

ESTIMACIÓN DE POTENCIALES EVOCADOS AUDITIVOS MEDIANTE TRANSFORMADA WAVELET Y DESCOMPOSICIÓN MODAL EMPÍRICA. UN ESTUDIO COMPARATIVO

Auditory brainstem response estimation using wavelet transform and empirical mode decomposition. A comparative study

RESUMEN

Los potenciales evocados auditivos del tronco cerebral (PEATC) son usados frecuentemente para fines diagnósticos, sin embargo su obtención se hace difícil por estar inmersos en actividad eléctrica fisiológica y no fisiológica. Para estimar esta señal, la promediación coherente es la técnica más utilizada y aceptada. Sin embargo, ésta presenta algunas limitaciones. Por esto se hace necesario buscar alternativas para estimar eficientemente la señal de PEATC.

El propósito de este trabajo es evaluar la transformada wavelet y la descomposición modal empírica como herramientas para mejorar el desempeño de la promediación coherente de PEATC buscando reducir el tiempo para obtener un registro confiable.

PALABRAS CLAVES: Potenciales evocados auditivos, transformada discreta wavelet, descomposición modal empírica.

ABSTRACT

The auditory brainstem response (ABRs) are commonly used for diagnostic purposes, however, are very difficult to obtain because they are immersed in a large amount of physiological and non physiological electrical activity. To estimate these potentials the coherent average is the most utilized and accepted technique, but this type of processing presents some drawbacks. Because of these, it becomes necessary to find different alternatives to estimate the ABR signal.

The purpose of this paper is to assess the wavelet transform and the empirical mode decomposition as a tool for improving the performance of the averaging coherent ABRs seeking to reduce the amount of time needed to obtain a reliable record.

KEYWORDS: Auditory evoked potential, empirical mode decomposition, discrete wavelet transform.

1. INTRODUCCIÓN

La respuesta de las vías nerviosas auditivas a la aplicación de un estímulo acústico adecuado se denomina potenciales evocados auditivos (PEA). Es posible utilizar diferentes patrones de estimulación, ubicación de electrodos y ventanas de análisis del registro, obteniéndose señales con distintos tipos de información relativa a las vías nerviosas involucradas [1].

Los PEA registrados en el cuero cabelludo están inmersos en una suma de otros potenciales generados por diversas fuentes fisiológicas (electroencefalograma EEG, electromiograma EMG, electrocardiograma ECG, etc.) y no fisiológicas (tensión de 50 Hz, señales de radio y ruido de instrumentación, entre otros). El caso más desfavorable se presenta con los potenciales evocados auditivos de tronco cerebral (PEATC), cuya amplitud máxima es del orden de 1 μ V, mientras que la amplitud del ruido es casi dos órdenes de magnitud mayor (alrededor de 50 μ V). Esta baja relación señal a ruido

(RSR), sumado al hecho de que los espectros de frecuencias de la señal y del ruido están superpuestos, hace que sea imposible obtener los PEATC con un solo registro o época [2]

Un registro de PEATC consiste en una sucesión de picos correspondientes a la actividad eléctrica del nervio auditivo y de las vías auditivas ascendentes dentro del tronco cerebral en los 10 ms siguientes a la aplicación de un estímulo [3]. El análisis de los PEATC reviste gran importancia, ya que diferentes características de la señal (presencia o no de picos, valores de latencias y amplitudes) permiten realizar diagnósticos, como estimación del umbral de audición o localización de patologías auditivas, entre otros [4, 5].

La promediación coherente o sincronizada es la técnica usual para estimar los PEA, y consiste en aplicar sucesivos estímulos auditivos a la persona y registrar las respuestas a cada uno de ellos. A cada una de estas respuestas se le denomina época. Posteriormente, estas

ALEJANDRO HERRERA

Ingeniero Biomédico
Asistente de Investigación
Escuela de Ingeniería de Antioquia
(EIA)
bmalejandro@eia.edu.co

CAROLINA POSADA

Estudiante Ingeniería Biomédica
Escuela de Ingeniería de Antioquia
(EIA) – Universidad CES
bmcarop@eia.edu.co

RUBEN ACEVEDO

Bioingeniero, M. Sc.
Profesor Adjunto
Universidad Nacional de Entre Ríos
Argentina
racevedo@bioingenieria.edu.ar

