

Extracción de características en interfaces cerebro computadoras mediante transformada wavelet discreta: Resultados preliminares

Iván E. Gareis[#], Gerardo G. Gentiletti[#], Rubén Acevedo^{#*}, Leonardo Rufiner^{#*}

[#]LIRINS, Facultad de Ingeniería, UNER, ^{*}SINC, Facultad de Ingeniería y Ciencias Hídricas, UNL
ivangareis@hotmail.com

Resumen—El objetivo de este trabajo es evaluar diferentes alternativas de extracción de características, con el objetivo de detectar la señal de potenciales evocados relacionados con eventos en interfaces cerebro computadoras. Se trata de minimizar el tiempo empleado y el error de detección en términos de sensibilidad y especificidad del método, buscando alternativas a la promediación coherente. En este contexto se presentan los resultados preliminares obtenidos al realizar la extracción de características mediante transformada wavelet discreta diádica utilizando distintas wavelets madre. Para el bloque de clasificación se utilizó un perceptrón de simple capa. Se compararon los resultados obtenidos para los casos de patrones temporales y patrones wavelets, obteniéndose una mejora en la tasa de clasificación, especificidad y sensibilidad para los patrones wavelets.

Palabras clave— Interfaz cerebro computadora, potenciales evocados relacionados con eventos, reconocimiento de patrones, transformada wavelet.

INTRODUCCIÓN

LA técnica comúnmente utilizada para mejorar la relación señal a ruido (RSR), y estimar la señal de potencial evocado (PE), es la promediación coherente o sincronizada. Esta consiste en aplicar sucesivos estímulos, y promediar la actividad eléctrica registrada en forma sincronizada con el instante de tiempo en que se aplica el estímulo.

Si bien la promediación es de uso masivo, presenta limitaciones que justifican buscar alternativas a esta. Una es el tiempo requerido para lograr una estimación confiable del PE, otra es que se asume que la señal de PE se mantiene constante de época a época y que el ruido (el resto de los potenciales) es del tipo blanco con media cero, hipótesis que no es cierta para la mayoría de los casos. En consecuencia el resultado obtenido de la promediación es una mala estimación de la señal real de PE [1].

Por otro lado, hay situaciones en las que no se necesita estimar la morfología de la señal y sólo basta con saber si está presente o no, la señal de potencial evocado. Este es el caso de las interfaces cerebro computadoras (ICC). Hay diversos paradigmas de ICC que se pueden agrupar inicialmente en dos grupos: a) Invasivas: en este caso el registro de señales se realiza en forma intracraneal o intracortical y b) No invasivas utilizando registro de superficie, de las cuales, las más comunes son las basadas en el electroencefalograma (EEG). En este último grupo se encuentra el paradigma basado en potenciales evocados relacionados a eventos (ERP, del inglés: Event Related Potentials) [2].

Los ERP son potenciales evocados de latencias mayores a 100 ms cuya manifestación depende de procesos psicológicos y conductuales, de los cuales el más estudiado es la respuesta P300. Cuando estímulos visuales, auditivos o somatosensoriales infrecuentes o particularmente significativos, son mezclados con estímulos frecuentes (o de rutina), los primeros evocan un potencial en el EEG sobre la corteza parietal con un pico positivo cercano a los 300 ms. Para estimar la señal de ERP es necesario mejorar la RSR inicial, ya que junto con estos se registra electroencefalograma y electromiograma entre otras señales que conforman el ruido contaminante.

Donchin *et al* [3] fueron los primeros en utilizar el P300 en una ICC con estímulos visuales organizados por renglones y columnas sobre una matriz de caracteres que permitían su deletreo. La ventaja más sobresaliente de esta técnica, es que no requiere de ningún entrenamiento para el usuario, al mismo tiempo que resulta natural en tareas de selección de objetivos (deletreo de palabras, selección de direcciones, selección de menús, etc.). Varios trabajos sugieren que las ICC basadas en P300 pueden alcanzar una tasa de comunicación de 25 bits/min operando en tiempo real [4].



Fig. 1: Arquitectura general de una ICC para control de dispositivos.

En la figura 1 se muestra la arquitectura general de una ICC propuesta por Millán *et al* [5], donde se describen los bloques funcionales de la misma, en el cual una persona (usuario) controla un dispositivo (por ejemplo una silla de ruedas motorizada). El usuario puede también monitorear el estado del dispositivo, a través de realimentaciones que le permiten determinar el resultado de sus esfuerzos por controlarlo.

