



USO DE LA COHERENCIA EN LA SELECCIÓN DE ELECTRODOS PARA LA CLASIFICACIÓN DE SEÑALES DE ELECTROENCEFALOGRAFÍA

R. Salazar-Varas^a, D. Gutiérrez^a

^aCentro de Investigación y Estudios Avanzados del IPN
Unidad Monterrey, Apodaca, Nuevo León.
rsalazar@cinvestav.mx, dgtz@ieee.org

RESUMEN

En este trabajo proponemos el uso de la conectividad existente entre diferentes regiones cerebrales para determinar la combinación óptima de sensores que sea capaz de discriminar señales de electroencefalografía (EEG) provenientes de diferentes estados mentales. Se calcularon las coherencias entre todas las combinaciones posibles dentro de una red de tres electrodos. Para formar el vector de características se utilizó la red de electrodos que brindara coherencias estadísticamente significativas para una clase y no para la otra. La clasificación se realizó mediante el discriminante de Fisher no lineal, donde el mejor plano de discriminación fue seleccionado en función de maximizar el área bajo la curva ROC (Receiver Operating Characteristics). Mediante ejemplos numéricos demostramos la aplicabilidad del método propuesto utilizando una base de datos reales provenientes de la competencia internacional de interfaz cerebro-computadora 2004.

INTRODUCCIÓN

Las interfaces cerebro-computadora (ICC) que utilizan datos de electroencefalografía (EEG) dependen de la exactitud en la clasificación de la actividad eléctrica cerebral para controlar en tiempo real dispositivos computarizados [1]. Recientemente se han propuesto muchos clasificadores lineales y no lineales, algunos de ellos reportan buen desempeño [2]. El método de Patrones Espaciales Comunes (Common Spatial Patterns, CSP) es ampliamente utilizado en aplicaciones de ICC para discriminar entre diferentes tareas de imaginación de actividad motora [3], pero su desempeño depende del conocimiento de las frecuencias de operación de los eventos a discriminar. En [4] propusimos un método basado en la eigendescomposición de las señales a partir de su modelo autorregresivo variante en el tiempo (time-varying autoregressive, TVAR) a fin de estimar la frecuencia específica para cada individuo, de modo que los CSP fueron calculados de las principales eigencomponentes. A pesar de que el método propuesto reportó alta efectividad en la clasificación de datos de EEG, la selección sistemática de los sensores continúa como una pregunta abierta.

Utilizar un gran número de electrodos puede mejorar el proceso de clasificación de las señales, sin embargo colocar muchos electrodos en la cabeza del usuario puede ser incómodo además de que se invierte mucho tiempo, lo que hace que el proceso de adquisición sea poco práctico para aplicaciones reales de ICC. Por otro lado, manejar información adquirida a través de un denso arreglo de electrodos es muy costoso en términos computacionales, lo que dificulta el procesamiento en tiempo real. Con el fin de obtener el mínimo número de sensores requeridos en la clasificación de señales de EEG sin afectar el desempeño en la clasificación, se han propuesto diferentes soluciones: en [5] se evalúan diferentes configuraciones de sensores predefinidas, buscando aquella que brinde una mejor clasificación, igualmente se propone un método mediante el cual se descartan sensores que no aportan información considerable al proceso de clasificación. Otras soluciones para seleccionar un número reducido de sensores ha sido modificando los filtros espaciales (incluyendo los CSP), con lo que se ha logrado que el número de sensores reduzca significativamente (por ejemplo [6, 7]).



En este trabajo presentamos una técnica basada en el hecho de que diferentes regiones cerebrales son simultáneamente activadas y se comunican entre sí cuando se realizan diversas tareas [8], proponemos que un arreglo óptimo de sensores puede ser determinado con base en la información de dicha conectividad. Una medida que provee información acerca de la conectividad es la coherencia. En [9] se estudian los patrones de coherencia producidos al realizar tareas de manera simultánea y sucesiva. Por otro lado, en [10] se hace uso de la coherencia para clasificar señales de EEG correspondientes al estado de reposo, imaginando un movimiento y ejecutando un movimiento. Sin embargo la coherencia no ha sido utilizada para estudiar la conectividad entre diferentes regiones cerebrales y cómo emplear esta información para la selección de sensores. Así pues, la búsqueda de sensores basada en la coherencia puede realizarse calculando la coherencia entre todas las posibles combinaciones de sensores y elegir la que tenga una coherencia significativa para una clase y no para la otra. Como un primer intento nuestro espacio de búsqueda fue de 9 electrodos, distribuidos sobre la zona frontal, central y parieto-occipital. A partir de la mejor combinación se obtienen las coherencias para formar el vector de características que emplea en el proceso de clasificación. El resto del trabajo está organizado de la siguiente forma: en la Sección 2 se expone el concepto de coherencia y se describe el método propuesto. En la Sección 3 mostramos la aplicabilidad de nuestro método mediante una serie de ejemplos numéricos con datos reales de EEG. Finalmente, en la Sección 4 se hace una pequeña discusión de los resultados.

