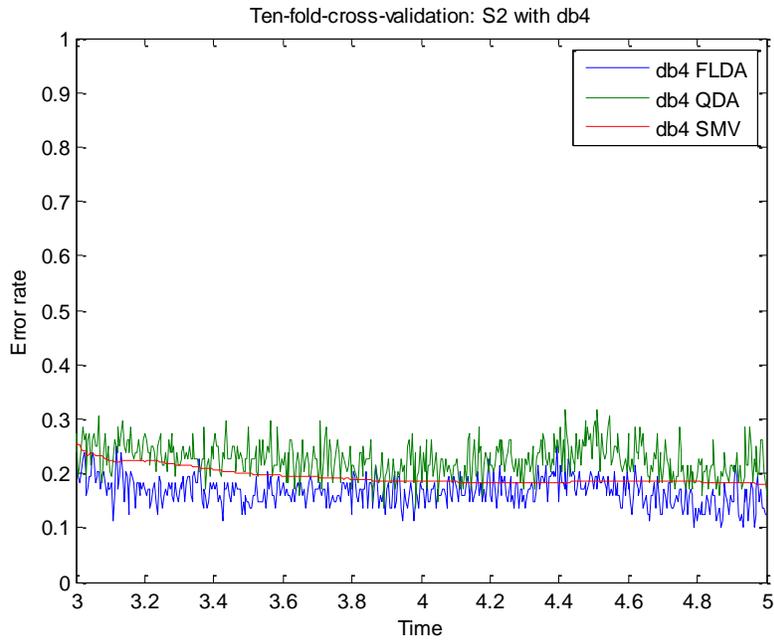


Fig. A.23. Error obtenido en el tiempo para S2 con db4.

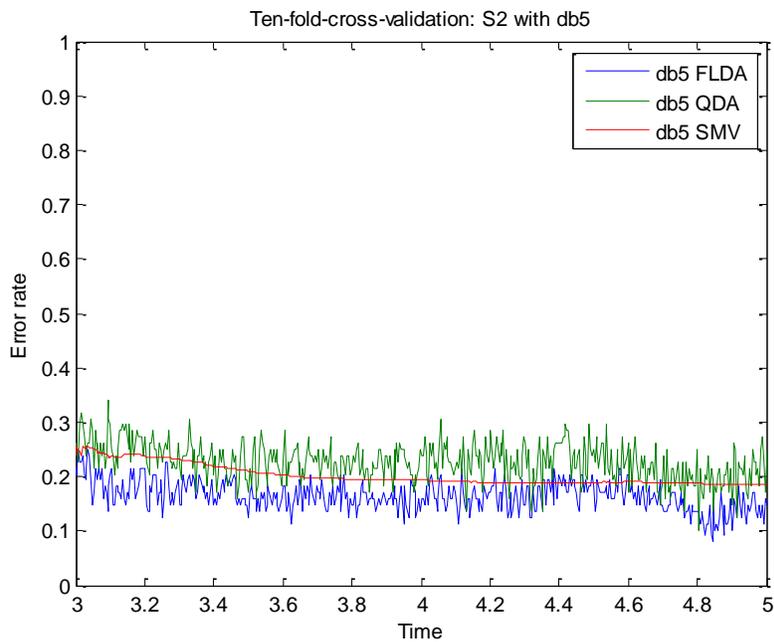


Fig. A.24. Error obtenido en el tiempo para S2 con db5.

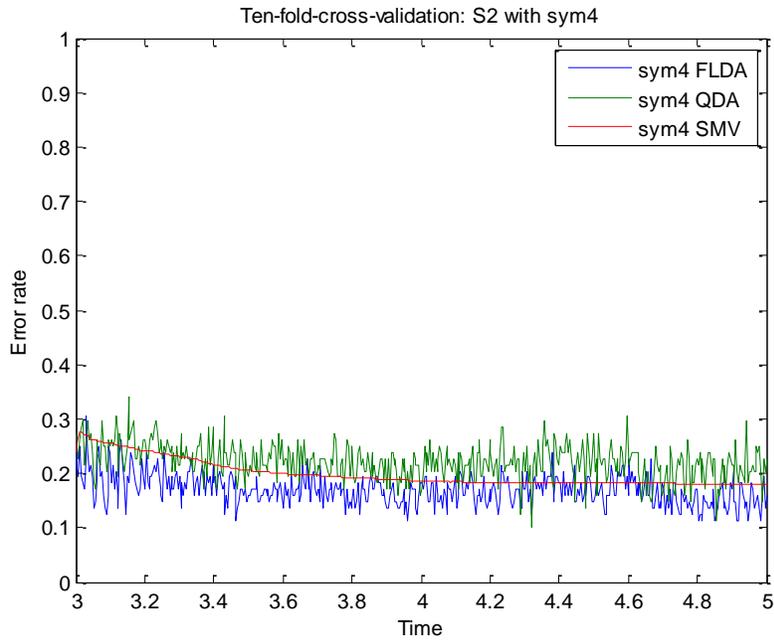


Fig. A.25. Error obtenido en el tiempo para S2 con sym4.