épocas se promedian para mejorar la RSR. En este proceso se supone a la señal de PEA estable en su amplitud y fase; y la señal ruido al no estar correlacionada con el estímulo es lo suficientemente variable para considerarse aleatoria y de valor medio cero [6, 7]. La ec. (1) representa el proceso de promediación coherente, donde $x_k[n]$ es la k-ésima época del potencial registrado, $s[n]$ es la señal de PEATC, $r_k[n]$ es la k-ésima señal de ruido con varianza σ^2 y K es el total de épocas.

$$y_K[n] = \frac{1}{K} \cdot \sum_{k=1}^K x_k[n] = s[n] + \frac{1}{K} \cdot \sum_{k=1}^K r_k[n] \quad (1)$$

En este contexto $r_k[n]$ se considera como la suma de señales aleatorias estacionarias con distintas varianzas, las cuales se suman definiendo una varianza resultante σ^2 . Para cada instante n , el término de ruido puede ser interpretado como un estimador del valor medio de una muestra de K datos. Tal promedio resulta en una nueva variable aleatoria que tiene el mismo valor medio y una varianza igual a $\frac{\sigma^2}{K}$ produciéndose un mejoramiento en la RSR en un factor de \sqrt{K} [8].

Si bien la promediación coherente es ampliamente aceptada, presenta ciertas limitaciones. La más importante es que supone que la señal de potencial evocado se repite exactamente igual en cada época, lo cual no es cierto en la mayoría de los casos [3]. Otra limitación a considerar es el tiempo necesario para obtener una estimación de la señal; lo cual se hace evidente en los PEATC. En este caso, la RSR inicial puede llegar a ser de -26 dB (1:20) y para obtener una morfología clara y definida de la señal de PEATC la RSR debe acercarse a un valor mínimo de 6 dB (2:1). Para esto es necesario promediar alrededor de 1600 épocas, y suponiendo una frecuencia de presentación de estímulos habitual de 11,11 estímulos por segundo (e.p.s.), este proceso insume alrededor de 2.5 minutos. Es importante destacar que para confirmar la morfología de la señal se deben realizar dos registros para verificar la replicabilidad de la señal de potencial evocado, por lo cual el tiempo necesario al menos se duplica [8]. En consecuencia, es importante explorar mejoras a la promediación coherente que disminuyan el tiempo requerido para la estimación confiable de la señal de PEATC. En la literatura hay diversas alternativas propuestas, entre las que se encuentra la aplicación de la transformada wavelet (WT) [9, 10] y la descomposición modal empírica. La transformada wavelet continua de

una señal $x(t)$ se define como $C_{a,b} = \int_{-\infty}^{\infty} x(t) \cdot \psi_{a,b}^*(t) dt$, donde $\psi(t)$ es la función wavelet, $C_{a,b}$ son los coeficientes wavelet, a y b son los parámetros de escala y traslación respectivamente. Definiendo $a_j = 2^j$ y $b_{j,k} = 2^{j \cdot k}$ con j, k

$\in \mathbb{Z}$, se obtiene la transformada wavelet discreta (DWT) [11]. La DWT permite realizar un análisis multiresolución en tiempo y frecuencia de una señal y de esta manera extraer los coeficientes wavelets que la caracterizan. Éstos pueden ser extraídos en distintos niveles de descomposición utilizando algún criterio, por ejemplo relacionado con las características de tiempo y frecuencia de la señal [12].

La descomposición modal empírica (EMD, del inglés Empirical Mode Decomposition) es una nueva técnica desarrollada por Huang *et al* para analizar señales no lineales y no estacionarias. Este método descompone un conjunto de datos en un número finito, y a menudo pequeño, de funciones modales intrínsecas (IMFs, del inglés Intrinsic Mode Functions). Consiste en un proceso adaptativo y por lo tanto, altamente eficiente, basado en la extracción de la energía asociada a varias escalas de tiempo intrínsecas [13, 14].

En este trabajo se evalúa la DWT y la EMD como herramientas para mejorar la RSR de los promedios, con el fin de utilizar menos épocas en la estimación de la señal de PEATC

2. MATERIALES Y MÉTODOS

2.1 MATERIALES

Se utilizaron ocho registros de PEATC pertenecientes a una base de datos generada en el Laboratorio de Audiología de la Universidad Autónoma Metropolitana (Iztapalapa) [15]. Corresponden a registros de siete adultos jóvenes, entre 20 y 35 años de edad, con un análisis audiológico previo (otoscopia, audiometría y timpanometría) para determinar la normal audición de los mismos. Se realizaron con un sistema NeuroScan® y los parámetros utilizados se muestran en la Tabla 1.

