

# Estimación de potenciales evocados auditivos del tronco cerebral mediante descomposición modal empírica

Carolina Posada Zuluaga<sup>1</sup>, Rubén C. Acevedo<sup>2</sup>

*1 Programa de Ingeniería Biomédica. Escuela de Ingeniería de Antioquia-Universidad CES, Colombia*

*2 Facultad de Ingeniería, Universidad Nacional de Entre Ríos, Argentina*

Recibido 12 de febrero de 2008. Aceptado 19 de mayo de 2008.

---

**Resumen**— Los potenciales evocados auditivos del tronco cerebral (PEATC) son frecuentemente usados para fines diagnósticos; sin embargo, su procesamiento se hace difícil porque están inmersos en una gran cantidad de ruido proveniente no solo de fuentes externas sino también fisiológicas. Hasta ahora el método más utilizado y aceptado para obtener un registro confiable es la promediación coherente, aunque presenta algunos inconvenientes. La descomposición modal empírica (EMD) es una técnica relativamente nueva que se usa para el procesamiento de señales no estacionarias como la mayoría de señales fisiológicas. Este método separa una señal, extrayendo la energía asociada a diferentes escalas de tiempo intrínsecas, en una suma finita de modos oscilatorios. El propósito de este trabajo fue evaluar la EMD como una herramienta para mejorar el desempeño de la promediación coherente de PEATC buscando reducir la cantidad de épocas necesarias para obtener un registro confiable. Para tal fin se analizó la reconstrucción de ocho registros usando solamente los modos 2, 3 y 4 resultantes de la EMD, los estudios determinaron que una reconstrucción de 800 épocas es aceptable.

**Palabras clave**— Descomposición modal empírica, Eliminación de ruido, Funciones modales intrínsecas, Potenciales evocados auditivos del tronco cerebral.

**Abstract**— The brainstem auditory evoked potentials (BAEPs) are commonly used for diagnostic purposes; however, processing becomes difficult because they are immersed in a large amount of noise coming not only from external sources but also from other physiological sources. So far the most widely used and accepted method to obtain reliable recording is the coherent averaging, but this type of processing presents some drawbacks. The empirical mode decomposition (EMD) is a relatively new technique which is used for processing of non-stationary signal like almost physiological signals. This method separates a signal, **extracting the energy** associated with various intrinsic time scales, into a finite set of oscillatory modes. **The purpose of the study was to assess the EMD as a tool for improving the performance of the averaging coherent BAEPs seeking to reduce the amount of epochs needed to obtain a reliable register. To this end, we have analyzed the reconstruction of eight registers using only modes 2, 3 and 4 resulting from the EMD, the studies found that a reconstruction of 800 epochs is acceptable.**

**Keywords**— Empirical mode decomposition, Denoising, Intrinsic mode functions, Brainstem auditory evoked potentials.

---

## I. 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 50/60 Hz, señales de radio y ruido de instrumentación, entre otros) [2]. Entre los PEA, se encuentran los potenciales evocados auditivos del tronco cerebral (PEATC), cuyo análisis 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 [3].

Un registro de PEATC consiste en una sucesión de picos correspondientes a la actividad eléctrica de las fibras del nervio auditivo y de las vías auditivas ascendentes dentro del tronco cerebral en los 10 ms siguientes a la aplicación de un estímulo [4], cuya amplitud máxima es del orden de 1  $\mu$ V, mientras que la amplitud del ruido es casi dos órdenes de magnitud mayor (alrededor de 50  $\mu$ V) [3]. 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 [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 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  $\sigma^2$  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  $\sum_{k=1}^K r_k[n]$  se considera como la suma de señales aleatorias estacionarias con distintas varianzas, las cuales se suman definiendo una varianza resultante  $\sigma^2$ . 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. Además, este tiempo es variable, ya que depende en cada caso de la amplitud de la respuesta del tronco cerebral, del ruido presente en el registro y de la colaboración del sujeto [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.

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 [9, 10].

