Extracción de características en ICC mediante métodos basados en diccionarios óptimos: Resultados preliminares

Rubén Acevedo[†], Gerardo Gentiletti[†], Verónica Medina Bañuelos^{*}, Leonardo Rufiner[#]
[†] LIRINS, Facultad de Ingeniería, Universidad Nacional de Entre Ríos.
^{*} LINI, Universidad Autónoma Metropolitana (México).
[#] SINC, Facultad de Ingeniería y Ciencias Hídricas, Universidad Nacional del Litoral. racevedo@bioingenieria.edu.ar

Resumen-El objetivo de este trabajo es detectar la señal de potenciales evocados relacionados con eventos (ERP) minimizando el tiempo empleado y el error medido en términos de sensibilidad y especificidad del método. Una interfaz cerebro computadora es un sistema que agrupa un bloque de instrumentación de biopotenciales, un bloque de extracción de características y un clasificador. En este contexto se presentan los resultados preliminares obtenidos al realizar la extracción de características mediante métodos de aproximación de señales basados en diccionarios. En particular el diccionario se obtiene a partir de descomposición wavelet multiresolución. El clasificador utilizado fue un perceptron de simple capa. Se compararon los resultados obtenidos en la detección de los ERP para los casos de patrones temporales y patrones wavelets. Los resultados obtenidos son muy buenos en el caso de señales simuladas, no siendo así para el caso de registros reales.

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

I. INTRODUCCIÓN

HAY diversos paradigmas de Interfaces Cerebro Computadora (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 y se denominan basadas en el EEG. Dentro de este último se encuentra el paradigma basado en potenciales evocados relacionados a eventos (ERP, event related potentials) [1].

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 el pico P300. Cuando estímulos visuales, auditivos o somatosensoriales infrecuentes 0 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 relación señal a ruido (RSR) inicial, ya que junto con estos se registra electroencefalograma (EEG), electromiograma (EMG) entre otras señales que conforman el ruido contaminante. El método utilizado para esto es la promediación coherente, que consiste en promediar la actividad eléctrica registrada de manera sincronizada con la aplicación del estímulo [2].

Donchin *et al* [3] fue el primero en utilizar el P300 en una ICC con estímulos visuales en 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]. Esta puede mejorarse mediante técnicas de selección de electrodos óptimos y nuevos algoritmos de reconocimiento de patrones.

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 o un brazo robot). El usuario puede también monitorear el estado del dispositivo, a través de realimentaciones que le permiten determinar el resultado de sus esfuerzos por controlarlo.



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

Recientemente, con el avance en algoritmos de aprendizaje maquinal 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 descripta se pueden distinguir claramente los siguientes bloques:

a) Instrumentación: básicamente consiste en los amplificadores con los cuales se hace el registro de electroencefalograma y/o potenciales evocados.

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

c) Clasificador: en general se utiliza un clasificador lineal debido a que el problema es de dos clases se suponen linealmente separables.

Los últimos dos bloques representan un sistema de reconocimiento de patrones aplicado, 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 detectar la señal de ERP minimizando el tiempo empleado en realizar la misma así como también el error medido en términos de sensibilidad y especificidad del método, utilizando un solo canal e registro. Se presentan los resultados preliminares de aplicar una técnica de descomposición de una señal basada en diccionario de átomos tiempo-frecuencia como método de extracción de características.

II. METODOLOGÍA

A. Registros de ERP

En el presente trabajo se utilizaron registros reales y registros simulados.

A.1 Registros reales

Se utilizaron amplificadores Grass® modelo 8-18-36, los parámetros de adquisición se indican en la tabla I. Se utilizó el software BCI2000 con autorización del Wadsworth Center (Albany, New York) [6].

Tabla I - Parámetros de adquisición.

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 mps		

Se adquirieron seis canales de EEG para poder determinar si variaba el desempeño del esquema propuesto en función de la posición de los electrodos.

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

A.2 Registros simulados

Se generaron utilizando un template al que se le sumó ruido blanco manteniendo la relación señal a ruido de los registros reales. El template se obtuvo mediante la promediación coherente del canal en el cual se obtuvo el mejor desempeño del método de extracción de características. Posteriormente se le aplicó un filtro pasa bajos para suavizarlo.