Parámetros	Descripción
Estímulo	Click de rarefacción
Frecuencia de presentación	17.5 e.p.s.
Enmascaramiento	No
Electrodos	Cz(+) / M1(-) / Fpz(G)
Banda de paso	100 – 3000 Hz
Frecuencia de muestreo	20000
Épocas	4000

Tabla 1. Parámetros de registro de PEATC

2.2 MÉTODOS

La modelización de los PEATC utilizando DWT se realizó en 3 pasos: a) descomposición de la señal de PEATC hasta un nivel N . El espectro de frecuencias de los PEATC tiene una componente de bajas frecuencias alrededor de los 100Hz, otra de medias frecuencias

alrededor de los 500Hz y un componente de altas frecuencias alrededor de 1kHz [8], por lo que una descomposición como la mostrada en la Figura 1, resulta adecuada para analizar esta señal, b) umbralamiento para seleccionar los coeficientes y c) reconstrucción de la señal utilizando transformada discreta wavelet inversa (IDWT) con el conjunto de coeficientes modificados. La metodología utilizada para obtener los coeficientes se describe en [16]. Se utilizó como función wavelet la Biortogonal 5.5 debido a los buenos resultados obtenidos por Bradley *et al.* en el análisis multiresolución de potenciales evocados [17].

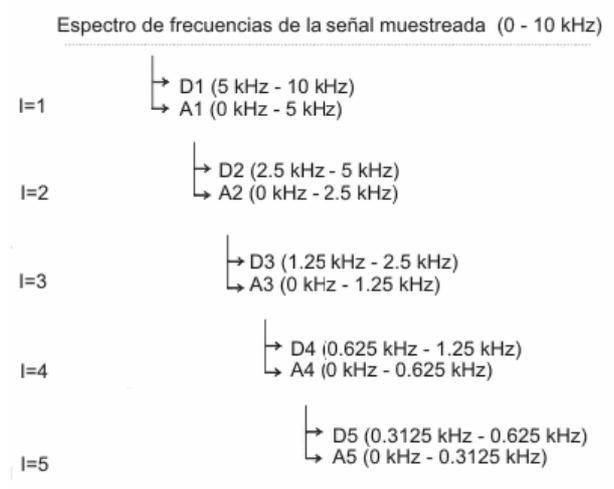


Figura 1. Descomposición de la señal de PEATC utilizando DWT.

La distribución de coeficientes obtenidos en el punto anterior puede verse como una máscara, la cual se aplica a un registro de PEATC para resaltar las características de la señal y disminuir el ruido simultáneamente.

Se realizó la descomposición modal empírica de un promedio coherente de 2000 épocas. A cada uno de los IMFs generados en el proceso se les estimó el espectro de frecuencias mediante periodograma y se comparó la distribución frecuencial de cada uno con el espectro de frecuencias del promedio de 2000 épocas, Figura 2. Después de encontrar la banda de frecuencia de cada modo resultante de la EMD, se usó esta información para reconstruir la señal con la suma de los IMFs que contenían las frecuencias de interés. La metodología utilizada para la selección de IMFs se describe en [18].

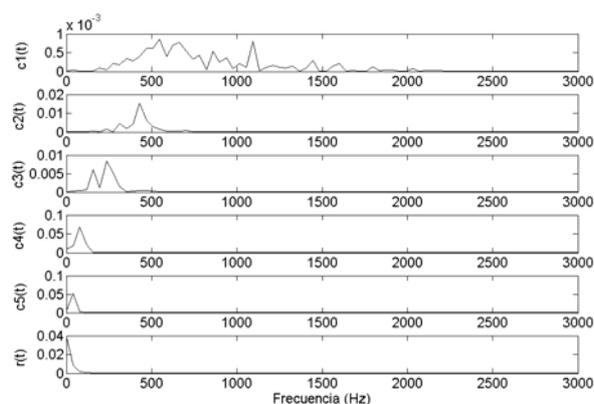


Figura 2. Periodogramas de los IMFs provenientes de una descomposición modal de un registro de PEATC de 2000 épocas.