El objetivo de este trabajo es estimar PEATC mediante la EMD de diferentes promedios con el fin de evaluar la técnica como herramienta para mejorar el desempeño de la promediación coherente. En este contexto, se midió la latencia de los picos III y V y la correlación de diferentes promedios procesados con EMD, para ser comparados con los mismos parámetros de registros de 2000 épocas procesados con promediación coherente.

## II. MATERIALES Y MÉTODOS

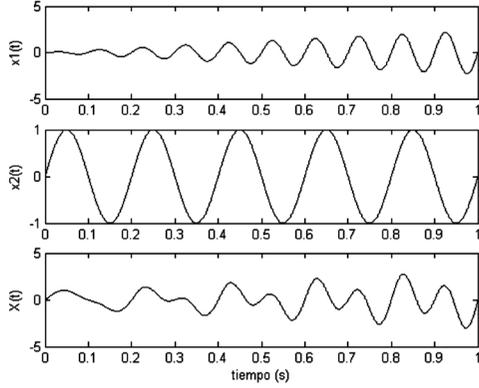
### *Descomposición Modal Empírica (EMD)*

Como ya se mencionó, la EMD separa una secuencia de datos en una suma de componentes (IMFs) y cada uno de estos contiene diferentes modos oscilatorios. Las funciones base o IMFs son derivadas directamente de la misma señal y solo se requiere que la señal tenga mínimo dos extremos: un máximo y un mínimo [9-12, 16].

La descomposición se hace de acuerdo a los lapsos de tiempo entre los mínimos y máximos locales; y los cruces consecutivos por cero. Así, una señal debe cumplir dos condiciones para ser un IMF: a) el número de extremos y el número de cruces por cero deben ser iguales o diferentes

por uno y b) en cualquier punto, el valor de la media de los envolventes definidos por los máximos y mínimos locales debe ser cero [9].

Por ejemplo, dada una señal  $X(t)$  que es la suma de dos señales, una modulada en amplitud y otra senoidal, como en la Fig. 1; el proceso de descomposición modal consiste en:



**Fig. 1.** Arriba, señal modulada en amplitud  $x_1(t) = 4 \text{ sen}(20\pi t) \text{ sen}(2\pi t/10)$ ; en medio, señal senoidal  $x_2(t) = \text{sen}(10\pi t)$  y abajo  $X(t) = x_1(t) + x_2(t)$ .

a) Identificar los mínimos y máximos locales e interpolar por medio de funciones splines cúbicas, así se obtiene una envolvente superior  $e_{\text{sup}}(t)$  y otra inferior  $e_{\text{inf}}(t)$ . Estas dos envolventes deben contener todos los datos entre sí.

b) La media entre estas dos envolventes es designada como:

$$m_1(t) = \frac{e_{\text{sup}}(t) + e_{\text{inf}}(t)}{2} \quad (2),$$

y luego se halla el residuo entre la señal original y la media,

$$x(t) - m_1(t) = h_1(t) \quad (3)$$

Lo ideal es que  $h_1$  fuera un IMF, pero antes debe cumplir las condiciones ya mencionadas, así que se sigue iterando. En la segunda iteración,  $h_1$  es tomada como la señal ( $h_1 - m_{11} = h_{11}$ ) y este procedimiento se puede realizar  $k$  veces ( $h_{1(k-1)} - m_{1k} = h_{1k}$ ) hasta obtener un IMF  $c_1 = h_{1k}$ .

c) El residuo entre los datos iniciales y el IMF resultante,

$$r(t) = X(t) - h_{1k}(t) \quad (4)$$

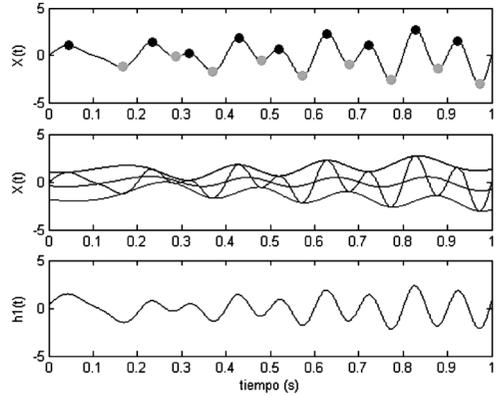
es tratado como una nueva señal para seguir haciendo el proceso y hallar otro IMF.

