



Universidad de Valparaíso
Facultad de Ingeniería
Escuela de Ingeniería Civil Biomédica

**Inspección De un Criterio y Metodología de
Selección de Bases Wavelet para Aplicación en
Señales de EKG**

Por

Diego Alonso Pardo Arcaya

Trabajo de Título para optar al Título de
Ingeniero Civil Biomédico

Prof. Guía: MSc. César Galindo Viaux

Marzo 2015

Dedicatoria

A Yiyo, David y Gonzalo, mis pilares.

Agradecimientos

Deseo agradecer a los académicos que a lo largo de mi vida han ayudado a forjar mi pensamiento y mi persona.

Quiero agradecer a mi profesor de pre-cálculo, don Isaac Acevedo por ayudarme sin pensar en recibir.

Deseo agradecer también a los académicos del Departamento de Ingeniería Biomédica de la Universidad de Valparaíso, pero en especial a Steren Chabert, a Cesar Galindo, a Rodrigo Salas y a Guillermo Avendaño, por su paciencia conmigo y por creer en mí.

Deseo también agradecer a mis amigos, pues estas páginas no existirían y no sería lo mismo sin ellos: Laura, Aland y Esteban.

Deseo también agradecer a mi familia por siempre estar a mi lado, por ser un núcleo tan cálido que me ha formado para ser siempre la mejor persona posible.

La pelea nunca ha sido tan dura, pero ustedes siempre estuvieron ahí para ayudarme.

Gracias.

Resumen

Palabras claves: Base Wavelet Óptima, Criterio de Selección, Extracción de Características.

Resumen: La Transformada de Wavelet es una herramienta muy versátil, pues tiene muchas bases con las que puede funcionar. No obstante, cuando se aplica la transformada a una señal, cada base tendrá un resultado distinto. La selección de la base wavelet óptima cobra relevancia, y ha sido abordada por distintos autores. En este trabajo, la selección de la base wavelet óptima se lleva a cabo buscando la similitud entre la función de escala madre y una señal EKG. Se extraen parámetros de dominio de frecuencia y de tiempo, y se crea un vector descriptor. El descriptor se aplica a las funciones de escala de wavelet y a las señales, y sus resultados se concentran en una matriz. Se aplica PCA a ésta última para obtener proyecciones de cada descriptor y se miden las distancias entre las proyecciones de las bases wavelet y las señales. La base con la menor distancia a la señal será la más similar. La aplicación de la metodología a dos grupos de complejos QRS, arrojó que las más similares fueron Db4 y Db9. Se compararon éstas con otras bases antes consideradas óptimas para EKG por otros autores, y sus resultados fueron similares. Por último, se validó los resultados de la metodología por medio de denoising de un registro EKG. Se concluyó que la similitud entre la base y la señal en gran parte hace a una base óptima respecto a otra, y permite desarrollar aplicaciones de búsqueda de patrones con mejores resultados.

Tabla de Contenidos

1. INTRODUCCIÓN	7
1.1. OBJETIVO PRINCIPAL	8
1.2. OBJETIVOS ESPECÍFICOS	8
2. ANÁLISIS DE LA PROBLEMÁTICA	9
3. MARCO TEÓRICO	10
3.1. TRANSFORMADA DE WAVELET	10
3.2. CARACTERÍSTICAS DE DESCRIPTORES	13
4. METODOLOGÍA	17
4.1. SEÑALES A ANALIZAR	18
4.1.1. COMPLEJOS QRS	18
4.1.2. BASES WAVELET	18
4.2. SELECCIÓN DE CARACTERÍSTICAS	19
4.3. MATRIZ DE DESCRIPTORES	20
4.4. PCA Y GENERACIÓN DEL SUB ESPACIO	21
4.5. MEDICIÓN DE DISTANCIAS	22
4.6. VALIDACIÓN DE BASES WAVELET ESCOGIDAS	23
5. RESULTADOS	24
5.1. GENERACIÓN DE SUB ESPACIOS BIH Y ESC	25
5.2. DISTANCIAS MEDIDAS EN SUB ESPACIOS BIH Y ESC	25
5.3. VALIDACIÓN DEL PROCESO DE SELECCIÓN DE BASES WAVELET	26
6. DISCUSIÓN	26
7. CONCLUSIONES	29
REFERENCIAS BIBLIOGRÁFICAS	31
ANEXOS	34

Inspección de un Criterio y Metodología de Selección de Bases Wavelet para Aplicación en Señales de EKG

Inspección de un
Criterio y
Metodología de
Selección de
Bases Wavelet
para Aplicación
en señales de
EKG

7

Diego Alonso Pardo Arcaya

Escuela de Ingeniería Civil Biomédica, Universidad de Valparaíso, Chile

Palabras claves: Wavelet, Criterio de Selección, Extracción de Características, Base Wavelet Óptima, Función de Escala Madre.

1. Introducción

La Transformada de Wavelet es una herramienta muy utilizada desde que se hizo conocida por Yves Meyer, Ingrid Daubechies y Stéphane Mallat, en la década de los 90 (Daubechies 1992, Meyer 1992, Mallat 1999), a pesar que la historia sitúa su nacimiento en la década de los 70. Lo cierto es que se ha vuelto muy popular gracias a la posibilidad de llevar a cabo un análisis a señales en tiempo y en frecuencia: podemos conocer qué frecuencias se desarrollan en qué período de tiempo, y por su buen desempeño en eliminación de ruido de señales, o *denoising*, y compresión de señales e imágenes (Haidekker 2011, Tang 2009).

Sin embargo, la característica que más llama la atención de esta transformada es su versatilidad a la hora de analizar una señal. La transformada de wavelet tiene como bases a un conjunto de funciones finitas, llamadas wavelet. Existe un repertorio muy grande de bases wavelet: entre las más conocidas se encuentran las familias Daubechies, Symmlet y Coiflet, y las bases únicas Morlet y Mexican Hat, entre muchas otras. Todas ellas comparten ciertas condiciones para ser consideradas wavelets, pero son muy distintas entre sí. Ésta es una de sus grandes fortalezas, pero a la vez una fuerte debilidad. Al aplicar la transformada de wavelet a una señal utilizando distintas bases wavelet, el resultado es distinto. Ante esto, una pregunta aparece: ¿qué base wavelet usar?

Muchos autores han intentado responder esta pregunta, ya sea creando una base wavelet que se ajuste a sus necesidades, o buscando la base que les entregue el mejor resultado posible. El proceso de crear una base wavelet para un problema en específico indica que existe una relación en la forma de la base y la señal o problema a analizar. De hecho, se sabe que una base wavelet considerada “óptima” genera coeficientes con valores significativamente altos y localizados, permitiendo detectar patrones en los coeficientes de las señales

analizadas (Takla 2006, Chapa 2000). Esto hace importante encontrar una metodología para definir la wavelet “óptima”.

El presente trabajo propone el comienzo de una serie de investigaciones para construir una metodología que permita identificar la base wavelet más adecuada para la señal a analizar. El concepto detrás de esta metodología es que se trabaje “a priori”. Actualmente, las investigaciones buscan aplicar la transformada a una señal específica, calcular un indicador desde los resultados y comparar algún parámetro para escoger la base wavelet adecuada. Este tipo de metodología la hemos llamado “a posteriori”, pues requiere aplicar la transformada para obtener la información necesaria. La metodología presentada propone seleccionar la base wavelet buscando similitudes entre la base y la señal. La forma de establecer una relación de similitud apunta a la extracción de características para convertirlas en un vector y compararlo con los vectores generados para otras señales. Así, se pueden comparar señales que pueden tener amplitudes y largos muy dispares, pero una misma morfología. Para desarrollar la metodología, se utilizará la plataforma MatLab (MatLab 2013a, TheMathWorks Inc. Natick, MA, USA)

El presente trabajo se divide en 7 secciones. La primera, descrita en la sección 2, busca adentrar al lector en las aproximaciones que otros autores han desarrollado, y dar luces de la metodología propuesta. En la sección 3 se entregan conceptos previos para entender el por qué se busca la similitud de señales. En la sección 4 se presenta la metodología, y en la sección 5, sus resultados. En la sección 6, se encuentra la discusión de los resultados, y por último, en la sección 7, las conclusiones de la experiencia.