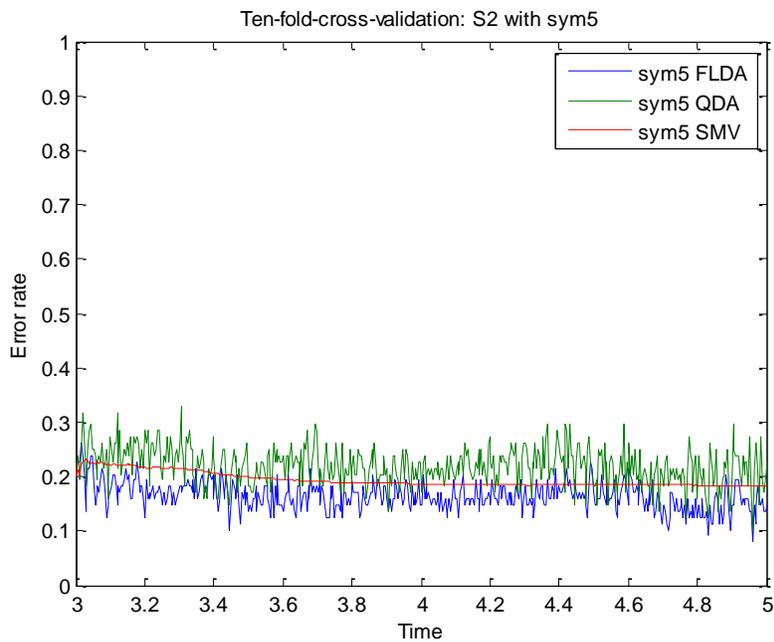


Fig. A.26. Error obtenido en el tiempo para S2 con sym5.

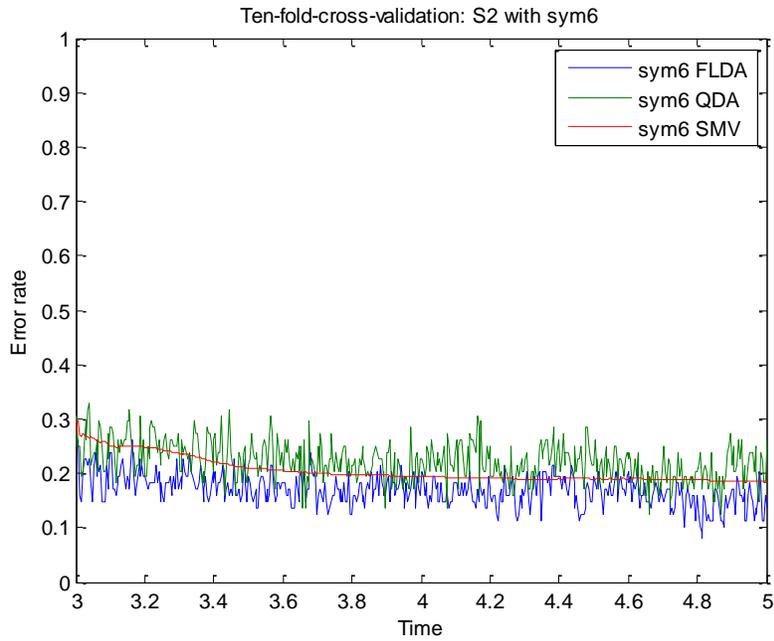


Fig. A.27. Error obtenido en el tiempo para S2 con sym6.

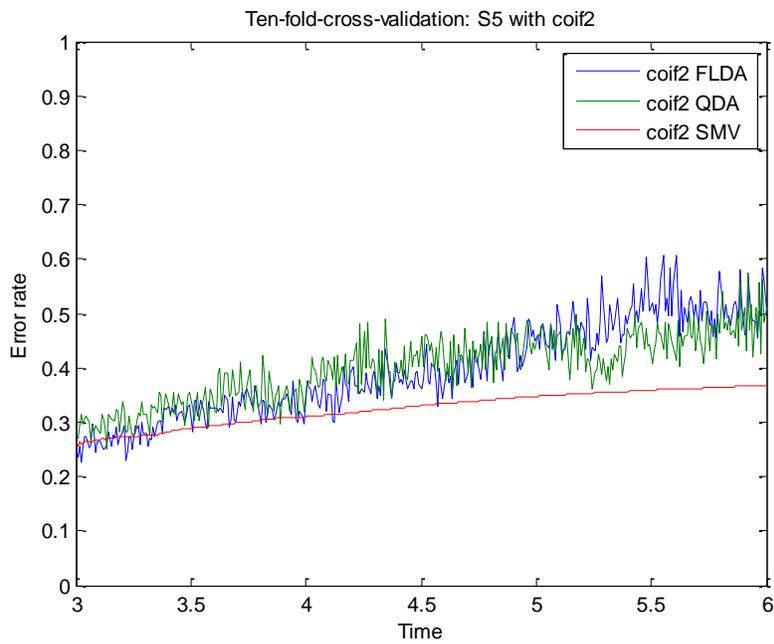


Fig. A.28. Error obtenido en el tiempo para S5 con coif2.