Este algoritmo continúa hasta que el residuo se hace muy pequeño y no tiene puntos de inflexión. Al final del proceso la señal puede ser expresada de la siguiente manera:

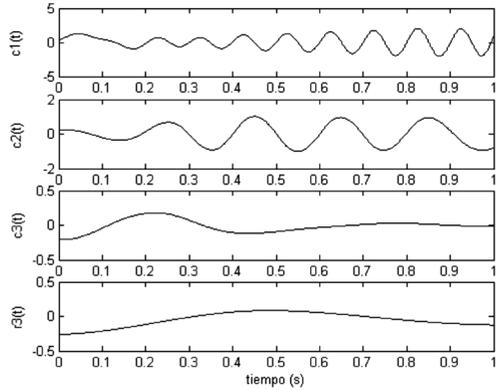
$$x(t) = \sum_{i=1}^k c_i(t) + r_k(t) \quad (5)$$

donde  $c_i(t)$  son los diferentes IMFs,  $k$  es el número total de IMFs y  $r_k(t)$  es el residuo final. Los IMFs son casi ortogonales entre sí y tienen media cero [9], por lo tanto, cada IMF contiene un ancho de banda en frecuencia y amplitud de la señal original [9, 10].

En la Fig. 2 se muestra la primera iteración al realizar descomposición modal de una señal que es la suma de dos señales, una de amplitud modulada y otra senoidal y los IMFs resultantes y el residuo final se muestran en la Fig. 3.



**Fig. 2.** Primera iteración de la EMD de la señal  $X(t)$ . Arriba se muestran los máximos y mínimos locales en azul y rojo respectivamente, al medio las envolventes superior (azul) e inferior (rojo) y la media entre ellas (magenta) y abajo el residuo resultante de la resta de la media de las envolventes y la señal original  $X(t)$ .



**Fig. 3.** IMFs y residuo final resultantes de la descomposición modal de  $X(t)$ .

Para investigaciones basadas en este método Huang *et al*, ha desarrollado varias funciones en Matlab. Una de ellas, *emd.mat*, tiene como parámetros de salida los IMFs propios de la señal, los índices de ortogonalidad y el número de iteraciones por cada modo; introduciendo un vector fila, que corresponde a la señal que se desea analizar y, otros parámetros que dependen de las necesidades del usuario.

Para lograr el objetivo propuesto se hizo primero una etapa de caracterización de los registros de PEATC mediante EMD y luego se procedió a la denoising (eliminación del ruido) de los mismos.

#### A. Caracterización de los PEATC

En el estudio, se trabajó con ocho registros de PEATC pertenecientes a una base de datos generada en el Laboratorio de Audiología de la Universidad Autónoma Metropolitana (Iztapalapa)[13]. 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.

Tabla 1. Parámetros de registro de PEATC

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 Hz
Épocas	4000

Se realizó la descomposición modal de un promedio de 2000 épocas, y esta se muestra Fig. 4. A los IMFs generados en el proceso se les estimó el espectro de frecuencias mediante periodograma (Fig. 6); y se comparó la distribución frecuencial de cada uno de estos con el espectro de frecuencias del promedio de 2000 épocas (Fig. 5).

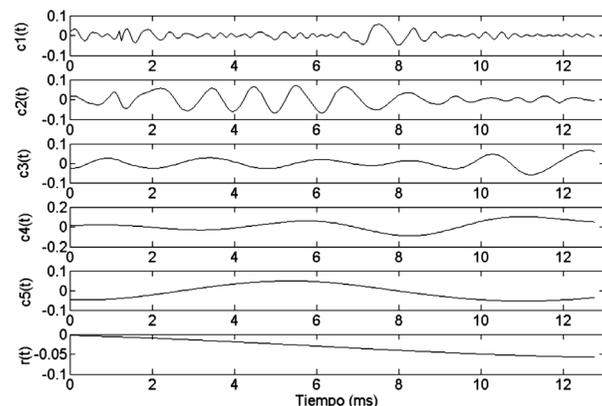


Fig. 4. Modos y residuo resultantes de la descomposición modal de un registro de PEATC de 2000 épocas.