Recientemente, con el avance en algoritmos de aprendizaje maquina y tecnologías de procesamiento digital, una parte de la investigación en ICC reside en la exploración de técnicas de extracción de características y clasificación de las señales registradas.

En una ICC como la descrita se pueden distinguir claramente los siguientes bloques:

a) Instrumentación: es la etapa de adquisición y acondicionamiento de la señal de electroencefalograma y/o potenciales evocados.

b) Extracción de características: tiene como objetivo generar una representación adecuada de la señal de entrada original (patrón, en el contexto del reconocimiento de patrones) que permita mejorar el desempeño del clasificador.

c) Clasificador: se utiliza un clasificador lineal asumiendo inicialmente que el problema es de dos clases linealmente separables.

Los últimos dos bloques son un sistema de reconocimiento de patrones donde el problema de clasificación a resolver tiene dos clases posibles: registros con respuesta (ERP) y registros sin respuesta.

El objetivo de este trabajo es evaluar diferentes alternativas de extracción de características para detectar la señal de potenciales evocados relacionados con eventos en interfaces cerebro computadoras minimizando el tiempo empleado en realizar la misma, así como también el error medido en términos de sensibilidad y especificidad del método. Se presentan los resultados preliminares de aplicar una técnica de descomposición de señales basada en diccionarios de átomos tiempo-frecuencia, el cual se genera a partir de descomposición wavelet multiresolución.

METODOLOGÍA

Registros de ERP

En el presente trabajo se utilizaron registros obtenidos con amplificadores Grass® modelo 8-18-36, los parámetros de adquisición de las señales de EEG se indican en la tabla I. Se utilizó el software BCI2000 desarrollado en el Wadsworth Center, Albany, New York [6].

TABLA I
 PARÁMETROS DE ADQUISICIÓN DEL EEG.

Parámetro	Descripción
Canales	Fz, Cz, Pz, Oz, C3 y C4 M1 (referencia) y M2 (tierra)
Banda de Paso	0,1 – 15 Hz
Frecuencia de Muestreo	1024 Hz

El número de registros con el que se trabajó fue de 800 registros, 400 con P300 y 400 sin P300. Si bien debido al método de obtención se cuenta con una mayor cantidad de registros sin P300 se prefirió mantener los conjuntos balanceados y no utilizar todos los registros.

Para minimizar la cantidad de información presentada al clasificador se realizaron sobre los registros temporales sucesivas decimaciones, con el objetivo de determinar cuanto se podía disminuir la cantidad de muestras de los patrones sin reducir el desempeño del clasificador. Se cuidó que no se produzca aliasing por medio de la aplicación de un filtro de Chebyshev. Por conveniencia las decimaciones se realizaron en múltiplos de dos.

Para la estimulación se utilizó el esquema del delectreador de Donchin [3], solo que se reemplazaron los caracteres del abecedario por íconos que controlan el movimiento de una silla de ruedas [7].

Extracción de características

Existen diversas alternativas para la extracción de características, algunas son más adecuadas que otras en función de la señal y de las características que se quieren

resaltar. En este trabajo se proponen métodos de aproximación basados en diccionarios como los usados en [8]. En este método la señal de interés se considera como un elemento de un espacio de señales que puede representarse en función de una base o diccionario que favorezca la expresión de las características buscadas. En este contexto se puede expresar la señal $\mathbf{x} \in \mathbb{R}^N$ en función de una base ortogonal mediante la ecuación $\mathbf{a} = \phi \cdot \mathbf{x}$ donde $\mathbf{a} \in \mathbb{R}^N$ es el vector expresado en la nueva base y ϕ es una matriz $\in \mathbb{R}^{N \times N}$ cuyas columnas son los elementos de la nueva base (también denominados átomos) [7].

En este trabajo no se utilizó la matriz ϕ en forma directa, sino que se calcularon los coeficientes \mathbf{a} mediante la descomposición multiresolución producida por la transformada wavelet diádica discreta (DDWT, diadic discret wavelet transform) utilizando wavelets de distintas familias, las cuales se seleccionaron por su similitud visual con el pico P300. En esta descomposición, en cada nivel, la señal se separa en una porción de altas frecuencias llamada detalle y en otra de bajas frecuencias llamada aproximación, cada una de las cuales tiene la mitad de las muestras de la señal original. La descomposición en el nivel posterior se realiza sobre esta última, obteniendo así al final una aproximación y tantos detalles como veces se repita este proceso. El conjunto de coeficientes obtenido de esta manera contiene la información completa de la señal original y no tiene redundancia.