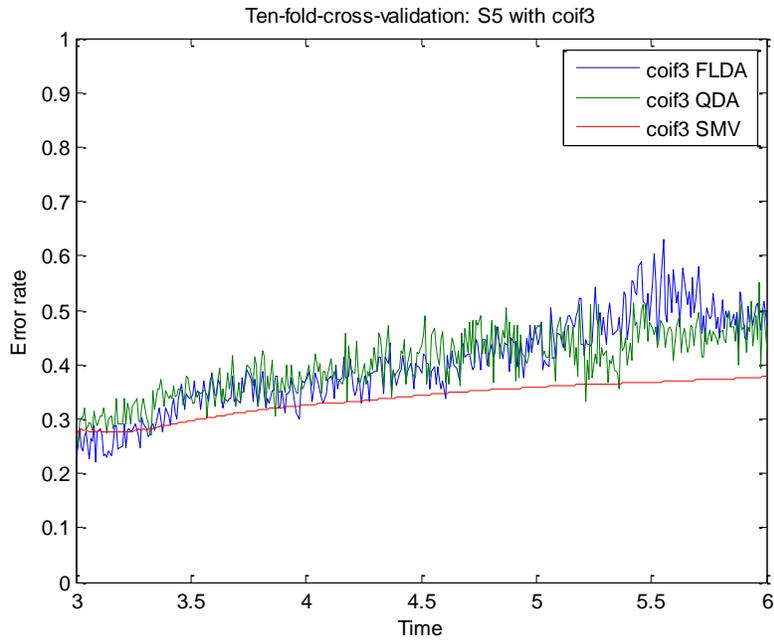


Fig. A.29. Error obtenido en el tiempo para S5 con coif3.

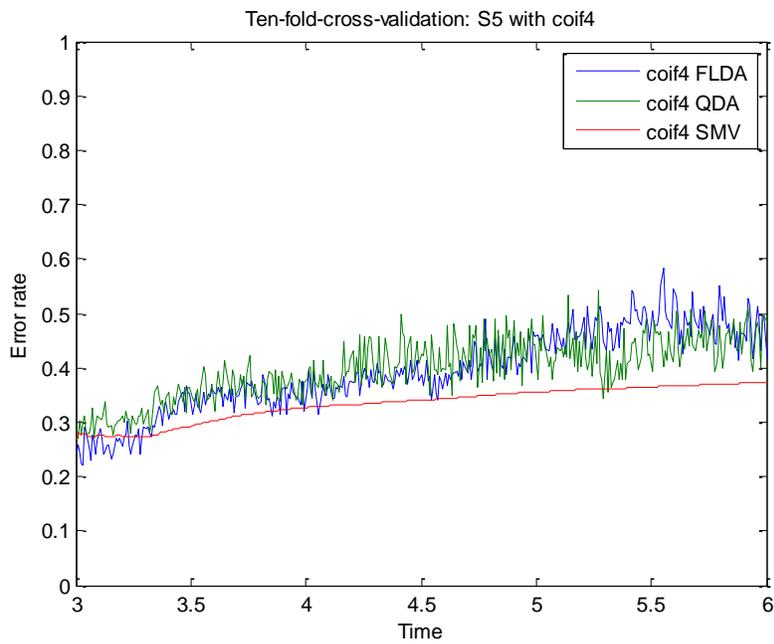


Fig. A.30. Error obtenido en el tiempo para S5 con coif4.

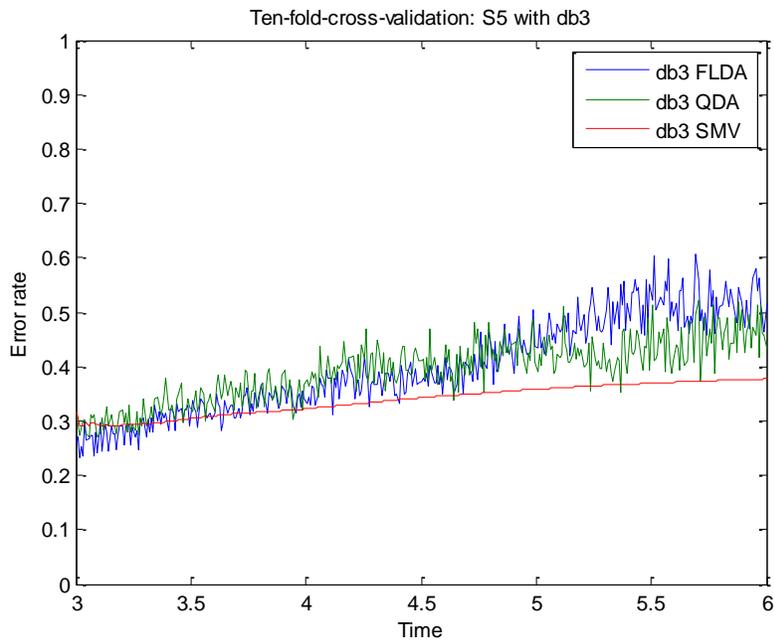


Fig. A.31. Error obtenido en el tiempo para S5 con db3.