1.1. Objetivo General

El objetivo general del presente trabajo es establecer una metodología de selección de la base wavelet más adecuada para analizar con la transformada de wavelet una señal de electrocardiografía (EKG). La metodología estará basada en encontrar la función de escala madre más similar a la señal de EKG. La relación de similitud se establecerá encontrando la menor distancia entre ambas funciones en un sub espacio generado por características temporales y de frecuencia que den cuenta de la forma de la onda de la función de escala de wavelet y la señal de EKG.

1.2. Objetivos Específicos

- a. Estudiar trabajos que tengan relación con la selección de bases wavelet más adecuadas basados en extracción de características, y escoger aquellas que describan la morfología de la forma de onda de una señal.
- b. Extraer características de señales de EKG y las funciones de escala madre que han sido etiquetadas como “óptimas” para analizar esas señales, y crear un vector que describa la forma de onda de la señal y de la función de escala madre.
- c. Buscar la menor distancia existente entre las señales de EKG y las funciones de escala madre en un sub espacio de características.

-
- d. Validar la metodología de selección, aplicando la transformada de wavelet en la señal de EKG con las bases wavelet que tienen la menor distancia respecto a la señal EKG.
 - e. Discutir los resultados y concluir.

2. Análisis de la problemática

El problema de la selección de la base adecuada para una señal en específico ha sido abordado por varios autores. De Sobral Cintra et al. (De Sobral Cintra 2004) en su trabajo aplican la transformada de wavelet para compresión de señales de electrogastrografía (EGG) humanas y caninas. Los investigadores emplearon distintas bases wavelet, y usaron la diferencia porcentual de la raíz cuadrática media en los resultados como una tasa de distorsión entre la señal original y su versión comprimida. Singh y Tiwari aplicaron *denoising* con bases wavelet ortogonales a registros EKG, y determinaron que la base wavelet con mejores resultados fue Daubechies 8 (Db8), al analizar los resultados minimizando el error de la raíz cuadrática media y la norma L1, entre otros (Singh 2006). Rafiee et al. concluyeron que la base Daubechies 44 (Db44) es adecuada para electromiografía (EMG), electroencefalografía (EEG) y amplitud de pulso vaginal (VPA) (Rafiee 2011). Además, indican que por su forma podría entregar buenos resultados al aplicarse en EKG. Aplicando *denoising* a registros EKG, Kania et al. plantean que Daubechies 1 (Db1) (también conocida como Haar) y Symmlet 3 (Sym3) entregan buenos resultados, si se utilizan al menos 4 niveles de descomposición piramidal (Kania 2007).

Un escenario similar se presenta al comprimir imágenes usando la transformada de wavelet. Nagaria et al. (Nagaria 2011) buscaron la base wavelet más adecuada en compresión de imágenes, estudiando las bases más utilizadas, como Haar, Daubechies, Symmlets, Coiflets, Meyer, Spline biortogonales y Spline inversas (Strang 1996). Para identificar la mejor base wavelet, utilizaron la minimización del error cuadrático medio, la maximización de la tasa de compresión y la maximización del valor *peak* de la tasa señal a ruido.

Estos métodos presentan dos elementos que son considerados problemáticos: una cantidad de bases wavelet acotada y una aproximación “a posteriori”.

Todos los métodos citados buscan aplicar la transformada de wavelet con una base particular y calcular algún indicador en sus resultados, para luego comparar este valor con el obtenido por la utilización de todas las demás bases wavelet. Al hecho de aplicar la transformada y obtener información desde sus resultados se le llamó “a posteriori”. Esta metodología reduce el alcance del estudio a las bases wavelet que sean incorporadas, además de estar sujeta a la capacidad de cálculo del estudio: de forma arbitraria, MatLab tiene implementada la familia Daubechies desde la base Db1 hasta Db45, mientras que en realidad, existen infinitas versiones para generar.

Por otra parte, existen muchas bases wavelet y se crean nuevas de forma constante. De los trabajos nombrados anteriormente, casi la totalidad utilizaron las bases más conocidas (Daubechies, Symmlet, Coiflet, *Mexican hat*). Todo esto condiciona que las bases wavelet más adecuadas para una señal sean escogidas de un grupo acotado, dejando de lado muchas que pueden entregar un mejor resultado con la misma señal.

De los trabajos nombrados anteriormente, sólo Rafiee et al. consideraron 324 bases de distintas familias: Haar, Daubechies, Symmlet, Coiflet, *Gaussian*, Morlet, Morlet compleja, *Mexican hat*, Spline biortogonal, Spline inverso, Meter, Meyer discreta, *Gaussian* compleja, Shannon y B-Spline frecuencial. No obstante, utilizan un método “a posteriori”, por lo que no consideró nuevas bases wavelet.

La presente propuesta se basa en el trabajo de De Chazal et al. (De Chazal 2004), de Phinyomark et al. (Phinyomark 2012, Phinyomark 2013). De Chazal et al. extrajeron características a registros EKG para clasificar sus frecuencias cardíacas mediante análisis de discriminantes lineales (LDA). Si bien no utilizan la transformada de wavelet, buscan caracterizar morfológicamente los registros EKG. Phinyomark et al. buscan clasificar distintos registros de EMG extrayendo distintas características y analizándolas con LDA. De Chazal et al. busca características propias de una señal EKG, como duración del complejo QRS, duración de la onda T o promedio de los intervalos RR, mientras que Phinyomark et al. poseen una lista de 50 características que incluyen algunas propias de EMG y otros valores estadísticos. De estos valores, algunos fueron escogidos para ser utilizados en el presente estudio.

3. Marco Teórico

A continuación se darán a conocer elementos importantes para comprender la metodología, los resultados y las conclusiones. En la sección 3.1. se describe la Transformada de Wavelet y su relación con las distintas bases wavelet. En la sección 3.2. se describen las características que serán utilizadas para conformar los descriptores.

3.1. Transformada de Wavelet

Una *Wavelet* es una función corta, oscilante y finita, que está modificada por un parámetro de escala, s , y por un parámetro de traslación o desplazamiento, u . Esto implica que posee la capacidad de dilatarse o comprimirse y de desplazarse. Su nombre viene del vocablo *wave* (onda en inglés) y el sufijo *-let* que indica que algo es corto, una versión pequeña; es decir, una “*ondita*”. La wavelet (formalmente, función de wavelet) corresponde a

$$\psi_{u,s}(t) = \frac{1}{\sqrt{s}} \psi\left(\frac{t-u}{s}\right), \quad Ec. 1$$

donde s es el parámetro de escala, u es el parámetro de desplazamiento, t es el tiempo y ψ es la función de wavelet madre, que corresponde a la wavelet primordial sin modificaciones.

La Transformada de Wavelet es una transformada matemática que permite descomponer una señal en la suma de muchas wavelets comprimidas y desplazadas. Esta transformada toma como supuesto inicial que todas las señales son cortas y cambian constantemente de manera brusca (Bedoya 2006). La suposición le da la capacidad de representar una zona específica de la señal con una o más wavelets, confiriéndole una muy buena resolución en tiempo, como también en frecuencia. La transformada de wavelet, matemáticamente, es la correlación entre la señal y una wavelet,

$$Wf(u, s) = \int_{-\infty}^{+\infty} f(t) \psi_{u,s}(t) dt, \quad Ec. 2$$

donde $f(t)$ es la señal, s es el parámetro de escala, u es el parámetro de desplazamiento, t es el tiempo y $\psi_{u,s}(t)$ es la función de wavelet.

Un factor clave de esta transformada es la versatilidad de las funciones de wavelet madre: existen muchas funciones capaces de ser una función de wavelet madre, aunque no cualquiera puede serlo. Para esto, se debe cumplir con la Condición de Admisibilidad (Peters 1998) dada por la desigualdad

$$\int_{-\infty}^{+\infty} \frac{|\Psi(\omega)|^2}{|\omega|} dt < \infty, \quad Ec. 3$$

donde $\Psi(\omega)$ es la transformada de Fourier de la función de wavelet madre, ω es la frecuencia y t es el tiempo. De esta desigualdad, se desprende que:

- a. La integral de la función de wavelet $\psi(t)$ debe ser 0.
- b. Debe tener soporte finito: debe existir un intervalo en el que la función tiene valor distinto a cero; y fuera de ese intervalo, la función debe tener valor cero.
- c. Debe ser finita: su energía es limitada y debe decaer.