MÉTODOS

En esta sección se hace una pequeña revisión del concepto de coherencia y explicamos el método de clasificación de señales de EEG basado en la coherencia.

Coherencia

Definamos $x_m(t)$ como el registro de EEG en el sensor $m = 1, 2, \dots, M$, realizado en los instantes de tiempo $t = 1, 2, \dots, N$. La coherencia entre las señales $x_i(t)$ y $x_j(t)$ está dada por [11]

$$\gamma_{ij}^2(f) = \frac{|P_{ij}(f)|^2}{P_i(f)P_j(f)}, \quad (1)$$

Donde $P_{ij}(f)$ es el espectro cruzado de $x_i(t)$ y $x_j(t)$, y $P_i(f)$ y $P_j(f)$ son los autoespectros de las señales. La coherencia es una medida del grado de correlación del espectro de energía de dos señales, en este caso provenientes de dos diferentes electrodos. Un elevado valor en la coherencia implica un alto grado de comunicación entre diferentes regiones del cerebro, mientras que valores bajos sugieren independencia [12].

En nuestro caso estamos interesados en analizar la conectividad entre tres sensores formando una red. Nos referiremos a las mediciones en estos sensores como $x_A(t)$, $x_B(t)$ y $x_C(t)$. Usando (1), se obtienen las siguientes coherencias: $\gamma_{AB}^2(f)$, $\gamma_{AC}^2(f)$ y $\gamma_{BC}^2(f)$. Estas coherencias pueden ser obtenidas para diferentes frecuencias de interés. En nuestro caso seleccionamos una frecuencia (denotada como f_s) en la que: (i) la coherencia fuera significativa (con un nivel de significancia de $\alpha=0.05$) y (ii) que se encontraran grandes diferencias entre las coherencias producidas por diferentes tareas. Las coherencias obtenidas dentro de la red seleccionada, fueron colocadas en un vector de características y , de tres dimensiones, donde cada dimensión corresponde a la coherencia $\gamma_{ij}^2(f_s)$, por ejemplo:

$$y = \begin{bmatrix} \gamma_{AB}^2 \\ \gamma_{AC}^2 \\ \gamma_{BC}^2 \end{bmatrix} \quad (2)$$

Debe notarse que a pesar de que en el presente trabajo evaluamos una red de tres sensores, el proceso se puede generalizar para un arreglo L-dimensional, obtenido de una red de sensores más compleja.

Clasificación



Ya que los datos han sido proyectados en un espacio de características de tres dimensiones, se requiere establecer límites para separar las características en grupos correspondientes a las clases que deseamos discriminar, para lo que empleamos el discriminante no lineal de Fisher. En el caso de dos clases, el hiperplano de discriminación es calculado mediante el siguiente vector normal [13]

$$v = \left(\beta \{\hat{R}_{yy}\}_1 + (1 - \beta) \{\hat{R}_{yy}\}_2 \right)^{-1} (\{\bar{y}\}_2 - \{\bar{y}\}_1) \quad (3)$$

Donde $\{\hat{R}_{yy}\}_i$ y $\{\bar{y}\}_i$ son la matriz de covarianza y la media de la clase $i=1,2$ respectivamente, y β es el factor de ajuste, que se encuentra en el rango $[0,1]$ y se elige con base en un criterio de optimización. En nuestro caso decidimos maximizar el área bajo la curva ROC (Receiver Operating Characteristics), donde 1 es resultado de una clasificación perfecta mientras que un área cercana a 0.5 indica que la clasificación se está llevando a cabo de manera aleatoria. La evaluación basada en el área bajo la curva ROC es utilizada en muchas áreas [14] y se ha utilizado para evaluar otros sistemas de ICC [15]. Por lo tanto el valor de β que maximice el área bajo la curva ROC generará nuestro mejor plano de discriminación. Bajo estas condiciones y dado un vector de características, la función de discriminación puede ser escrita como

$$D = v^T y + v_o, \quad (4)$$

Con

$$v_o = - \frac{\beta \sigma_1^2 v^T \{\bar{y}\}_2 + (1 - \beta) \sigma_2^2 v^T \{\bar{y}\}_1}{\beta \sigma_1^2 + (1 - \beta) \sigma_2^2}, \quad (5)$$