Previo a la aplicación de la DDWT se filtraron los registros por medio de un filtro digital pasa bajos de Chebyshev de octavo orden con frecuencia de corte en 100Hz y luego se submuestrearon. Luego de realizar la DDWT se seleccionaron diferentes conjuntos de detalles y en la aproximación, y con los coeficientes de estos se formaron vectores, uno correspondiendo a cada época. Estos vectores constituyeron los grupos de patrones a presentar al clasificador. Los conjuntos de coeficientes se seleccionaron tomando en cuenta que los ERP están formados por componentes de baja frecuencia, por tanto siempre se tomaron las aproximaciones y se fue incrementando la cantidad de detalles utilizados, siempre conservando los detalles de escalas mayores y agregando detalles de escalas pequeñas para poder determinar que escalas de los detalles tenían información útil para el proceso de clasificación.

Clasificador

Suponiendo que las clases del problema son linealmente separables, se utilizó una red neuronal tipo perceptron de simple capa (PSC).

Desempeño

Como índices para evaluar el desempeño del sistema completo de clasificación se utilizaron el porcentaje de clasificación, la sensibilidad y la especificidad; los cuales se definen de la siguiente manera:

$$\% \text{ clasificación} = \frac{\text{Patrones bien clasificados}}{\text{Total de patrones}}$$

$$\text{sensibilidad} = \frac{VP}{VP + FN}$$

$$\text{especificidad} = \frac{VN}{VN + FP}$$

donde VP son verdaderos positivos, VN son verdaderos negativos, FP son falsos positivos y FN son falsos negativos.

RESULTADOS

Registros

En la figura 2 se muestran los promedios de 300 épocas con ERP obtenidos en registros simultáneos en las posiciones Fz, Cz, Pz, Oz, C3 y C4, en sus respectivas posiciones. Se puede observar que la morfología de los ERP depende de la posición de los electrodos en los cuales se hace el registro, y en algunas posiciones se manifiesta mejor que en otras. De los 6 canales se seleccionaron Cz y Oz, en los cuales el pico P300 se manifiesta con diferente polaridad, y se evaluó como influye la cantidad de épocas a promediar en el análisis en tiempo y frecuencia.

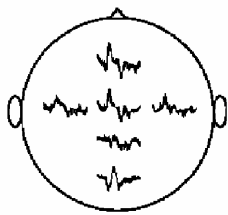


Fig. 2: Morfología de los ERP en diferentes posiciones.

En la figura 3 se muestran los escalogramas resultantes de aplicar Transformada Wavelet Continua (muestreada), utilizando la wavelet Biorthogonal 3.9 (seleccionada por su similitud morfológica visual con el pico P300), a promedios de 10, 5, y 1 época de registros con P300 y sin P300 en los canales Cz y Oz. Estos muestran diferencias entre canales para los casos en que hay P300 (con ERP) y que no hay (sin ERP), aunque es más evidente en el canal Oz. Esta situación para el caso del canal Oz se mantiene en forma marcada para promedios de hasta 3 épocas (no mostrados aquí). Debido a este análisis preliminar se trabajó con los datos del canal Oz con promedios de hasta dos épocas, descartando por ahora los datos de los patrones no promediados (una sola época ó *single trial*).

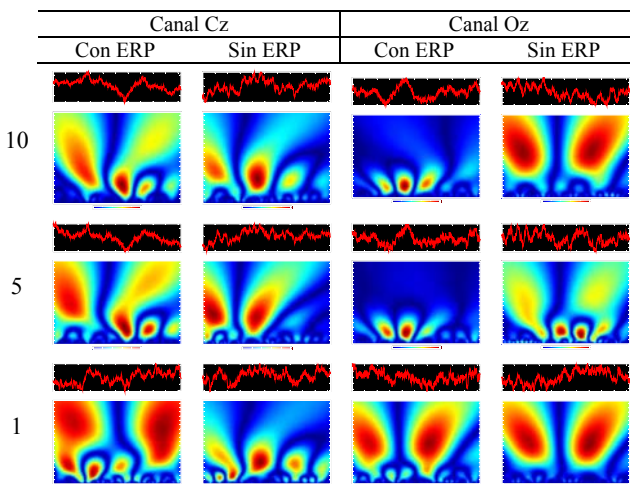


Fig. 3 – Escalogramas para Cz y Oz en función de las épocas promediadas

En la figura 4 se muestran los promedios obtenidos en registros correspondientes al canal Oz, en línea llena se presenta el promedio correspondiente a 300 épocas con

P300 mientras que en línea punteada el correspondiente a épocas sin P300.