B. Extracción de características

Hay diversas alternativas, algunas son más adecuadas que otras en función de la señal y de las características que se quieran resaltar. En este trabajo se propone utilizar métodos de aproximación basados en diccionarios [7], en este contexto la señal de interés se considera como un elemento de un espacio de señales que puede representarse en función de una base que favorezca la expresión de las características buscadas. En este contexto se puede expresar la señal $\mathbf{x} \in \mathbf{R}^{N}$ en función de una base ortogonal mediante la ecuación $\mathbf{a} = \phi . \mathbf{x}$ donde $\mathbf{a} \in \mathbf{R}^{N \times N}$ cuyas columnas son los elementos de la nueva base denominados átomos [7].

En este trabajo la matriz ϕ se construyo mediante la descomposición multiresolución producido por la transformada wavelet diádica discreta (DDWT, diadic discret wavelet transform) utilizando la wavelet Biorthogonal 3.9 [7], la cual se seleccionó por su parecido morfológico con el pico P300. En la figura 2 se muestra la descomposición hasta el nivel 5, en cada nivel la señal se separa en una porción de altas frecuencias llamada *detalle* y en otra de bajas frecuencias llamada *aproximación*; la descomposición en el nivel posterior se realiza sobre esta última. En la figura 2 el i-ésimo detalle se representa como Di, mientras que la i-ésima aproximación como Ai.

Señal de ERP (0 – 512 Hz)				
Nivel=1	D1 (256 – 512 Hz) A1 (0 – 256 Hz)	512 muestras		
Nivel =2	D2 (128 – 256 Hz) A2 (0 – 128 Hz)	256 muestras		
Nivel =3	D3 (64 – 128 Hz) A3 (0 – 64 Hz)	128 muestras		
Nivel =4	D4 (32 – 64 Hz) A4 (0 – 32 Hz)	64 muestras		
Nivel =5	D5 (16 – 32 Hz) A5 (0 – 16 Hz)	32 muestras		
		/		

Fig. 2- Descomposición multiresolución de ERP.

Este proceso también puede verse como la aplicación de filtros pasa bajos y pasa altos sucesivas veces; donde en cada nivel además hay un proceso de decimación por 2 de la señal [7]. Siguiendo este razonamiento, y considerando que la señal de ERP original tiene 1024 muestras, la *aproximación* y *detalle* en el nivel 5 tienen 32 muestras cada una. Debido a que el ancho de banda de los ERP es menor a 10Hz, entonces toda la información se encuentra en la aproximación del nivel 5 (A5). Entonces, aplicando la descomposición multiresolución sobre el promedio de M épocas (patrón temporal de 1024 muestras) se obtiene un nuevo patrón de 32 muestras.

C. Clasificador

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

D. Desempeño

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

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

Sensibilidad = $\frac{\text{VP}}{\text{VP} + \text{FN}}$
Especificidad = $\frac{\text{VN}}{\text{VN} + \text{FP}}$

III. RESULTADOS

A. Registros

En la figura 3.a se muestran los promedios de 300 épocas obtenidos en registros simultáneos en las posiciones Fz, Cz, Pz, Oz, C3 y C4. Los trazos en rojo corresponden a registros donde no se produjo respuesta (ERP) y en azul corresponden a registros donde si hay respuesta. 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.



Fig.3 - Morfología de los ERP.

De los 6 canales se seleccionaron Cz (figura 3.b) y Oz (figura 3.c), en lo 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.

En la figura 4 se muestran los escalogramas resultantes de aplicar Transformada Wavelet Continua (muestreada), utilizando la wavelet Biorthogonal 3.9, a promedios de 10, 5 y 1 época de registros con P300 y sin P300 en los canales Cz y Oz. En el caso de promedios de 10 épocas se observan diferencias en los escalogramas de ambos canales para los casos en que hay P300 (con ERP) y que no hay (sin ERP); aunque es mas evidente en el canal Oz. Esta situación se mantiene para promedios de 5 épocas pero se pierde en el caso de una sola época.