Y

$$\sigma_i^2 = v^T \{\hat{R}_{yy}\}_i v, \quad (6)$$

Finalmente, la regla de discriminación está dada por

$$\begin{aligned} \text{Si } D > 0, \quad y &\in \text{ Clase 1} \\ \text{En otro caso, } y &\in \text{ Clase 2} \end{aligned} \quad (7)$$

EJEMPLOS NUMÉRICOS

Mostramos la aplicabilidad del método propuesto a través de una serie de ejemplos numéricos usando datos reales de EEG pertenecientes a la base de datos IVa de la competencia internacional BCI 2004 [16]. La base de datos corresponde a eventos de imaginación motora, y está compuesta por el registro de cinco voluntarios sanos (denominados aa, al, av, aw y ay) quienes fueron sentados cómodamente frente a una pantalla en la que se indicó durante 3.5s el movimiento a imaginar: mano derecha (denotado a partir de ahora como clase 1) o pies (denotado como clase 2). Las indicaciones fueron intercaladas por periodos de longitud variable (1.75-2.25s) en que se solicitó al voluntario que se relajara. La actividad cerebral fue registrada usando un arreglo de 118. Para cada sujeto se tienen 280 mediciones (140 por cada clase), con una frecuencia de muestreo de 1000Hz. Como parte del preprocesamiento realizamos un submuestreo a 250Hz y aplicamos un filtro pasabanda, con frecuencias de corte de 3Hz y 30Hz.

La mejor red de sensores se eligió evaluando la significancia estadística (con $\alpha=0.05$) según lo propuesto en [17]. La red de sensores con mayor significancia en una clase y menor en la otra, fue la seleccionada. Dentro de cada red, se evaluaron las coherencias en un rango de frecuencias de 1Hz a 30Hz, como se muestra en la Figura 1, a fin de encontrar el valor de fs (descrito en la sección de Métodos).

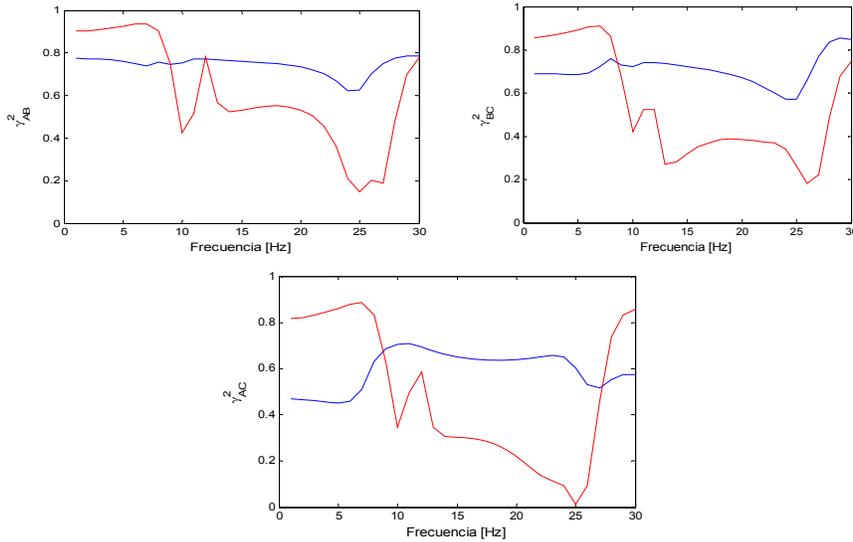


Figura 1. Coherencias en la red de electrodos seleccionada para el voluntario *al*.