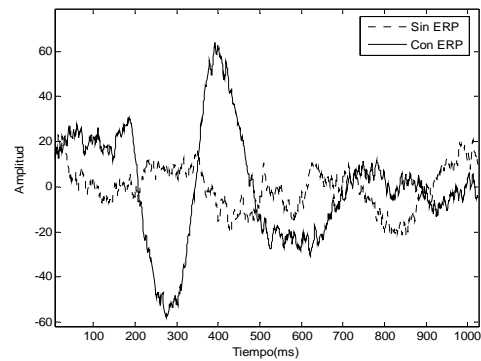


Fig. 4: Épocas con y sin P300 en el canal C4.

Submuestreo

La tabla II resume los resultados obtenidos al evaluar el desempeño del sistema completo pasando al clasificador los registros temporales de promedios de dos épocas decimados. Como se puede ver en la tabla el desempeño del sistema fue máximo al realizar una decimación por cuatro. En base a esto se definió este nivel de decimación para ser utilizado como paso previo a la aplicación de la DDWT

TABLA II
DESEMPEÑO DEL SISTEMA PARA DISTINTAS DECIMACIONES

Decimacion	Aciertos	Sensibilidad	Especificidad
Sin decimacion	63,00 %	0,70	0,56
x 2	63,00 %	0,72	0,54
x 4	64,75 %	0,75	0,54
x 8	61,75 %	0,69	0,54
x 16	62,00 %	0,70	0,53
x 32	62,00 %	0,69	0,55

Detección de ERP

Considerando los resultados obtenidos en el análisis tiempo-escala, y en el desempeño del sistema para los patrones temporales decimados, se generaron conjuntos de patrones temporales que consistieron en promedios de 2 épocas correspondientes al canal Oz, todos decimados cuatro veces. A partir de estos se generaron patrones con los coeficientes wavelets correspondientes a la aproximación del nivel seis y de diferente cantidad de detalles, resultante de la descomposición multiresolución de cada uno de los patrones temporales. Siempre se conservó la aproximación seis debido a que se asumió que en ese nivel de escala los patrones con los que se trabajó contenían información sobre la presencia del ERP. Esta suposición, se basa en el hecho de que la onda del ERP está compuesta por armónicos de bajas frecuencias al ser obtenida por métodos no invasivos. Al tener cada detalle y cada aproximación la mitad de las muestras de la señal en el nivel anterior se llega, partiendo de una señal de 256 muestras, a tener sólo cuatro muestras en la sexta aproximación. En la tabla III se detalla la cantidad de muestras de cada uno de los detalles en los diferentes niveles así como también el de la aproximación en el nivel seis para los patrones decimados de 256 muestras.

TABLA III

CANTIDAD DE MUESTRAS PARA DISTINTOS NIVELES DE DESCOMPOSICIÓN

Profundidad de descomposición		Número de muestras
Nivel 1	Detalle	128
Nivel 2	Detalle	64
Nivel 3	Detalle	32
Nivel 4	Detalle	16
Nivel 5	Detalle	8
	Aproximación	4

Partiendo de los patrones promediados de a dos y decimados se realizó la descomposición de estos con wavelets de distintas familias, obteniendo de estas forma los coeficientes wavelet correspondientes a la aproximación en el nivel seis y los detalles de los niveles uno, dos, tres, cuatro, cinco y seis. Luego se tomaron los coeficientes correspondientes a la aproximación del nivel seis y los correspondientes a los detalles de diferente cantidad de niveles, y se formó con estos los patrones que fueron presentados al clasificador.

En la tabla IV se presenta el desempeño del PSC en la detección de patrones correspondientes a descomposiciones con algunas wavelets representativas de patrones formados con promedios de dos épocas, tomando en cada caso diferente cantidad de detalles.