Existen muchas funciones que pueden ser clasificadas como funciones de wavelet madre. Las más conocidas son Haar, Daubechies, Symmlet y Coiflet (Daubechies 1992, Mallat 1999).

De acuerdo a la ecuación *Ec.2*, la transformada de wavelet es continua, pues es una integral en el tiempo. Así, se puede representar una señal con wavelets modificadas por escalas continuas y desplazamientos infinitesimales. Pero en este estudio, y para ser capaces de analizar señales digitales, se requiere trabajar con la forma discreta de la transformada. Es decir, se trabajará con desplazamientos y escalas en valores discretos.

Para llevar a cabo el proceso de cambio de continua a discreta, se introducen dos conceptos clave: la Red Diádica y el Diezmado. Para hacer discreta la escala, la forma más común de hacerlo es mediante una red diádica (Faundez 2001, Burrus 1998, Chui 1997), donde la escala pasa a ser una base de 2. Es decir,

$$s = 2^i, \quad i \in \mathbb{Z}, \quad Ec. 4$$

donde i es el nuevo parámetro de escala de la wavelet. Luego, se hace discreto el desplazamiento. La variación del parámetro u en una escala no es igual en una escala mayor o menor; por esto, se hace necesaria una modificación en función de la escala en la que se desplace. El diezmo logra que la traslación vaya de acuerdo a la escala. Esto se logra con

$$u = 2^i \cdot n, \quad i, n \in \mathbb{Z}, \quad Ec. 5$$

donde i es el nuevo parámetro de escala y n es el nuevo parámetro de desplazamiento. Por último, se cambia la variable de tiempo t por la variable discreta k . Reemplazando las variables antes descritas en las ecuaciones *Ec.1* y *Ec.2*, la transformada de wavelet discreta queda

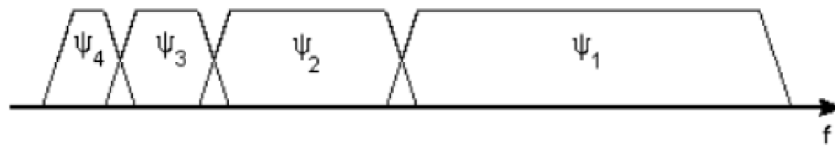
$$W(2^i, 2^i n) = \sum_k \Psi_{i,n}^* \cdot f(k), \quad \Psi_{i,n}^*(k) = \frac{1}{\sqrt{2^i}} \psi\left(\frac{k}{2^i} - n\right), \quad Ec. 6$$

$$i, n \in \mathbb{Z}$$

donde i es el parámetro de escala discreta, n es el desplazamiento discreto, $f(k)$ es la señal discreta, k es el tiempo discreto y $\Psi_{i,n}^*$ es la función de wavelet en forma discreta. Notar que $\Psi_{i,n}^*$ no es lo mismo que $\psi_{u,s}$.

La red diádica, al escalar en potencias de 2, logra que cada vez que se avance una escala, se reduzca a la mitad el espectro en frecuencia que la función de wavelet analiza (Bedoya 2006). En la figura 1 se muestra un esquema de la forma en que se aborda el espectro en frecuencia con cada nivel que se avanza. En cada nivel i , la función de wavelet analiza la mitad del espectro con las frecuencias más altas, dejando la otra mitad del espectro para ser dividida nuevamente en el nivel $i+1$.

Figura 1: División del espectro en frecuencia por análisis diádico (Bedoya 2006).



Cuando se aplica la transformada de wavelet, se está aplicando una correlación (como se ve en la *Ec.2*) entre la función de wavelet y la señal (Tang 2009). La correlación con la función de wavelet madre modificada genera el análisis de la mitad del espectro con frecuencias altas; por tanto, el análisis de la mitad del espectro sobrante es realizado por otra función. Ésta es llamada Función de Escala Madre, y es el complemento ortogonal de la función de wavelet madre (Faundez 2001). Las dos funciones en conjunto son llamadas Base Wavelet.

De acuerdo a lo explicado anteriormente, la función de wavelet madre actúa como un filtro pasaaltos, mientras que la función de escala madre actúa como un filtro pasabajos. La aplicación de las dos funciones en una señal es lo que se conoce como Algoritmo Piramidal (Faundez 2001, Sede Bogota DNSAV). Al ser aplicado, cada filtro genera coeficientes que (gracias al diezmado) corresponden a la mitad del largo total de la señal analizada. Los coeficientes que genera el filtro pasaaltos es llamado Detalle, y los que genera el filtro pasabajos es llamado Aproximación. En la figura 2 se puede ver el proceso esquematizado.

Inspección de un
Criterio y
Metodología de
Selección de
Bases Wavelet
para Aplicación
en señales de
EKG

Como se dijo más arriba, la función de wavelet madre está unida a la función de escala madre. Por lo tanto, cada una de las wavelets existentes posee ambas funciones. En el presente estudio, nos centraremos de forma específica en la función de escala madre, pues representa el filtro pasabajos; es decir, la aproximación de la señal.

$$\begin{aligned}
 \text{ESCALA 0} \quad & f(t) = \left[\sum_{i=0}^7 S_i^{[0]} \phi_i^{[0]}(t) \right] \\
 \text{ESCALA 1} \quad & f(t) = \sum_{i=0}^3 d_i^{[1]} \psi_i^{[1]}(t) + \left[\sum_{i=0}^3 S_i^{[1]} \phi_i^{[1]}(t) \right] \\
 \text{ESCALA 2} \quad & f(t) = \sum_{i=0}^3 d_i^{[1]} \psi_i^{[1]}(t) + \sum_{i=0}^3 d_i^{[2]} \psi_i^{[2]}(t) + \left[\sum_{i=0}^3 S_i^{[2]} \phi_i^{[2]}(t) \right] \\
 \text{ESCALA 3} \quad & f(t) = \sum_{i=0}^3 d_i^{[1]} \psi_i^{[1]}(t) + \sum_{i=0}^3 d_i^{[2]} \psi_i^{[2]}(t) + \sum_{i=0}^3 d_i^{[3]} \psi_i^{[3]}(t) + \left[\sum_{i=0}^3 S_i^{[3]} \phi_i^{[3]}(t) \right]
 \end{aligned}$$

Figura 2: Esquema de Algoritmo Piramidal. Se aprecia cómo en cada nivel la Aproximación es separada en una nueva Aproximación y un Detalle (Galindo 2009).

3.2. Características de Descriptores

La caracterización de las funciones de escala de las bases wavelet se realiza a través de la construcción de los descriptores. El descriptor es un vector conformado por características que den cuenta de la forma de la onda a analizar. De esta forma, dos señales o funciones que visualmente son similares tendrían dos descriptores o puntos que en el sub espacio de características se encontrarían a una distancia cercana. Para crear los descriptores, se buscaron parámetros e indicadores en distintas investigaciones donde el objetivo fue identificar la base wavelet óptima para alguna aplicación y señal específica, entendiendo que se la identificó entre un grupo acotado de bases. Se escogió como base la lista que Phinyomark et al. (Phinyomark 2012) desarrollaron, con 50 características generales sobre la forma de la onda e información propia de la señal analizada.

De todas éstas, en este estudio se utilizarán sólo algunas de las características generales de dominio temporal y de dominio de las frecuencias. Las características en el dominio del tiempo escogidas para el presente trabajo se explican a continuación.

Promedio: valor estadístico, entrega información sobre la tendencia de la señal sobre o bajo la línea del cero.