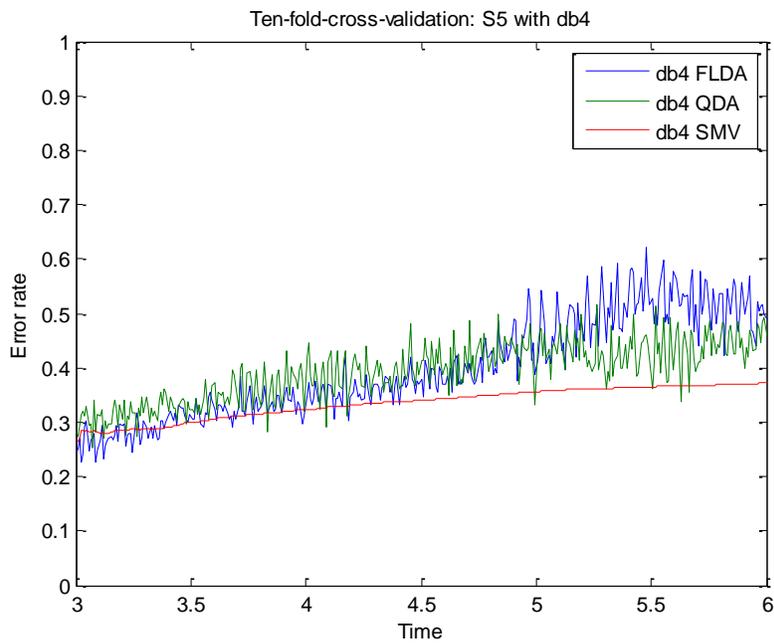


Fig. A.32. Error obtenido en el tiempo para S5 con db4.

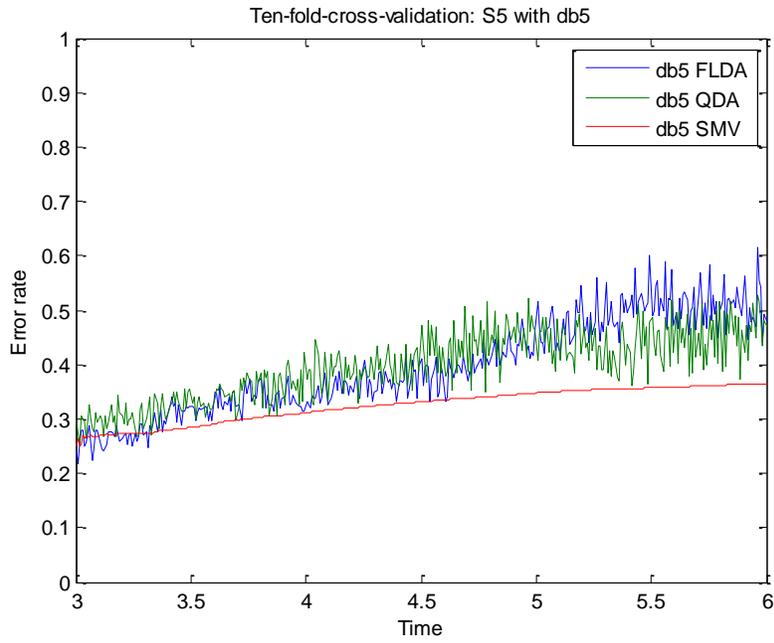


Fig. A.33. Error obtenido en el tiempo para S5 con db5.