En la Tabla 1 se presentan la combinación de sensores utilizada para cada voluntario, junto con la frecuencia seleccionada. Conociendo la mejor combinación de sensores y la frecuencia f_s , se formó el vector de características y a partir de (2), para cada medición de cada clase. La gráfica generada al proyectar estos vectores de características se muestra en la Figura 2, en la que se observa un patrón de agrupamiento para cada una de las clases. Dada esta distribución de los datos, se utilizó el discriminante no lineal de Fisher, descrito en la Sección 2.2.

Voluntario	Sensores	Frecuencia [Hz]
aa	F3, Fz, C3	16
al	F3, C3, Cz	13
av	Fz, Cz, C4	15
aw	Fz, F4, C4	13
ay	F3, Fz, C3	15

Tabla 1. Selección de sensores y frecuencias para cada voluntario

La clasificación se realizó entrenando el sistema con el 50% de los datos y el restante 50% fue utilizado para la evaluación. Para evaluar el desempeño del método de clasificación se utilizó el área bajo la curva ROC. El proceso de clasificación y evaluación se repitió 200 veces, cambiando en cada iteración el conjunto de datos de entrenamiento y evaluación. El área promedio obtenida para cada sujeto se presenta en la Tabla 2 junto con su desviación estándar. Igualmente, en la columna denominada "Área CSP" se muestran los resultados obtenidos para esta misma base de datos a partir del trabajo expuesto en [4]. Cabe destacar que, a pesar de que las áreas no disminuyen significativamente con el método propuesto, con la ventaja de que sólo se requiere el uso de 3 electrodos, los cuales han sido seleccionados en función de la conectividad entre regiones cerebrales para un estado mental determinado.

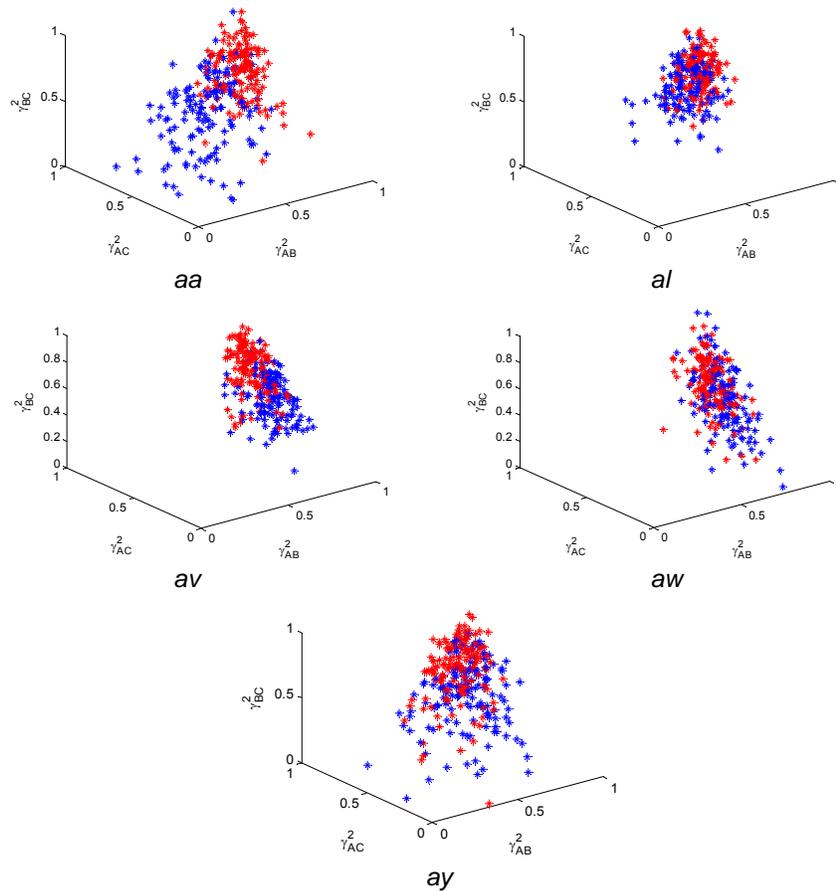


Figura 2. Gráfica del vector de características basado en la coherencia para cada voluntario.