Valor Medio Absoluto: (MAV, por sus siglas en inglés) es un promedio de los valores absolutos de la señal. Es un indicador comúnmente empleado en análisis de EMG (Huang 2013) (Phinyomark 2013). Se utilizará como comparación frente al Promedio. Matemáticamente, corresponde a

$$MAV = \frac{1}{N} \sum_{i=1}^N |x_i|, \quad Ec. 7$$

donde N es el largo de la señal, $i=1,2,3,\dots,N$ y x_i es el valor de la señal en el instante i .

Momentos Centrales Normalizados: valores utilizados en trabajos de Hamedi et al., Phinyomark et al. y Saridis & Gootee (Hamedi 2012, Phinyomark 2013, Saridis 1982), sobretodo en señales de EMG. Por trabajar en estas señales, dada su alta frecuencia, utilizaron los valores absolutos de los momentos para evitar valores negativos en los momentos impares. En el presente trabajo, se busca detallar la forma de la señal por lo que los momentos no serán utilizados con valor absoluto. Los momentos centrales normalizados segundo, tercero y cuarto son ampliamente utilizados en estadística y análisis de señales: varianza, sesgo y curtosis.

Además, dos trabajos extienden la utilización de los momentos centrales normalizados, dejando abiertas las posibilidades de utilizar hasta un k -ésimo momento. En el trabajo de Phinyomark et al., los autores emplean el quinto momento central normalizado como complemento a las características a extraer de señales de EMG, mientras que Samanta y Al-Balushi (Samanta 2003) analizaron las vibraciones de rodamientos en maquinarias, capturando señales vibratorias para determinar el nivel de desgaste de piezas y posible cambio de éstas. Una de las características extraídas corresponde al sexto momento central normalizado. Basado en estas experiencias, es que en el presente trabajo se utilizarán desde el segundo momento hasta el décimo.

Cruces por Cero: entrega una medida de frecuencia aproximada, contando las veces que la línea del cero es cruzada. Phinyomark et al. utilizan una fórmula que modificaron, añadiendo un umbral para hacer más robusta la característica frente a una señal ruidosa. Dado que en las señales utilizadas aquí, el ruido se ha eliminado previamente, se obviará la utilización del umbral dejándolo igual a cero (a esta fórmula la llamaremos ZC1). Por otra parte, Deng et al. (Deng 2008) trabajaron en clasificación de instrumentos musicales, extrayendo para esto distintos indicadores de dominio temporal y de frecuencia. Entre ellos, rescatan una tasa de cruces por cero, donde normalizan la cuenta de pases por cero con la cantidad de puntos analizados. Llamaremos a esta forma ZC2. En otra investigación, Tzanetakis y Cook (Tzanetakis 2002) buscan clasificar géneros musicales extrayendo características y analizándolas con técnicas de *Machine Learning*. El cruce por

ceros es obtenido por evaluación de valores positivos contra negativos con la función *sign*, donde ante valores positivos de x , incluyendo cero, *sign* entrega 1, y ante valores negativos de x , entrega 0. Llamaremos a esta forma ZC3. Ante la variedad de fórmulas, se utilizarán las tres buscando estudiar su desempeño en señales de EKG y funciones de escala madre. Las representaciones matemáticas son

$$ZC1 = \sum_{i=1}^N \text{sign}(x_i \cdot x_{i+1}), \quad \text{Ec. 8}$$

$$ZC2 = \frac{\sum_{i=1}^N |\text{sign}(x_i) - \text{sign}(x_{i-1})|}{2N}, \quad \text{Ec. 9}$$

$$ZC3 = \frac{1}{2} \sum_{i=1}^N |\text{sign}(x_i) - \text{sign}(x_{i-1})|, \quad \text{Ec. 10}$$

donde N es el largo de la señal, $i=1,2,3,\dots,N$ y x_i es el valor de la señal en el instante i .

Longitud de Forma de Onda: (WL, por sus siglas en inglés) corresponde a la longitud acumulativa de la forma de onda de la función analizada. Se utiliza como una medida de complejidad de la señal (Phinyomark 2012). Matemáticamente corresponde a

$$WL = \sum_{i=1}^{N-1} |x_{i+1} - x_i|, \quad \text{Ec. 11}$$

donde N es el largo de la señal, $i=1,2,3,\dots,N$ y x_i es el valor de la señal en el instante i .

Raíz Cuadrática Media: en electrónica análoga, es un estimado de la potencia entregada por una señal sinusoidal (i.e. electricidad doméstica). Respecto a señales en general, es una estimación de la magnitud de la función analizada que se utiliza de manera muy amplia.

Relación de Áreas Positiva y Negativa: esta característica es la única del conjunto de descriptores que fue creada exclusivamente para el estudio. Corresponde a un cuociente entre la suma y la resta de las áreas positivas y negativas de la señal, cuyo valor indica hacia qué lado se encuentra más inclinada respecto a la línea del cero (eje de las abscisas). Se le llamó *posneg* (positivo-negativo), y matemáticamente, se puede expresar como

$$\text{posneg}(X) = \frac{\sum(x_i \cdot \text{sign}^+(x_i)) - \sum(x_i \cdot \text{sign}^-(x_i))}{\sum(x_i \cdot \text{sign}^+(x_i)) + \sum(x_i \cdot \text{sign}^-(x_i))}, \quad \text{Ec. 12}$$

$$\text{sign}^+(x) = \begin{cases} 1, & \text{si } x \geq 0 \\ 0, & \text{si } x < 0 \end{cases}, \quad \text{sign}^-(x) = \begin{cases} 1, & \text{si } x < 0 \\ 0, & \text{si } x \geq 0 \end{cases}, \quad \text{Ec. 13}$$

donde $X=[x_1, x_2, x_3, \dots, x_n]$ es la señal a analizar, y sign^+ y sign^- son funciones que actúan en función del signo de x . El dominio de posneg corresponde a $\text{dom}=\mathbb{R}$ y su recorrido corresponde a $\text{rec}=[0,1]$.

Si la señal tiene mayor área positiva que negativa (es decir que tiene un sesgo hacia arriba de la línea del cero), posneg será positivo. Si el área negativa desaparece, dejando sólo al área positiva, posneg será muy cercano o igual a 1. Si la señal tiene mayor área negativa que positiva, el valor de posneg será negativo. Si el área positiva desaparece, dejando sólo al área negativa, posneg será muy cercano o igual a -1. Por último, cuando las dos áreas sean iguales, posneg será igual a 0. La figura 3 muestra este comportamiento de forma gráfica, para una mejor comprensión. En el anexo 1 se encuentra la función programada para ser utilizada en MatLab. Así, la función posneg entrega información valiosa sobre la distribución de la señal respecto al eje de las abscisas, complementando y cerrando las características del dominio temporal.

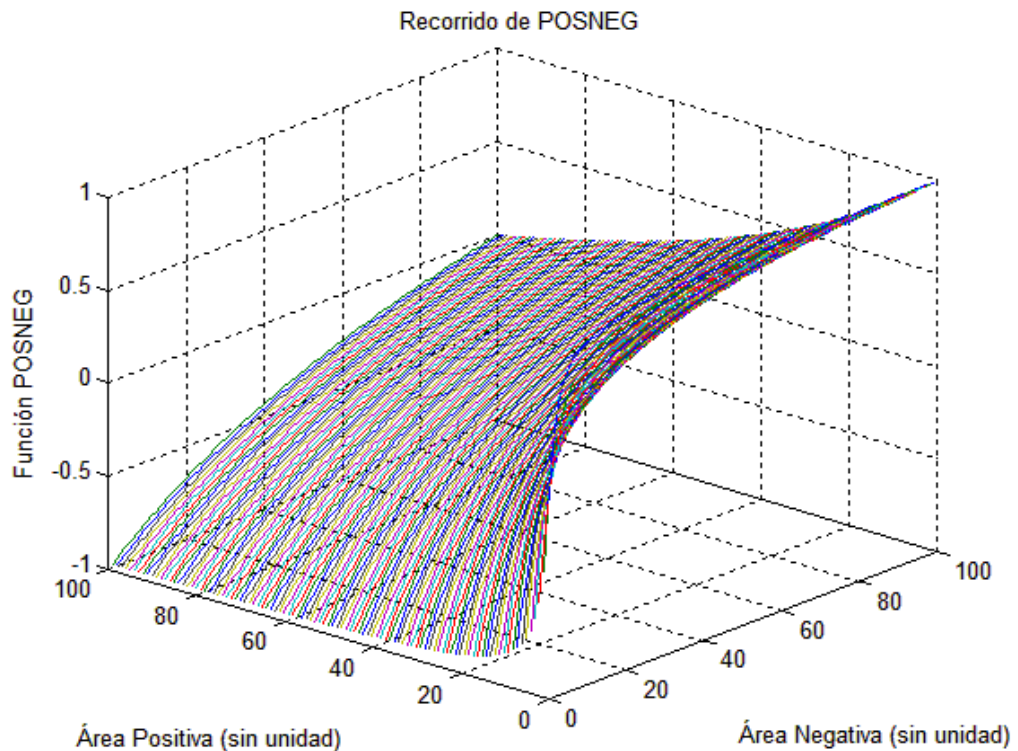


Figura 3: Recorrido de la función POSNEG. Se aprecia cómo los dos bordes más proximales de la curva son asíntotas.