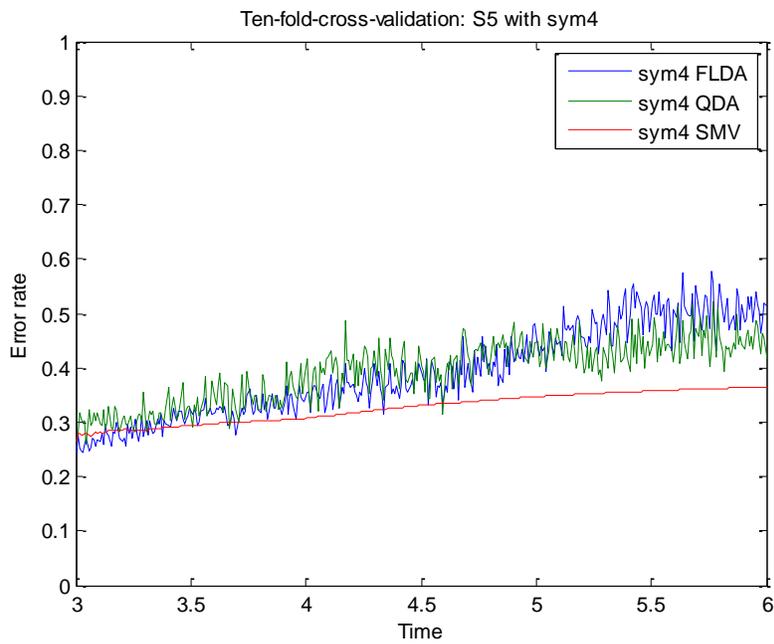


Fig. A.34. Error obtenido en el tiempo para S5 con sym4.

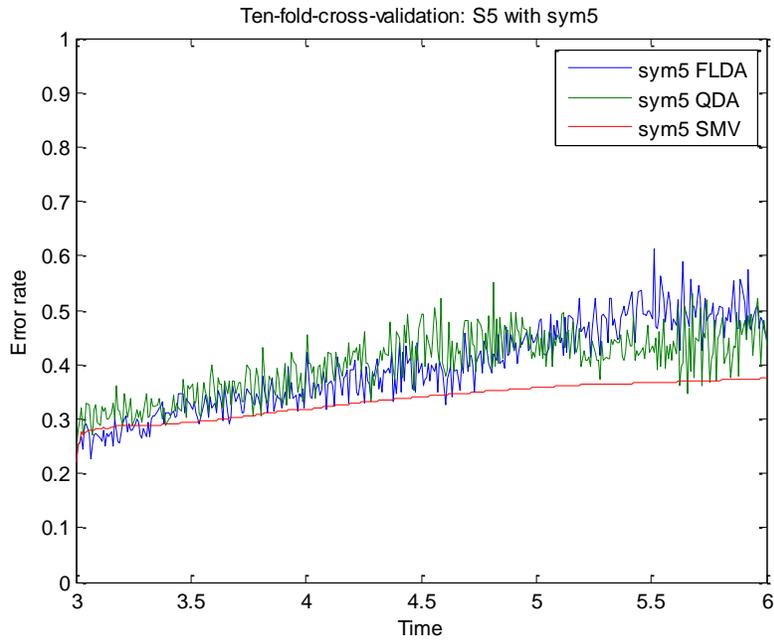


Fig. A.35. Error obtenido en el tiempo para S5 con sym5.

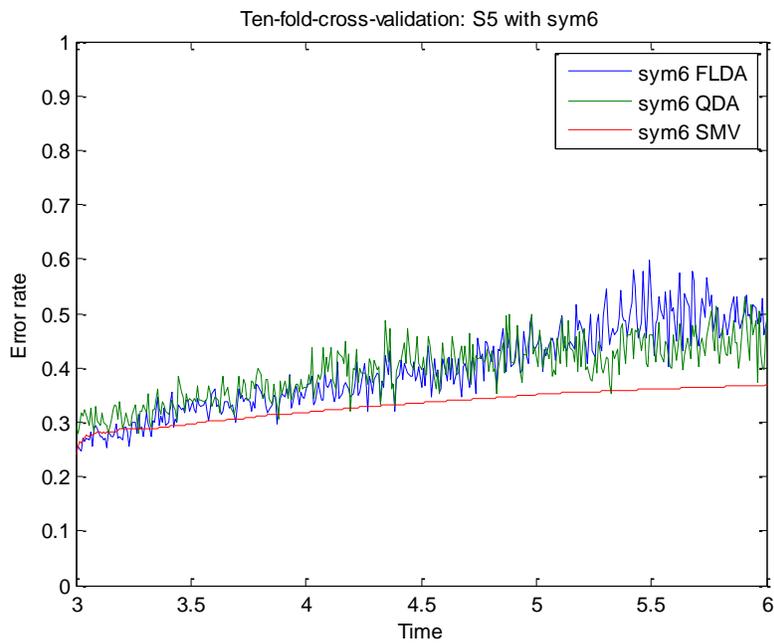


Fig. A.36. Error obtenido en el tiempo para S5 con sym6.