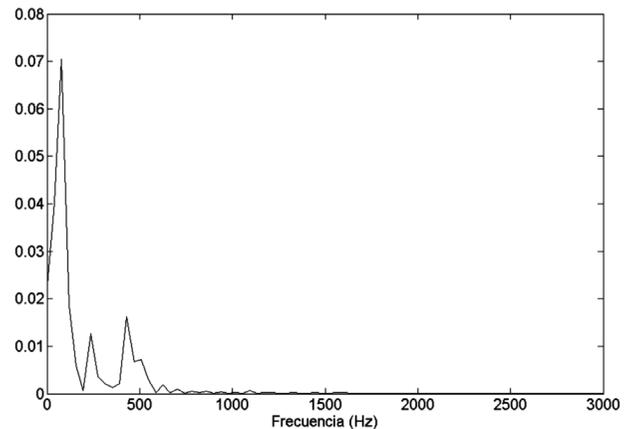


Fig. 5. Periodograma de un registro PEATC de 2000 épocas.

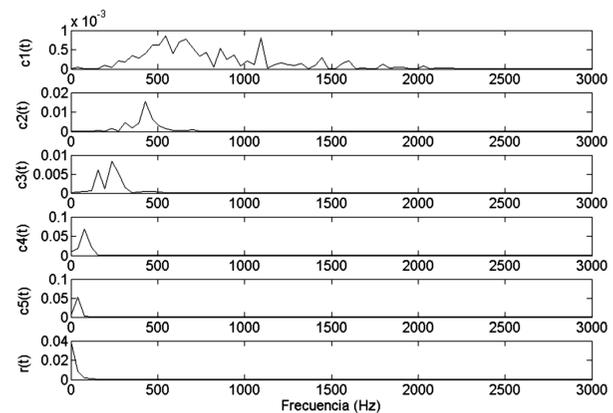


Fig. 6. Periodogramas de los IMFs provenientes de una descomposición modal de un registro de PEATC de 2000 épocas.

#### B. Denoising de PEATC

Se hizo promediación coherente de los registros de PEATC con diferentes cantidades de épocas (200, 400, 600, 800, 1000, 1200, 1400, 1600, 1800 y 2000) y a los promedios resultantes se les aplicó la función de EMD desarrollada por Huang *et al.*

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.

#### C. Análisis estadístico

En la práctica clínica, para fines diagnósticos, uno de los parámetros más utilizados es la latencia (tiempo de aparición del pico) de los picos III y V [2], por lo tanto se estimó la latencia de estos. De esta manera se halló una latencia media con su respectiva desviación estándar de las reconstrucciones realizadas a las diferentes cantidades de épocas con los ocho registros.

También se halló la correlación entre cada reconstrucción y la promediación coherente de 2000 épocas que se uso como referencia.

Las promediaciones, la descomposición modal empírica, el análisis espectral y todo el estudio en sí, fueron ejecutados en Matlab 7.0.1®.

### III. RESULTADOS Y DISCUSION

En la mayoría de los casos la señal se descomponía en 4 ó 5 modos más el residuo (Fig. 4), esto se debe a la distribución espectral de una señal de PEATC, que tiene picos sobresalientes a los 100, 500 y 1000 Hz (frecuencia baja, media y alta respectivamente) como se puede observar en la Fig. 5. Cada una de estas regiones aporta energía a determinados picos de la señal. La energía de los picos I y III corresponde en parte a la región de frecuencia media y en parte a la de frecuencia alta. La mayor parte de la energía del pico II corresponde a la componente de frecuencia alta y la energía del pico IV corresponde a la componente de frecuencia media, al igual que la mayor parte de la energía del pico V [8, 14, 15].