Las características utilizadas en el dominio de la frecuencia son las siguientes:

Frecuencia Media: corresponde a la frecuencia promedio de la señal. Se calcula como la sumatoria de las frecuencias multiplicadas por sus potencias, y divididas por la potencia total. Matemáticamente corresponde a

$$MeanFreq = \frac{\sum_{j=1}^M f_j P_j}{\sum_{j=1}^M P_j}, \quad Ec. 14$$

donde M es la cantidad de frecuencias, $j=1,2,3,\dots,M$, f_j es la frecuencia en la posición j y P_j es la potencia de la frecuencia en la posición j .

Frecuencia Mediana: corresponde a la frecuencia cuya potencia divide al espectro de frecuencia a la mitad. Matemáticamente se define como

$$MedianFreq = \frac{1}{2} \sum_{j=1}^M P_j, \quad Ec. 15$$

donde M es la cantidad de frecuencias, $j=1,2,3,\dots,M$ y P_j es la potencia de la frecuencia en la posición j .

Frecuencia Peak: corresponde a la frecuencia cuya potencia es la máxima del espectro. Matemáticamente

$$PeakFreq = \max(P_j), \quad Ec. 16$$

donde $j=1,2,3,\dots,M$ y P_j es la potencia de la frecuencia en la posición j .

4. Metodología

El presente trabajo se llevó a cabo en seis etapas. En primer lugar, se seleccionaron las señales a utilizar: los complejos QRS y las bases wavelet. La sección 4.1 describe los procesos necesarios para obtener los complejos QRS y las bases wavelet. En segundo lugar, se seleccionaron las características que conformarán el descriptor, detallados en la sección 4.2. En la tercera etapa, se obtuvieron los descriptores de cada complejo QRS y base wavelet y se generó una matriz. Esto es detallado en la sección 4.3. En la cuarta etapa, se analizó la matriz antes nombrada empleando PCA. Los componentes resultantes fueron ordenados de mayor a menor respecto a su aporte a la varianza total, y se seleccionaron los primeros para crear un sub espacio de características. La aplicación de PCA, así como la generación del sub espacio se describe en la sección 4.4. En la quinta etapa, se identificaron los descriptores en este sub espacio recién creado y se midieron las distancias euclidianas entre los descriptores de bases wavelet y los de complejos QRS. La descripción del proceso se encuentra en la sección 4.5. Por último, las bases wavelet asociadas a las menores distancias registradas fueron utilizadas para aplicar *denoising* a

un registro EKG mediante la transformada de wavelet. Este proceso de validación se describe en la sección 4.6.

4.1. Señales a analizar

Se utilizarán dos tipos de señales: la señal biológica a analizar (los complejos QRS) y la señal que es objeto de estudio (la base wavelet). De forma específica, se busca trabajar con la función de escala madre de la base wavelet.

4.1.1. Complejos QRS

Los complejos QRS fueron extraídos de la base de datos “*The QT Database*” (Laguna 1997), obtenida desde el depósito Physionet (Goldberger 2000) (accesible desde el sitio <http://www.physionet.org/>). Esta base de datos fue diseñada para la evaluación de algoritmos de detección de ondas P, T y complejos QRS en señales de EKG.

Para este estudio, se utilizaron dos conjuntos de registros EKG de sujetos sanos, *i.e.* sin patologías cardíacas: un grupo de señales del Hospital Beth Israel de Nueva York (BIH) y de la Sociedad Europea de Cardiología (ESC). Se tomaron las señales de EKG y se separaron complejos QRS mediante marcadores en archivos facilitados por la plataforma ATM de Physionet. Esto, para asegurar que los descriptores extraigan información del núcleo de la señal de EKG y no del conjunto. Se tomaron en total 488 complejos QRS de los registros BIH (que llamaremos BIH-QRS) y 809 complejos de los registros ESC (ESC-QRS). Una vez obtenidos, se identificaron los *peaks* R en base a los marcadores de la base de datos y se alinearon.

4.1.2. Bases Wavelet

Las bases wavelet se pueden clasificar de distintas formas, y de acuerdo a distintos criterios, como si contienen números reales o complejos, o por la longitud de sus filtros (Daubechies 2008). La plataforma MatLab, en su *toolbox* para procesar y aplicar wavelets, las clasifica en 5 grupos: ortogonales, biortogonales, wavelets con función de escala, wavelets sin función de escala y wavelets complejas sin función de escala. Se buscaron las bases wavelet que formaron parte del estudio en publicaciones y trabajos académicos. De todas las bases wavelet existentes, se emplearon aquellas pertenecientes a los tres primeros tipos, y que se encuentran implementadas en MatLab o cuya implementación fuese posible. Las bases resultantes de este proceso son enumeradas en las tablas 1, 2 y 3, donde se encuentran incluidas las bases óptimas: Haar, Db3, Db8, Db44 y Sym3.

Las funciones de escala madre de cada base wavelet fueron obtenidas desde MatLab mediante la función *wavefun*. Ésta permite generar la función de escala y la función de wavelet utilizando un método iterativo, controlado por un parámetro llamado *iter*: a mayor valor de *iter*, mayor es la frecuencia de

muestreo empleada para generar las funciones, y mayor es la cantidad de puntos resultantes. Para estandarizar la obtención de las funciones de escala, se buscó el valor de *iter* por defecto, que equivale a una frecuencia de muestreo igual a 256 Hz.

Bases Wavelet Ortogonales									
Haar		Daubechies				Symmlet			
1	Haar	2	Db2	14	Db24	26	Sym1	35	Sym14
		3	Db3	15	Db27	27	Sym2	36	Sym16
		4	Db4	16	Db28	28	Sym3	37	Sym19
		5	Db6	17	Db30	29	Sym5	38	Sym20
		6	Db8	18	Db33	30	Sym7	39	Sym22
		7	Db9	19	Db34	31	Sym8	40	Sym25
		8	Db12	20	Db36	32	Sym10	41	Sym26
		9	Db15	21	Db39	33	Sym11	42	Sym27
		10	Db17	22	Db41	34	Sym13	43	Sym30
		11	Db18	23	Db42				
		12	Db20	24	Db44				
		13	Db21	25	Db45				

Tabla 1: Bases wavelet ortogonales que son parte del estudio.

4.2. Selección de características

Utilizando como base la lista de parámetros de Phinyomark, se crearon los descriptores con las características generales descritas anteriormente. Éstas corresponden a características de dominio temporal, incluyendo además algunas del dominio de la frecuencia para estudiar su aporte en información.

Bases Wavelet Ortogonales					
Coiflet		Battle-Lemarie		Meyer	
44	Coif1	49	Lem1	52	Meyr
45	Coif2	50	Lem2	53	Dmey
46	Coif3	51	Lem3		
47	Coif4				
48	Coif5				

Tabla 2: Bases wavelet ortogonales que son parte del estudio.

Las características del dominio temporal son: Promedio, Valor Medio Absoluto, Momentos Centrales (desde el segundo hasta el décimo), ZC1, ZC2, ZC3, Longitud de Forma de Onda, Raíz Cuadrática Media y POSNEG.