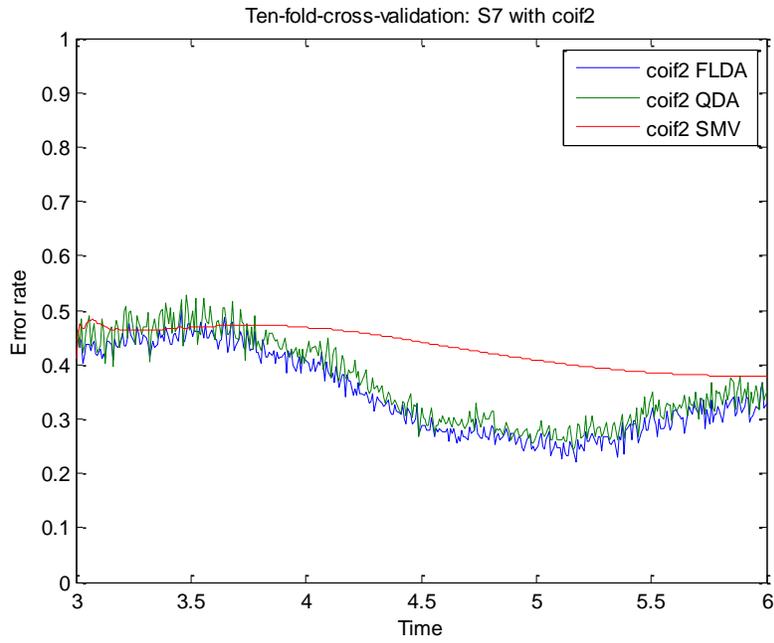


Fig. A.37. Error obtenido en el tiempo para S7 con coif2.

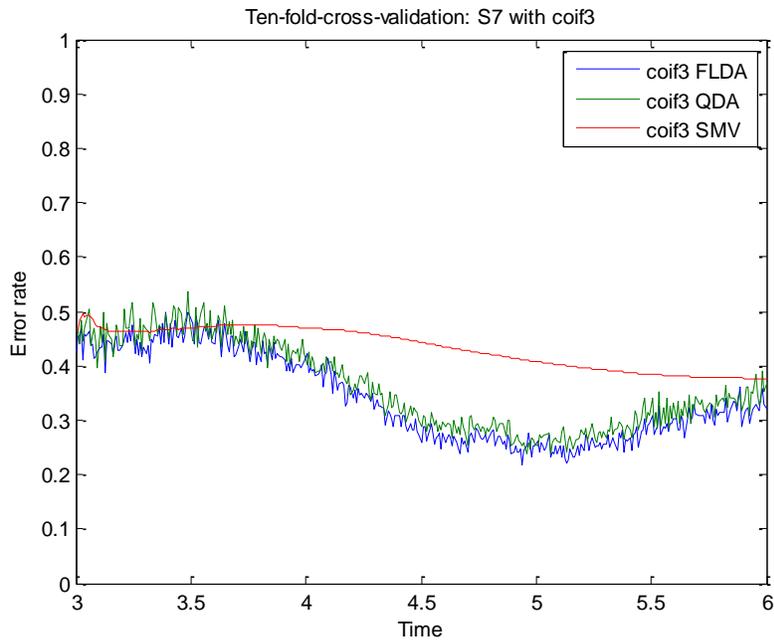


Fig. A.38. Error obtenido en el tiempo para S7 con coif3.

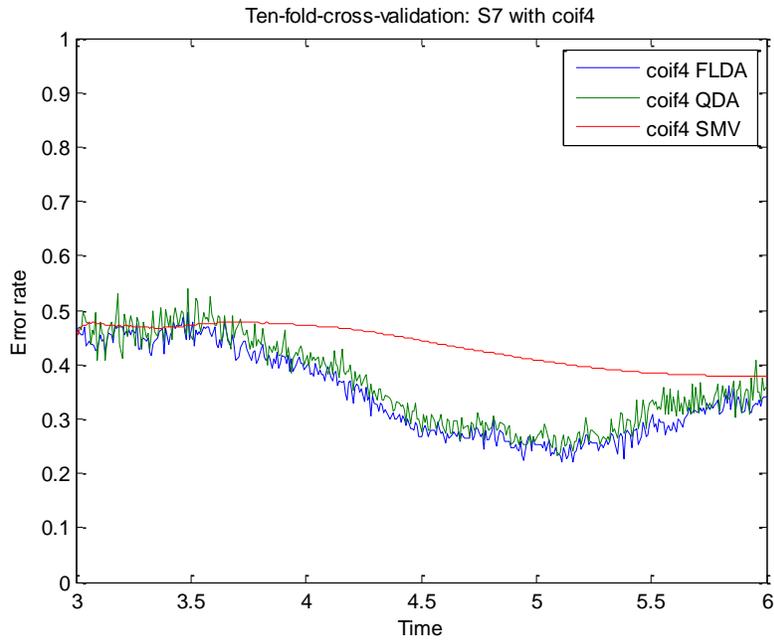


Fig. A.39. Error obtenido en el tiempo para S7 con coif4.