En la Fig. 4, correspondiente a la descomposición modal de un promedio de 2000 épocas de PEATC, se puede observar que en el primer modo se concentran las oscilaciones rápidas mientras que en los últimos las señales se hacen cada vez más suaves. Esto se puede comprobar al hacer el análisis frecuencial de estos modos (Fig. 6). Después de analizar los espectrogramas de los modos resultantes al hacer la descomposición modal a señales promediadas de 200, 400, 600, 800, 1000, 1200, 1400, 1600, 1800 y 2000 épocas; se encontró que en todos los casos el primer modo era el que introducía más ruido, conteniendo frecuencias por encima de los 1000 Hz; por lo tanto se fijó reconstruir la señal teniendo en cuenta solo los modos 2, 3 y 4.

Los modos 2, 3 y 4 aportan a la señal sus frecuencias características, aunque el primer modo también aporta a la frecuencia alta de la señal original, su mayor contenido es ruido.

Debido a que una de las variables más significativas a la hora de evaluar un PEATC es la latencia de los picos, esta variable fue la más determinante durante el estudio. En las Tablas 2 y 3 se muestran las latencias medias y sus desviaciones estándar para los picos III y V de las reconstrucciones realizadas, a partir de las reconstrucciones con 800 épocas se comienzan a estabilizar y la diferencia respecto al promedio coherente de 2000 épocas es muy pequeña. En la Tabla 4, se muestran las correlaciones

( $\rho$ ) de las reconstrucciones realizadas respecto a la promediación coherente de 2000 épocas.

**Tabla 2.** Latencias medias y respectivas varianzas del pico III

Épocas	Latencia media (ms)	$\sigma$	Épocas	Latencia media (ms)	$\sigma$
200	3,594	0,204	1200	3,519	0,175
400	3,625	0,290	1400	3,444	0,199
600	3,544	0,246	1600	3,438	0,192
800	3,488	0,125	1800	3,363	0,230
1000	3,463	0,127	2000*	3,469	0,059

Se muestran las latencias medias del pico III y sus desviaciones estándar para las reconstrucciones realizadas teniendo como base los IMFs 2, 3 y 4 de las promediaciones de las diferentes cantidades de épocas para cada registro.

\*Promediación coherente de 2000 épocas.

**Tabla 3.** Latencias medias y respectivas varianzas del pico V

Épocas	Latencia media (ms)	$\sigma$	Épocas	Latencia media (ms)	$\sigma$
200	5,569	0,295	1200	5,675	0,144
400	5,581	0,162	1400	5,688	0,155
600	5,438	0,213	1600	5,606	0,296
800	5,613	0,074	1800	5,663	0,323
1000	5,694	0,176	2000*	5,681	0,156

Se muestran las latencias medias del pico V y sus desviaciones estándar para las reconstrucciones realizadas teniendo como base los IMFs 2, 3 y 4 de las promediaciones de las diferentes cantidades de épocas para cada registro.

\*Promediación coherente de 2000 épocas.

**Tabla 4.** Correlaciones medias y respectivas desviaciones estándar de las reconstrucciones realizadas a cada registro con las diferentes cantidades épocas.

Épocas	$\rho$ media	s	Épocas	$\rho$ media	s
200	0,409	0,186	1200	0,708	0,161
400	0,499	0,151	1400	0,842	0,061
600	0,594	0,130	1600	0,773	0,114
800	0,674	0,133	1800	0,830	0,058
1000	0,772	0,069	2000*	0,851	0,094

\*Promediación coherente de 2000 épocas

### IV. CONCLUSIÓN

Como se puede notar, las correlaciones no son tan altas, 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, que fueron el parámetro más determinante.

Basados en esto, se encontró que para un promedio de 800 épocas (Fig. 7), la reconstrucción obtenida ya era aceptable.