Las características del dominio de la frecuencia son: Frecuencia Media, Frecuencia Mediana y Frecuencia Peak. Como elementos previos a tener en cuenta, cada una de ellas requiere conocer el espectro de potencia de la señal. Para esto, utilizaremos su estimación: el cálculo de la densidad espectral de potencia (PSD, por sus siglas en inglés) y su frecuencia de muestreo. La frecuencia de muestreo de los registros EKG es de 250 Hz. Las frecuencias de muestreo de las funciones de escala son casi todas similares, a excepción de la función de la base Meyer: corresponde a 16,063 Hz. La frecuencia de muestreo de todas las demás funciones es $256 \pm 0,214$ Hz. Para corregir su frecuencia, se calculó nuevamente su función de escala con la función *wavefun*, ahora ajustando el valor de *iter* para obtener una frecuencia cercana a 256 Hz. Al alcanzar un valor para *iter*=11, la frecuencia de muestreo de Meyer es de 256,0625 Hz, y la nueva frecuencia de muestreo de todas las funciones de escala es de $256 \pm 0,2123$ Hz. La tabla 4 presenta las frecuencias de muestreo de las funciones de escala.

Bases Wavelet Biortogonales							
Spline				Spline Inverso			
54	Bior1.1	61	Bior3.1	69	Rbio1.1	76	Rbio3.1
55	Bior1.3	62	Bior3.3	70	Rbio1.3	77	Rbio3.3
56	Bior1.5	63	Bior3.5	71	Rbio1.5	78	Rbio3.5
57	Bior2.2	64	Bior3.7	72	Rbio2.2	79	Rbio3.7
58	Bior2.4	65	Bior3.9	73	Rbio2.4	80	Rbio3.9
59	Bior2.6	66	Bior4.4	74	Rbio2.6	81	Rbio4.4
60	Bior2.8	67	Bior5.5	75	Rbio2.8	82	Rbio5.5
		68	Bior6.8			83	Rbio6.8

Tabla 3: Bases wavelet biortogonales que son parte del estudio.

4.3. Matriz de descriptores

Para el cálculo de los descriptores, se creó una función que agrupa a todas las características, llamada *vfeat*. En el anexo 2 se expone el código con las fórmulas utilizadas. Esta función entrega un vector con las características explicadas anteriormente. El próximo paso corresponde a la aplicación de PCA, y la función de MatLab encargada de esto requiere como dato de entrada una matriz con todos los descriptores a utilizar. En este caso, los descriptores de los complejos QRS y de las bases wavelet. La matriz será de dimensiones $n \times p$, donde n es la cantidad de descriptores y p es la cantidad de características de cada descriptor. En este estudio, la matriz será llamada *desc* y tendrá dos versiones: una por cada grupo de complejos QRS. La matriz creada con los complejos BIH-QRS será de dimensiones 571×20 , compuesta por 83 descriptores de bases wavelet y 488 descriptores de complejos QRS. La matriz creada con los complejos ESC-QRS será de dimensiones 892×20 , compuesta por 83 descriptores de bases wavelet y 809 descriptores de complejos QRS.

4.4. PCA y generación del sub espacio

Se aplicó el Análisis de Componentes Principales (PCA, por sus siglas en inglés) a la matriz de descriptores. PCA es un método matemático para reducción de dimensiones de grandes grupos de datos. El método crea nuevas variables, llamadas componentes principales, por combinación lineal de las variables existentes, haciendo que todos los componentes principales sean ortogonales entre sí. De esta forma se evita la redundancia. El método crea tantos componentes principales como variables originales existan, proyectando la mayor parte de la variabilidad de los datos sobre la menor cantidad de componentes. Así, la reducción de dimensiones se logra tomando en cuenta sólo los primeros componentes, buscando que en conjunto reúnan sobre el 70% o 90% (Jolliffe 2002) de la varianza total.

En MatLab, la función que permite este análisis es *princomp*. Los productos más importantes de la función, y en los que nos basaremos para la investigación, son la matriz *score* y el vector *latent*. *Score* posee las coordenadas de los puntos en el nuevo sub espacio creado. Esta matriz es de idénticas dimensiones a la matriz de entrada de la función. *Latent* contiene las covarianzas de cada componente; es decir, su representación de la varianza total de los datos.

Se aplicó la función *princomp* a ambas matrices *desc*, obteniendo dos matrices *score* y dos vectores *latent*. Para generar el sub espacio de componentes, se estableció que se considerarían las componentes cuyo porcentaje de varianza represente al menos el 80% respecto de la varianza total. Si aquel porcentaje fuese alcanzado con una o dos componentes, se considerarían las tres primeras para desarrollar la investigación.

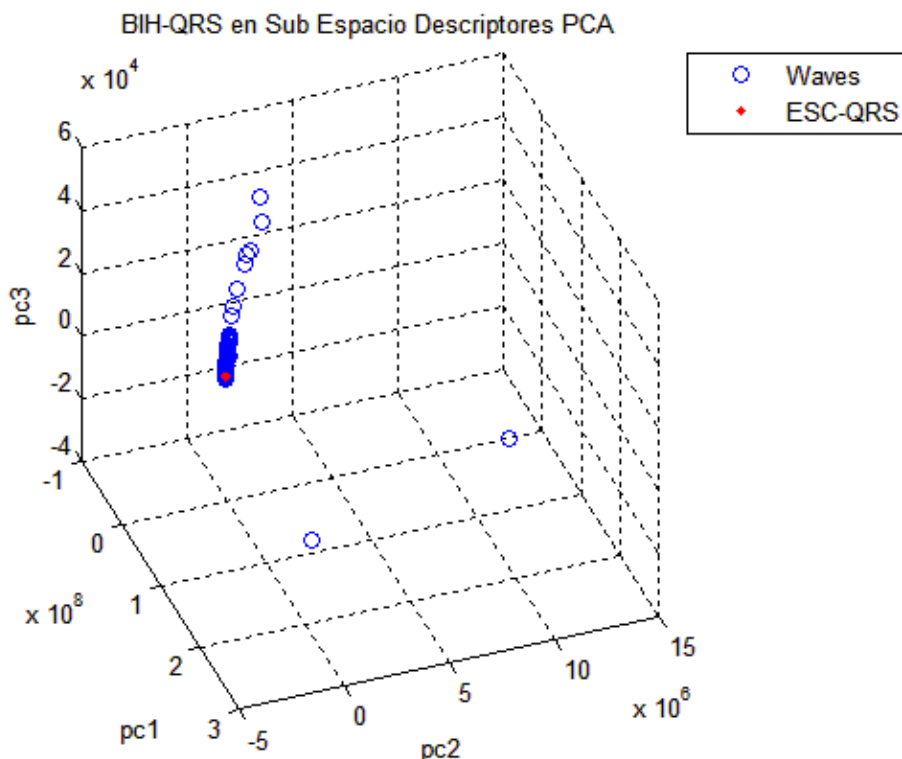


Figura 4: Sub espacio BIH. Se aprecia cómo los complejos QRS quedan por debajo del grupo de bases wavelet.

4.5. Medición de distancias

Una vez creado el sub espacio de características, estaremos frente a dos escenarios: los descriptores de los complejos BIH-QRS y de las bases wavelet por un lado, y por otro, los descriptores de los complejos ESC-QRS y de las bases wavelet. Se busca identificar la menor distancia entre los descriptores extraídos de las bases wavelet y los descriptores de los complejos QRS. Cada descriptor de base wavelet se trabajará de forma individual, mientras que los de complejos QRS deben ser considerados como un solo punto. Por esto, se calculó el centro del conjunto BIH-QRS promediando sus puntos; de la misma forma se calculó el centro del conjunto ESC-QRS. En base a estos puntos se medirán las distancias euclidianas, buscando identificar el descriptor de la base wavelet que presente la menor distancia hacia el centro del conjunto de complejos QRS.