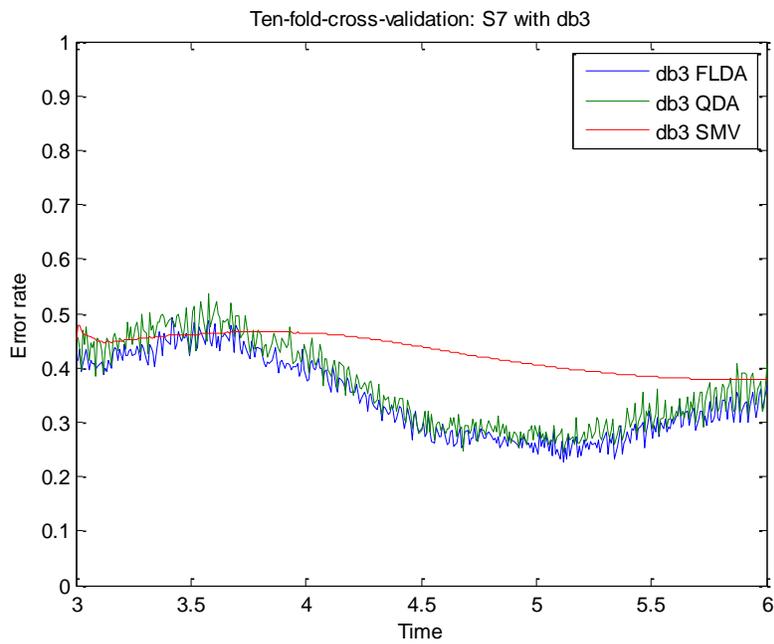


Fig. A.40. Error obtenido en el tiempo para S7 con db3.

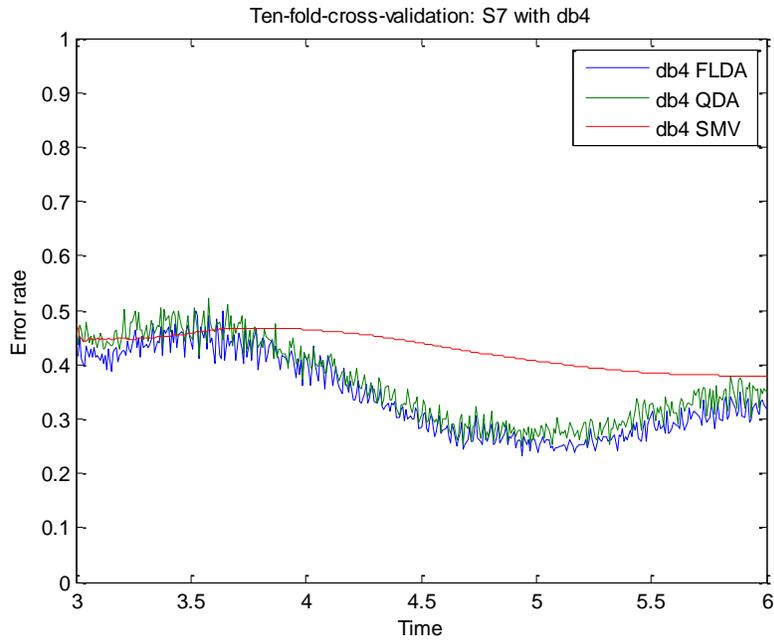


Fig. A.41. Error obtenido en el tiempo para S7 con db4.

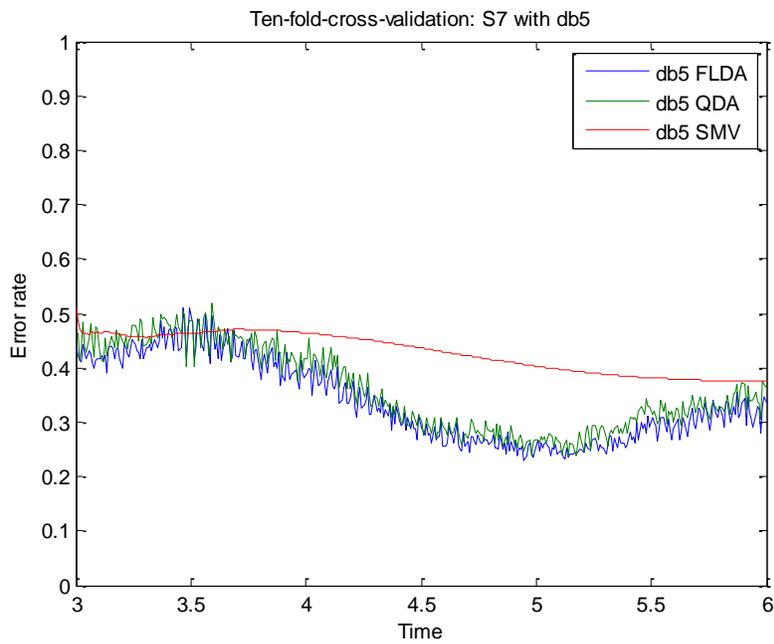


Fig. A.42. Error obtenido en el tiempo para S7 con db5.