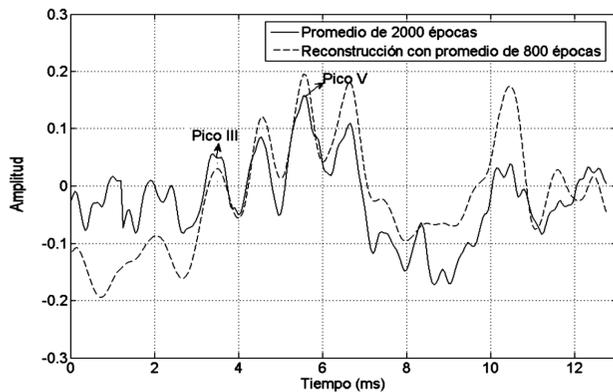


Fig. 7. En línea continua se muestra un registro de PEATC de 2000 épocas y en la línea discontinua una reconstrucción con los IMFs 2, 3 y 4 después de realizar la descomposición modal a un registro de PEATC de 800 épocas.

El procesamiento de señales mediante EMD, es un área en la que todavía hay mucho por investigar. Como trabajos a futuro se planea usar EMD como técnica para extracción de características de señales de PEATC que sirvan para reconocer señales con o sin respuesta, lo cual sería muy útil en el diagnóstico de patologías. También se podrían analizar los resultados con otras técnicas de EMD como EMD on-line y EMD local, entre otras mencionadas por Kan Zeng et al., Cheng Juncheng et al., Gabriel Rilling et al., Patrick Flandrin et al., etc.

## REFERENCIAS

[1]. Misulis K. Spehlmann's Evoked Potentials: Primer Visual, Auditory and Somatosensory Evoked Potentials in Clinical Diagnosis. *Butterworth-Heinemann*, 1994.

[2]. Rockville, MD. *Short latency auditory evoked potentials*, American Speech-Language-Hearing Association, 1987.

[3]. Keith C. *Evoked Potentials in Clinical Medicine*. Lippincot-Raven, 3° ed., 1997.

[4]. Jacobson J. *The Auditory Brainstem Response*. College Hill Press, 1985.

[5]. Niedermeyer E., Lopes da Silva F. *Electroencephalography: basic principles, clinical applications and related fields*, Lippincott Williams & Wilkins, 2005.

[6]. Biurrun J., Escobar S., Acevedo R. Mejoramiento de la relación señal a ruido en potenciales evocados mediante filtrado digital. *XV Congreso Argentino de Bioingeniería*, Argentina.

[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 premedicación. *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]. Liang H., Bressler S., Desimone R., Fries P. Empirical mode decomposition: a method for analyzing neural data. *Neurocomputing*, 65-66, 801-807, 2005.

[10]. Huang N.E., Shen Z., Long S.R., Wu M.L., Shih H.H., Zheng Q., Yen N.C., Tung C.C., Liu H.H.. The empirical mode decomposition and Hilbert spectrum for nonlinear and nonstationary time series analysis, *Proc. Roy. Soc. London A*, 454, 903-995, 1998.

[11]. Rösler H.A. Study on the Empirical Mode Decomposition. Tesis de Grado, Maestría, University of Amsterdam, 2002.

[12]. Deléchelle E., Lemoine J., Niang O. Empirical Mode Decomposition: An Analytical Approach for Sifting Process. *IEEE Signal Processing Letters*, 12, 764-767, 2005.

[13]. Gentiletti G., Yañez O., Cornejo J. Evaluation of Automatic Identification Algorithms for Auditory Brainstem Response used in Universal Hearing Loss Screening. *Procs. of 25 th Annual International Conf. of the IEEE Eng. Med Biol. Society*, Cancun, Mexico, 2003.

[14]. Boston J. Spectra of Auditory Brainstem Responses and Spontaneous EEG. *IEEE Trans. On Biomed. Eng.*, 36, 407-414, 1998.

[15]. Hernández F., Flores T., Peñalosa Y. Instituto Nacional de Comunicación Humana. Registros electrofisiológicos para el diagnóstico de la patología de la comunicación humana, México, D.F; Secretaría de Salud, 327, 1996.

[16]. Rilling G., Flandrin P., Goncalves P. On Empirical Mode Decomposition and its Algorithms, *IEEE Sig. Proc. Lett.*, 2003.