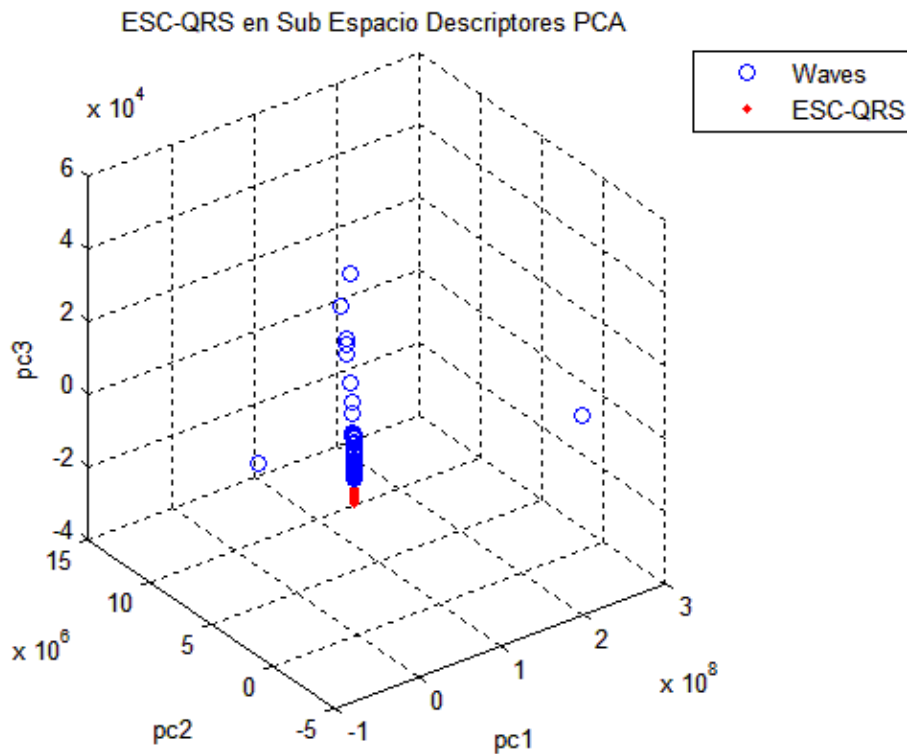


Figura 5: Sub espacio ESC. Se aprecia en este sub espacio que los complejos QRS quedan al descubierto.

4.6. Validación de bases wavelet escogidas

La última etapa corresponde a aplicar la transformada de wavelet sobre una señal utilizando las bases wavelet identificadas, buscando comprobar si hay un cambio en los resultados al utilizar una base u otra. Para esto, se descompondrá un registro EKG con la transformada de wavelet, empleando ambas bases que presentaron la menor distancia. Además, se emplearán las bases identificadas como óptimas por otros trabajos: Haar, Db3, Db8, Db44 y Sym3. En total, se utilizarán 7 bases wavelet.

Se analizarán los coeficientes de la descomposición wavelet que generen cada una de las distintas bases. Se busca que la base que esté mejor adaptada genere la menor cantidad de coeficientes, y que estos sean de valor significativamente alto. Para corroborarlo, se utilizarán 5 ciclos EKG aislados (extraídos de "The QT Database", desde registros no utilizados previamente en el estudio), a los cuales se les aplicará *denoising* empleando cada una de las bases wavelet en cuestión. Enseguida, se aplicará la transformada inversa para reconstruir el registro EKG, se calculará la raíz cuadrada media del error (RMSE) para cada reconstrucción y se promediarán. Al final, se comparará el valor promedio de RMSE de cada una de las bases wavelet analizadas.

Se aplica la transformada de wavelet a un nivel tal que permita la descomposición total o muy cercana. Singh y Tiwari llegan al nivel 10 de descomposición, con señales de 512 puntos de largo. Las señales que se utilizarán en el presente estudio alcanzan en promedio 873 puntos, por lo que se utilizarán los mismos 10 niveles de descomposición.

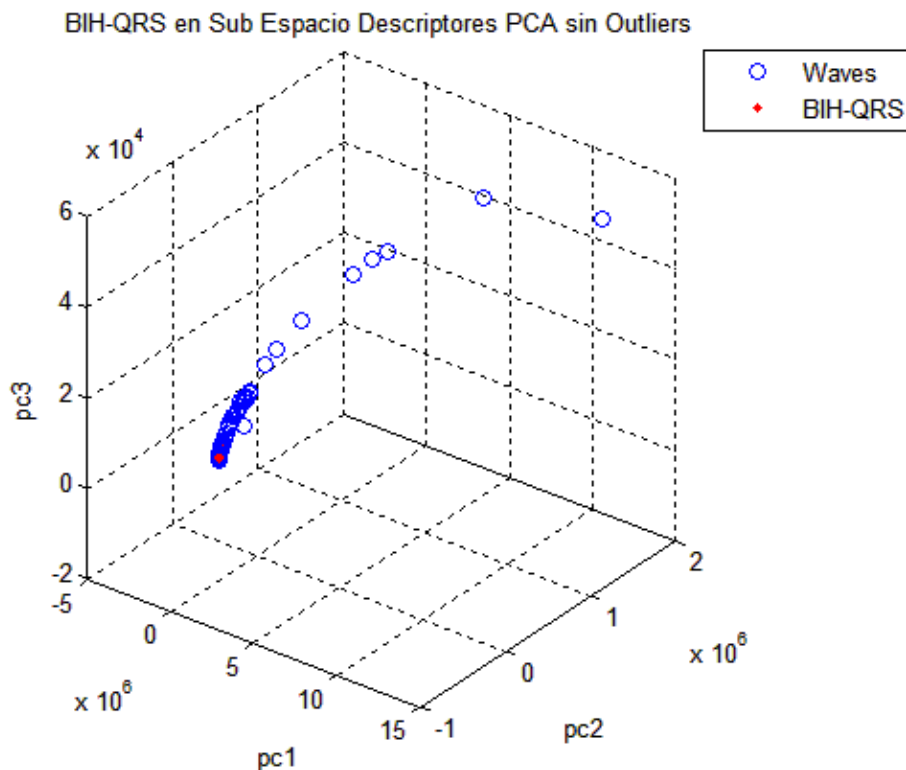


Figura 6: Sub espacio BIH, en ausencia de outliers.

5. Resultados

Se aplicó PCA a los dos bloques de complejos QRS por separado (BIH-QRS y ESC-QRS). Luego, se identificaron las dimensiones de un sub espacio de características usando la información del vector *latent* y de la matriz *score*, generando el sub espacio. Sus resultados se describen en la sección 5.1. Después, se midieron las distancias entre complejos QRS y entre bases wavelet, descritas en la sección 5.2. Por último, las bases que presentaron la menor distancia hacia los complejos QRS fueron utilizadas para aplicar *denoising* con la transformada de wavelet a registros EKG. La descripción de los resultados se encuentra en la sección 5.3.

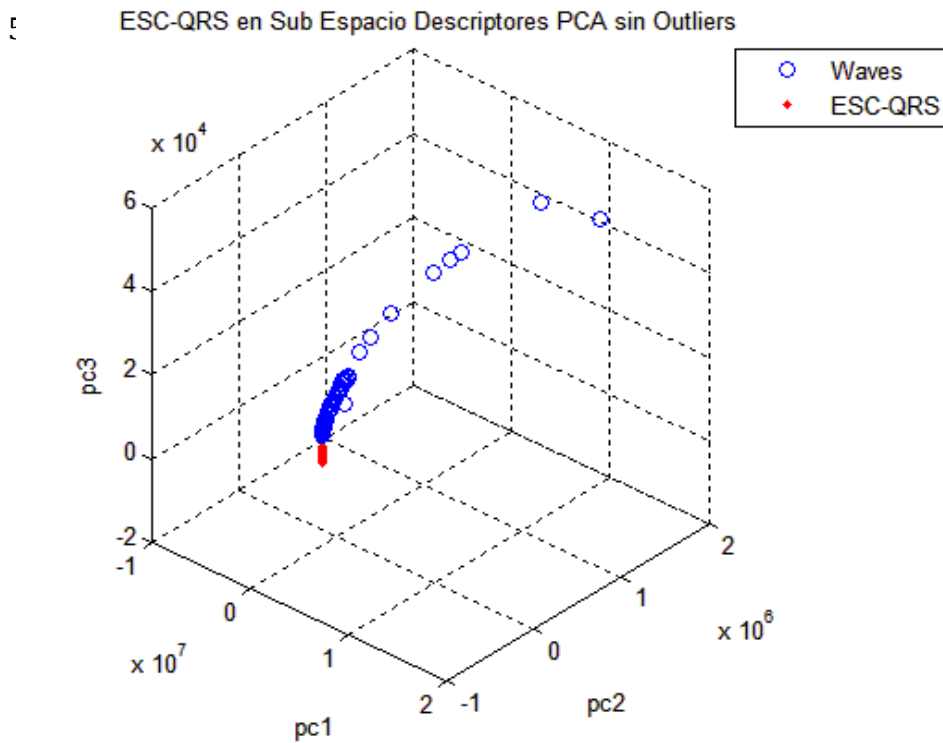


Figura 7: Sub espacio ESC, en ausencia de outliers.