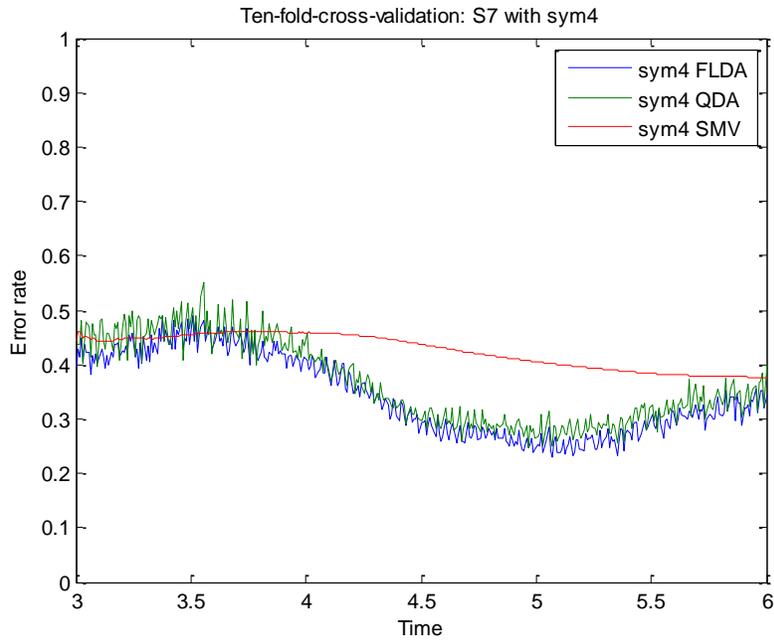


Fig. A.43. Error obtenido en el tiempo para S7 con sym4.

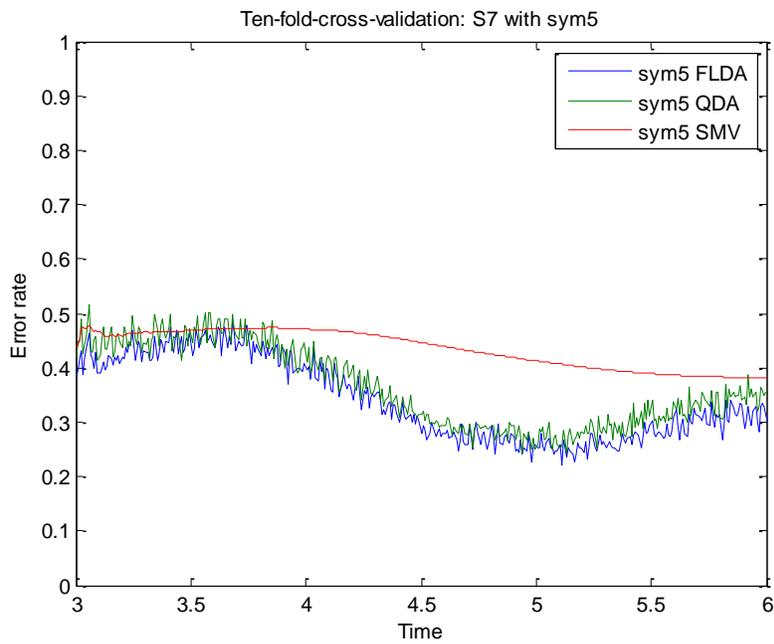


Fig. A.44. Error obtenido en el tiempo para S7 con sym5.

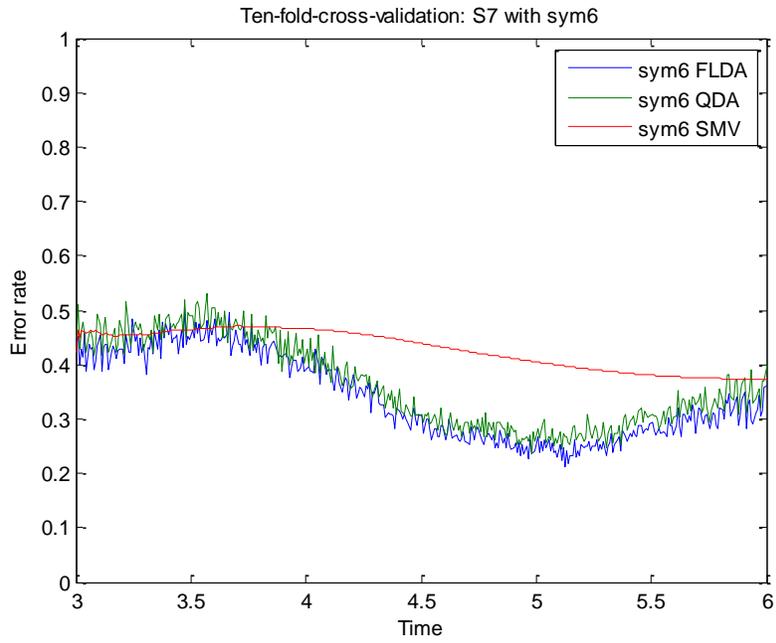


Fig. A.45. Error obtenido en el tiempo para S7 con sym6.