Se utilizaron los promedios coherentes de 200, 400, 600, 800 y 1000 épocas. En el caso del análisis con DWT se descompuso la señal para obtener los coeficientes wavelet que los caracterizan, luego de esto, se le aplicó la máscara de coeficientes obtenida previamente y se realizó la IDWT. Para el análisis con EMD se realizó la descomposición modal de cada uno de estos registros, se reconstruyó a partir de los IMFs cuyo espectro de frecuencia corresponde con el espectro de frecuencia de la señal de PEATC. Ambos procesos se pueden ver como un *denoising* de las señales, eliminando elementos ruidosos de la señal, y resaltando características importantes de la señal.

Para analizar el desempeño en cada caso se evaluó el error cuadrático medio (ECM) y el coeficiente de correlación (ρ) entre el promedio coherente de 2000 épocas y la señal resultante del proceso de *denoising*. Debido a que las latencias de los picos III y V son los parámetros clínicos más importantes [8], adicionalmente se estimaron estas para cada caso. Para esto se utilizaron 8 registros, los registros de aquellos sujetos donde alguno de los picos no fuera distinguible o no estuviera presente no fueron tenidos en cuenta.

3. RESULTADOS

3.1 DENOISING DE LAS ÉPOCAS

En las Figuras 3 y 4 se muestran las curvas del ρ y ECM entre el promedio coherente de 2000 épocas y las señales resultantes del proceso de *denoising* utilizando la DWT.

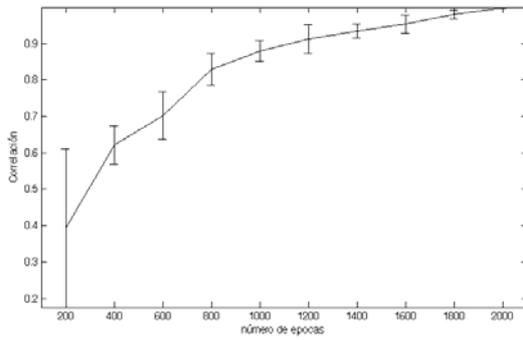


Figura 3. Evaluación de ρ medios y sus respectivas desviaciones estándar para diferentes cantidades de épocas aplicando mascara de coeficientes obtenidos a partir de DWT.

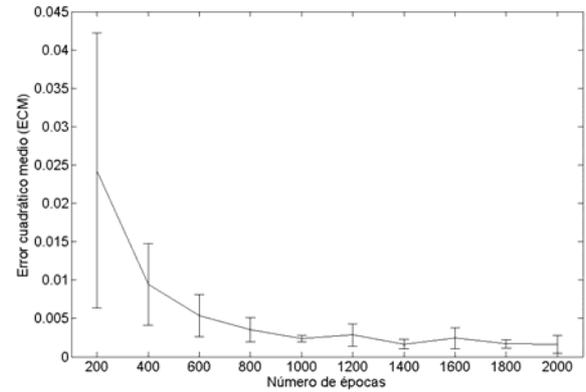


Figura 6. ECM y respectivas desviaciones estándar de las reconstrucciones realizadas a cada registro utilizando EMD.

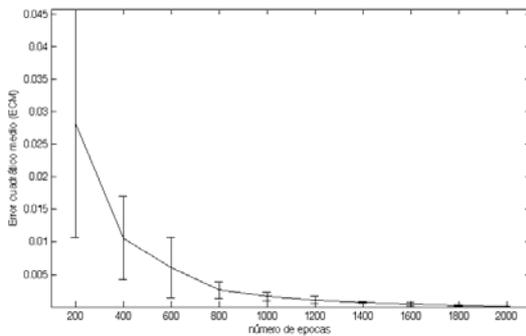


Figura 4. ECM y respectivas desviaciones estándar para diferentes cantidades de épocas aplicando mascara de coeficientes obtenidos a partir de DWT.

En las figuras 5 y 6 se muestran las curvas del ρ y ECM entre el promedio coherente de 2000 épocas y las señales resultantes del proceso de *denoising* utilizando la EMD como herramienta de análisis.

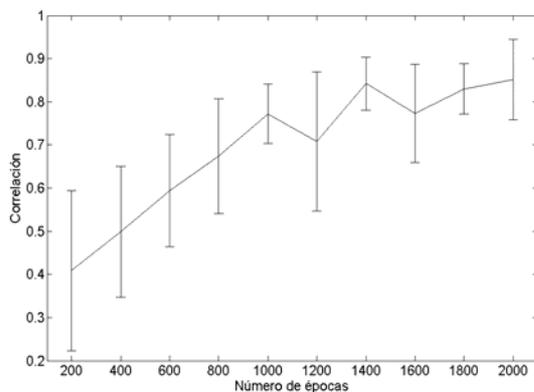


Figura 5. Evaluación de ρ medios y respectivas desviaciones estándar de las reconstrucciones realizadas a cada registro utilizando EMD.