Fig. 4 - Comparación entre los canales Cz y Oz.

En base a esto se realizó el análisis mas detallado en el canal Oz, el cual se muestra en la figura 5. Se puede observar que las diferencias entre los escalogramas obtenidos para promedios con y sin P300 se mantienen hasta 3 épocas.



Fig.5 - Análisis tiempo-escala del canal Oz.

B. Detección de ERP

Considerando los resultados obtenidos en el análisis tiempo-escala, se generaron conjuntos de patrones temporales que consintieron en promedios de 20, 10, 9, 8, 7, 6, 5, 4, 3 y 2 épocas correspondientes al canal Oz. A partir de estos se generaron patrones con los coeficientes wavelets correspondientes a la aproximación del nivel 5 resultante de la descomposición multiresolución de cada uno de los patrones temporales.

En la figura 6 muestra el desempeño del PSC en la detección de patrones (temporales y wavelets) correspondientes a registros simulados. Se muestra que el porcentaje de clasificación es mejor para el caso de los patrones wavelets, ya que para promedios de tan solo 5 épocas se logra el 100 % de detección.





En la figura 7 se muestra el desempeño del PSC en el caso de patrones generados a partir de registros reales.



Fig.7 - Porcentaje de clasificación en registros reales.

Complementando información sobre el desempeño del PSC, se muestran los valores de sensibilidad y especificidad para registros simulados (tabla II) y para registros reales (tabla III).

Tabla II - Registros simulados. Sensibilidad v especificidad

	Temporales		Wavelets	
Épocas	Sens.	Espec.	Sens.	Espec.
20	1	1	1	1
10	1	0.98	1	1
9	1	0.96	1	1
8	0.95	0.95	1	1
7	0.96	9.94	1	1
6	0.94	0.94	1	1
5	0.98	0.90	1	1
4	0.93	0.90	0.99	0.97
3	0.95	0.86	0.98	0.98
2	0.90	0.86	0.97	0.96

Tabla III – Registros reales
Sensibilidad v especificidad

~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~						
	Temporales		Wavelets			
Épocas	Sens.	Espec.	Sens.	Espec.		
20	0.9	0.65	0.85	0.70		
10	0.85	0.60	0.85	0.60		
9	0.81	0.86	0.77	0.66		
8	0.83	0.64	0.85	0.60		
7	0.84	0.60	0.84	0.57		
6	0.78	0.57	0.78	0.60		
5	0.77	0.62	0.82	0.62		
4	0.81	0.58	0.77	0.56		
3	0.78	0.57	0.77	0.59		
2	0.72	0.56	0.72	0.56		

#### **IV. CONCLUSIONES**

Si bien este trabajo es el puntapié inicial en esta línea de investigación, se lograron resultados esperanzadores en el caso de registros simulados, no ocurrió lo mismo para los registros reales. Esto puede deberse a varios motivos, entre los que se pueden mencionar la elección de la función wavelet madre, el clasificador utilizado, las características de los registros. En el futuro se pretende lograr un diccionario óptimo construido "a medida" para resolver el problema de detección de P300 en el contexto de una ICC.

#### REFERENCIAS

- Gabriel Gentiletti, Carolina Tabernig, Rubén Acevedo, "Interfaz [1] cerebro-computadora: Estado del artey desarrollo en Argentina", Revista Argentina de Bioingeniería (SABI), Vol. 13 No. 1, pp. 13-19, Junio 2007, ISSN: 0329-5257
- Rubén Acevedo, Claudia Bonell, Gabriel Gentiletti, Julio Aldonate, [2] Carlos Mercuri, Sergio Escobar, José 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.
- L. A. Farrel y E. Donchin, "Taking off the top of your head: toward [3] a methal prothesis utilizing event related brain potentials", Electroencephalography and Clinical Neurophysiology, vol. 70, pp. 510-523, 1988.
- R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. [4] 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
- Rufiner H. L., "Análisis y representación de la voz mediante técnicas no convencionales", Tesis de Doctorado en Ingeniería, [7] Universidad de Buenos Aires, Argentina, 2005