TABLA IV

DESEMPEÑO DEL SISTEMA PARA DIFERENTES WAVELETS Y DETALLES

Wavelet madre: Biortogonal 1.5			
Detalles	Aciertos[%]	Sensibilidad	Especificidad
1-6	61,00	0,68	0,54
2-6	61,50	0,71	0,52
3-6	59,50	0,69	0,50
4-6	63,75	0,75	0,53
5-6	65,00	0,77	0,53
Wavelet madre: Coiflets 4			
Detalles	Aciertos	Sensibilidad	Especificidad
1-6	61,00	0,70	0,52
2-6	64,25	0,75	0,54
3-6	66,75	0,75	0,58
4-6	64,00	0,76	0,52
5-6	64,75	0,75	0,55
Wavelet madre: Daubechies 9			
Detalles	Aciertos	Sensibilidad	Especificidad
1-6	63,25	0,71	0,55
2-6	63,75	0,72	0,55
3-6	65,25	0,70	0,60
4-6	66,25	0,72	0,60
5-6	69,25	0,74	0,64
Wavelet madre: Symlets 8			
Detalles	Aciertos	Sensibilidad	Especificidad
1-6	63,00	0,73	0,53
2-6	64,50	0,73	0,56
3-6	62,50	0,72	0,53
4-6	65,75	0,76	0,55
5-6	60,50	0,72	0,49

CONCLUSIONES

Del análisis de los resultados se desprende que el mejor desempeño es superior para el caso de la descomposición con wavelets Daubechies 9, siendo esto cierto en general para todas las wavelets pertenecientes a esta familia que fueron utilizadas. Se observa también que en algunos casos

el desempeño del sistema es incluso menor que utilizando los patrones representados en una base temporal. Es importante resaltar también que se observó en general una sensibilidad mayor que la especificidad, resultado que de ser consistente debe ser tomado en cuenta en el diseño de la interfaz con el objetivo de mejorar su rendimiento.

Se realizó además un análisis similar al anterior pero utilizando promedios 2, 3, 4, 5, 6, 7, 8, 9 y 10 épocas llegando a alcanzar valores de clasificación del 75% con promedios de diez épocas y aplicando DDWT y de hasta un 70% con promedios de diez épocas presentando al PSC los registros temporales.

Si bien este trabajo es un paso inicial en esta línea de investigación, se lograron resultados esperanzadores. Ya que utilizando un clasificador muy sencillo y probablemente no óptimo, y utilizando los registros de solo un canal se logró obtener --con promedios de únicamente dos épocas-- porcentajes de clasificación aceptables.

Sin embargo para alcanzar un desarrollo aplicable en forma práctica se deberá mejorar el desempeño del sistema; considerando por ejemplo más canales así como también el uso de otro tipo de clasificadores. En el futuro se pretende lograr un diccionario óptimo construido "a medida" para mejorar la detección de P300 en el contexto de una ICC.

AGRADECIMIENTOS

Este trabajo fue financiado por la Universidad Nacional de Entre Ríos (PID 6101) y por el Programa de Cooperación Internacional STIC-AMSUD (09STIC01).

REFERENCIAS

- [1] R. Acevedo, C. Bonell, G. Gentiletti, J. Aldonate, C. Mercuri, S. Escobar, J. Biurrun, "Potenciales evocados auditivos de tronco cerebral: Aspectos tecnológicos del registro", Revista Argentina de Bioingeniería (SABI), Vol. 12 No. 1, pp. 10-17, Junio 2006.
- [2] G. Gentiletti, C. Tabernig, R. Acevedo, "Interfaz cerebro-computadora: Estado del arte y desarrollo en Argentina", Revista Argentina de Bioingeniería (SABI), Vol. 13 No. 1, pp. 13-19, Junio 2007, ISSN: 0329-5257
- [3] L. A. Farrel y E. Donchin, "Taking off the top of your head: toward a methal prothesis utilizing event related brain potentials", Electroencephalography and Clinical Neurophysiology, vol. 70, pp. 510-523, 1988.
- [4] R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain-computer interfaces for communication and control," Clinical Neurophysiology., vol. 113, pp.767-791, 2002.
- [5] J. del R. Millán, P. W. Ferrez, F. Galán, E. Lew, "Non-invasive brain-machine interaction", International Journal of Pattern Recognition and Artificial Intelligence, vol. 20, pp. 1-13, 2007.
- [6] <http://www.bci2000.org/BCI2000/Home.html>
- [7] G.G. Gentiletti Faenze, J.G. Gebhart, M. Richard, E. García Cossio, R.C. Acevedo, V. Medina Bañuelos, "Robotics Simulation Platform for BCI Application: A Wheelchair Driving Example Using P300 Paradigm", Proceedings of the 4th International Brain-Computer Interface Workshop and Training Course, Graz (Austria), 2008
- [8] Rufiner H. L., "Análisis y representación de la voz mediante técnicas no convencionales", Tesis de Doctorado en Ingeniería, Universidad de Buenos Aires, Argentina, 2005.