3.2 ANÁLISIS DE LATENCIAS

La Tabla 2 muestra las latencias medias para los picos III y V con la desviación estándar correspondiente de las reconstrucciones hechas utilizando la mascara de coeficientes hallada utilizando la DWT. Para 200 y 400 épocas algunos picos fueron poco distinguibles o ausentes, por lo cual se descartaron estos valores.

Épocas	Pico III [ms]	Pico V [ms]
200*	3.466 ± 0.130	5.492 ± 0.136
400**	3.433 ± 0.198	5.438 ± 0.119
600	3.462 ± 0.192	5.523 ± 0.166
800	3.456 ± 0.143	5.550 ± 0.162
1000	3.434 ± 0.152	5.548 ± 0.128
2000°	3.438 ± 0.051	5.570 ± 0.079

*Pico III y Pico V ausente o no distinguibles en un registro.

**Pico III ausente o no distinguible en un registro

° Promediación coherente sin aplicar *denoising*

Tabla 2. Latencias de los picos III y V de las reconstrucciones hechas utilizando la DWT.

En la Tabla 3 se muestran las latencias medias y sus desviaciones estándar para los picos III y V de las reconstrucciones realizadas utilizando la EMD. Para 200, 400 y 600 épocas algunos picos no fueron distinguibles o ausentes en el registro, por lo que estos valores no fueron tenidos en cuenta para realizar los cálculos.

Épocas	Pico III [ms]	Pico V [ms]
200*	3,594 ± 0,204	5,569 ± 0,295
400**	3,625 ± 0,290	5,581 ± 0,162
600**	3,544 ± 0,246	5,438 ± 0,213
800	3,488 ± 0,125	5,613 ± 0,074
1000	3,463 ± 0,127	5,694 ± 0,176
2000 °	3.438 ± 0.051	5.570 ± 0.079

*Pico III ausente o no distinguibles en un registro.

**Pico III y Pico V ausente o no distinguible en un registro

° Promediación coherente sin aplicar *denoising*

Tabla 3. Latencias de los picos III y V de las reconstrucciones hechas utilizando la EMD.

4. CONCLUSIONES

Como se puede notar las correlaciones no son tan altas para la reconstrucción de la señal utilizando EMD, esto se debe a que con la eliminación del primer IMF para la reconstrucción se pierde la información de frecuencia alta de la señal. Sin embargo para efectos prácticos, en la mayoría de los casos se mantienen las latencias de los picos III y V. Basados en esto y en la tendencia a la estabilidad de la correlación y el error cuadrático medio (Figura 3 y Figura 4), se encontró que para un promedio de 800 épocas, la reconstrucción obtenida ya era aceptable. Esto apoyado en el hecho que para 200, 400 y 600 épocas algunos picos ya no eran distinguibles o incluso ausentes en los registros analizados.

Basados en la morfología de la onda y en el análisis de latencias de los picos III y V, las pruebas realizadas para determinar el número mínimo de épocas para modelar correctamente la señal utilizando la DWT, mostraron que un número adecuado sería de 800 a 1000 épocas de promediación lo que eventualmente ayudaría a disminuir el tiempo de registro. Incluso con 600 épocas se podría tener una reconstrucción mejor que la obtenida con 800 épocas utilizando la EMD, debido a que con 600 épocas utilizando la DWT se pudieron evaluar todas las latencias de los registros, y los valores de correlación y ECM estuvieron mas alto y cercano respectivamente que aquellos encontrados utilizando la EMD.

Las razones para esto es que con los IMFs (2, 3 y 4) escogidos para la reconstrucción, se esta eliminando un contenido importante del componente de altas frecuencias y de medias frecuencias, por esta razón el coeficiente de correlación presenta una tendencia mas inestable que el hallado mediante la DWT. Esto se evidencia en el coeficiente de correlación del promedio de 2000 épocas de la señal aplicando estrategia de *denoising* con EMD vs. la señal de 2000 épocas sin aplicar *denoising*. En este punto se observa que eliminando el primer IMF se esta eliminando información útil para el estudio.