Voluntario	Áreas coherencia	Áreas CSP
aa	0.783±0.026	0.813±0.049
al	0.945±0.013	0.958±0.031
av	0.732±0.030	0.694±0.027
aw	0.884±0.020	0.841±0.022
ay	0.771±0.030	0.861±0.038

Tabla 2. Área bajo la curva ROC

CONCLUSIONES

Presentamos un método que permite seleccionar una red de sensores mediante la cual es posible discriminar señales de EEG provenientes de dos diferentes estados mentales. Así mismo se presentó un método de clasificación basado en la coherencia entre mediciones en la red de sensores seleccionada. Probamos que el método propuesto basado en la coherencia tiene el potencial de clasificar señales de EEG logrando un buen número de clasificaciones correctas trabajando con un número reducido de sensores.



BIBLIOGRAFÍA

- [1] D. J. McFarland G. Pfurtscheller J. R. Wolpaw, N. Birbaumer y T. M. Vaughan, "Brain-computer interfaces for communication and control," *Clinical Neurophysiology*, vol. 113, pp. 767–791, Marzo 2002.
- [2] M. Fatourehchi A. Bashashati and R. K. Ward, "A survey of signal processing algorithms in brain-computer interfaces based on electrical brain signals," *Journal of neural engineering*, vol. 4, pp. R32–R57, Marzo 2007.
- [3] J. Miller-Gerking H. Ramoser and G. Pfurtscheller, "Optimal spatial filtering of single trial EEG during imagined hand movement," *IEEE Transactions on biomedical engineering*, vol. 8, pp. 441–446, Diciembre 2000.
- [4] D. Gutiérrez and R. Salazar-Varas, "Using eigenstructure decompositions of time-varying autoregressions in common spatial patterns-based EEG signal classification," *Biomedical signal processing and control*, vol. 7, pp. 622–631, Abril 2012.
- [5] T. Dickhaus C. Sannelli and S. Halder, "On optimal channel configurations for SMR-based brain-computer interfaces," *Brain Topography*, vol. 23, pp. 186–193, Febrero 2010.
- [6] R. K. Ward X. Yong and G. E. Birch, "Sparse spatial filter optimization for EEG channel reduction in brain-computer interface," in *International conference on acoustics, speech and signal processing*. IEEE, 2008, vol. XXXVII, pp. 417–420.
- [7] K. K. Ang M. Arvaneh, C. Guan and C. Quek, "Optimizing the channel selection and classification accuracy in EEG-based BCI," *IEEE Transactions on biomedical engineering*, vol. 58, pp. 1865–1873, Junio 2011.
- [8] Ch. Neuper G. Pfurtscheller, A. Stancák Jr., "Event-related synchronization (ERS) in the alpha band an electrophysiological correlate of cortical idling: A review," *International journal of psychophysiology*, vol. 24, pp. 39–46, Noviembre 1996.
- [9] S. Okazaki S.T. Okuhata and H. Maekawa, "EEG coherence pattern during simultaneous and successive processing tasks," *International Journal of Psychophysiology*, vol. 72, pp. 89–96, 2009.
- [10] L. B. Felix C. A. Maia A. P. Souza, S. A. S. Filho and C. J. Tierra-Criollo, "Classification of imaginary movements using the magnitude-squared coherence feature extractor," in *Biosignals and biorobotics conference*. IEEE, 2012, vol. III, pp. 1–6.
- [11] S. Sanei and J.A. Chambers, *EEG signal processing*, Wiley, 2007.
- [12] S. Weiss and H. M. Mueller, "The contribution of EEG coherence to the investigation of language," *Brain and language*, vol. 85, pp. 325–343, 2003.
- [13] T. Cooke and M. Peake, "The optimal classification using a linear discriminant for two point classes having known mean and covariance," *Journal of multivariate analysis*, vol. 82, pp. 379–394, Enero 2002.
- [14] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognition Letters*, vol. 27, pp. 861–874, 2006.
- [15] D. Gutiérrez and D. I. Escalona-Vargas, "EEG data classification through signal spatial redistribution and optimized linear discriminants," *Computer methods and programs in biomedicine*, vol. 97, pp. 39–47, 2010.
- [16] G. Curio G. Dornhege, B. Blankertz and K.R. Müller, "Boosting bit rates in non-invasive EEG single-trial classifications by feature combination and multi-class paradigms," *IEEE Transactions on biomedical engineering*, vol. 51, pp. 993–1002, Junio 2004.
- [17] L. Faes, G. D. Pinna, A. Porta, R. Maestri y G. Nollo, "Surrogate Data Analysis for Assessing the significance of the Coherence Function," *IEEE Transactions on biomedical engineering*, vol. 51, pp. 1156–1166, Julio 2004.