En el análisis de la señal utilizando DWT se utiliza para la reconstrucción los coeficientes de las sub-bandas correspondientes al contenido frecuencial de la señal de PEATC. Adicionalmente realizando una selección de los

coeficientes se puede tener un acceso simultáneo a la escala (frecuencia) y tiempo, este análisis multiresolución resulta optimo para este tipo de señales no estacionarias y eventualmente ayudaría a reducir el tiempo necesario para obtener un estimación confiable de la señal (600 o 800 épocas es aceptable).

Como trabajo a futuro se planea usar estas técnicas para extracción de características de señales de PEATC que sirvan para reconocer promedios con o sin respuesta, lo cual sería muy útil en el diagnóstico de hipoacusias en neonatos así como también implementación de interfaces cerebro computadoras.

5. REFERENCIAS

- [1] Misulis K., Spehlmann's Evoked Potentials: Primer Visual, Auditory and Somatosensory Evoked Potentials in Clinical Diagnosis, Butterworth-Heinemann, 1994.
- [2] Niedermeyer E., Lopes da Silva F., Electroencephalography: basic principles, clinical applications and related fields, Lippincott Williams & Wilkins, 2005.
- [3] Keith C. Evoked Potentials in Clinical Medicine. Lippincott-Raven, 3° ed., 1997.
- [4] Jacobson J. The Auditory Brainstem Response. College Hill Press, 1985.
- [5] Niedermeyer E, Lopes da Silva F (2005) Electroencephalography: basic principles, clinical applications and related fields. Lippincott Williams & Wilkins.
- [6] Biurrún J., Escobar S., Acevedo R. Mejoramiento de la relación señal a ruido en potenciales evocados mediante filtrado digital. XV Congreso Argentino de Bioingeniería, Argetina.
- [7] Torres A., Pérez M., Sotero R., Rioja L., Eimil E. Caracterización de la relación señal/ruido de los potenciales evocados auditivos de estado estable durante la promediación. Revista CENIC Ciencias Biológicas, 35, 191-196, 2004.
- [8] Acevedo R., Bonell C., Gentiletti G. et al. Potenciales evocados auditivos de tronco cerebral: Aspectos tecnológicos del registro. Revista Argentina de Bioingeniería (SABI), 12, 10-17, 2006.
- [9] Sita G., Ramakrishnan A. G., Wavelet Domain Nonlinear Filtering for Evoked Potential Signal Enhancement, Comput Biomed Res 33: 431-446, 2000.
- [10] Quian R., Garcia H., Single-trial event-related potentials with wavelet denoising, Clin Neurophysiol 114: 376-390, 2003.
- [11] Samar V., Rao R., Swartz K., Wavelet Analysis of Neuroelectric Waveforms: A Conceptual Tutorial, Brain Lang 66: 7-60, 1999.
- [12] Maglione J. L., Pincilotti M., Acevedo R., Estimation of the Auditory Brainstem Response's wave V by means of Wavelet Transform, Procs. of

- 25 th Annual International Conf. of the IEEE Eng. Med Biol. Society, Cancun, Mexico, 2003.
- [13] Liang H., Bressler S., Desimone R., Fries P. Empirical mode decomposition: a method for analyzing neural data. *Neurocomputing*, 65-66, 801-807, 2005.
- [14] N.E. Huang, Z. Shen, S.R. Long, M.L. Wu, H.H. Shih, Q. Zheng, N.C. Yen, C.C. Tung and H.H. Liu, "The empirical mode decomposition and Hilbert spectrum for nonlinear and nonstationary time series analysis," *Proc. Roy. Soc. London A*, Vol. 454, pp. 903-995, 1998.
- [15] Gentiletti G, Yañez O, Cornejo J (2003) Evaluation of Automatic Identification Algorithms for Auditory Brainstem Response used in Universal Hearing Loss Screening. En *Proceedings of 25 th Annual International Conf. of the IEEE Eng. Med Biol. Society, Cancun, Mexico, 2003*.
- [16] Herrera A., Acevedo R., Biurrun J. Promediación coherente mejorada mediante transformada wavelet de potenciales evocados auditivos de tronco cerebral. en *IFMBE Proceedings 18, IV Congreso latinoamericano de Ingeniería Biomédica 2007*, pp 183-187.
- [17] Bradley A P, Wilson W J (2004) On Wavelet Analysis of auditory evoked potentials. *Clin Neurophysiol* 115: 1114-1128.
- [18] Posada C., Acevedo R. Mejoramiento de la Relación Señal a Ruido de Potenciales Evocados Auditivos mediante Descomposición Modal Empírica. En proceso de publicación en *Revista Ingeniería Biomédica EIA*.