5.1. Generación de sub espacios BIH y ESC

Las covarianzas en el vector *latent* extraído del conjunto BIH-QRS muestran que el primer componente describe prácticamente en su totalidad la distribución de los vectores, representando un 99,9% de la información. Por esto, para crear el sub espacio de características, se consideraron las tres primeras componentes. Éste será el sub espacio BIH. De la misma manera, las covarianzas en el vector *latent* extraído del conjunto ESC-QRS muestran que el primer componente representa un 99,9%, por lo que también se utilizarán las tres primeras componentes. Éste será el sub espacio ESC. La tabla 4 presenta los valores de las covarianzas de ambos grupos.

Vector LATENT					
	BIH	ESC		BIH	ESC
PC1	4,86642E+14	3,12328E+14	PC11	0,145900912	0,803368861
PC2	2,8223E+11	1,81073E+11	PC12	0,02051492	0,40782371
PC3	23943225,74	18737394,98	PC13	0,01370297	0,232135869
PC4	2190809,01	6299680,354	PC14	0,007617686	0,017325517
PC5	197654,439	134851,2342	PC15	0,006679451	0,006109734
PC6	4160,635989	4149,345376	PC16	0,00300035	0,005334855
PC7	1786,47229	2679,474529	PC17	0,001067688	0,002353095
PC8	137,2688572	172,3938603	PC18	0,000334566	0,000940128
PC9	2,194708877	154,9447421	PC19	5,10743E-05	2,91931E-05
PC10	0,569043893	7,628853757	PC20	8,06888E-06	1,04489E-06

Tabla 4: Valores de
vectores LATENT.
Contienen las
varianzas parciales de
cada componente.

5.2. Distancias medidas en sub espacios BIH y ESC

Se calcularon las distancias en cada sub espacio entre el promedio de los descriptores de los complejos QRS y cada uno de los descriptores de las bases wavelet. Para el sub espacio BIH, la menor distancia hasta el centro de los descriptores QRS corresponde a Db9, mientras que para el sub espacio ESC, la menor distancia corresponde a Db4. En los anexos 3 y 4 se exponen todas las distancias.

En las figuras 4 y 5 se puede ver el resultado de graficar las matrices score de ambos grupos. Es decir, se aprecia el sub espacio BIH y QRS, con los descriptores de los complejos QRS y de las bases wavelet. Tanto en las tablas de distancias como en las figuras de los sub espacios, se pueden apreciar puntos muy lejanos, u *outliers*. Si bien son parte de los datos, sus valores pueden alterar el análisis de la mayoría de ellos. Por lo tanto, se considerarán de forma aparte. La figura 6 y 7 muestra los dos sub espacios sin *outliers* en BIH y ESC.

5.3. Validación del proceso de selección de bases wavelet

Las bases seleccionadas por el proceso anterior más las bases consideradas óptimas por otros trabajos son utilizadas como validación. Las bases seleccionadas de forma específica son: Db4 y Db9; y las bases consideradas óptimas son: Haar, Db3, Db8, Db44 y Sym3. En los anexos 5, 6, 7 y 8 se aprecian las funciones de escala madre de todas las bases nombradas.

Las bases se probaron aplicando *denoising* con la transformada de wavelet hasta 10 niveles, a 5 ciclos EKG aislados. Como umbral para eliminar coeficientes, se empleó el umbral universal de Donoho y Johnston. Después, se aplicó la transformada inversa, y al resultado se le calculó el RMSE. Como se utilizaron 7 bases wavelet, se tendrán 35 valores de RMSE. Estos valores se promediaron por base, obteniendo un RMSE promedio para cada una de ellas. La tabla 5 contiene los resultados totales y sus promedios.

6. Discusión

De acuerdo al vector *latent* del primer y segundo sub espacio, la primera característica representa el 99,9% de la varianza total. Esto quiere decir que bastaría con utilizar el primer componente para representar a todos los datos; no obstante, esta elección depende tanto de las características que conforman el descriptor, como de las señales a analizar. Es por esto, que esta elección puede ser modificada. De acuerdo a la literatura, hay varios caminos por los que definir una cantidad de componentes a escoger. Jolliffe (Jolliffe 2002) separa las metodologías en tres grandes grupos: las formas más simples – entre las que se encuentran la utilizada en este estudio – que básicamente se basa en métodos considerados más subjetivos, pero que en la práctica funcionan. El segundo grupo está basado en pruebas de hipótesis. Según el autor, en la práctica sobreestiman la cantidad de variables necesaria. El tercer grupo está basado en un número de reglas estadísticas. Estos métodos están dirigidos para situaciones en las que la toma de decisión no puede ser ayudada por métodos visuales, como gráficos entre dos componentes, o porcentaje de representatividad respecto a la varianza total. Por lo general, son computacionalmente complejos. En el presente estudio, la decisión del método estudiado responde a la baja cantidad de variables utilizadas (20 características) respecto a estudios donde se utilizan 50 características (Phinyomark 2013). Para este caso, se recomienda utilizar como mínimo 3 dimensiones, pues se adapta a las mejoras que se puedan hacer al descriptor, como también se adapta a las distintas señales que puedan ser analizadas mediante el método presentado.

RMSE							
EKG	Db4	Db9	Haar	Db3	Db8	Db44	Sym3
1	29,2812	32,3436	37,5055	26,7812	24,0750	45,0328	26,7812
2	26,1896	28,5238	29,3198	26,2953	31,7195	45,8215	26,2953
3	31,4359	33,4614	37,5684	32,5279	32,7840	49,2817	32,5279
4	23,5272	33,6371	33,2006	23,3007	24,6673	36,5543	23,3007
5	28,9509	30,5091	34,2065	25,0031	32,4557	47,8330	25,0031
Promedios RMSE							
-	27,8770	31,6950	34,3602	26,7816	29,1403	44,9047	26,7816

Respecto a las distancias medidas, las menores distancias correspondieron a Db4 (BIH-QRS) y Db9 (ESC-QRS). En el anexo 7 se aprecian las funciones de escala madre de cada base, y en el anexo 8, un ejemplo de los complejos QRS empleados en cada caso. Se aprecia una gran similitud, no sólo en cuanto a la presencia de la campana central, sino también en la estrechez de ésta.

Respecto a las bases consideradas óptimas por trabajos anteriores, sus distancias a los complejos QRS en comparación con las menores distancias, se encuentran destacadas en los anexos 3 y 4. De las cinco consideradas, Haar tuvo el peor desempeño, siendo uno de los *outliers* en todo momento. La base Db44 se mantuvo lejos de los complejos QRS, quedando en el lugar N°69 en ambos sub espacios. De acuerdo al anexo 6, su función de escala madre presenta gran oscilación, factor que se estima fue el causante de su distancia. La base Db8 en el sub espacio ESC estuvo en el lugar N°25; no obstante, en el sub espacio BIH estuvo en el lugar N°8. Su función de escala madre presenta 2 cúspides positivas y dos cúspides negativas, donde la primera positiva es la predominante (anexo 6). Esta característica podría ser la responsable de que se ubicara cerca en BIH, pero lejos en ESC, donde la campana es mucho más ancha. Las bases Db3 y Sym3 actuaron de forma coordinada: en el espacio BIH se ubicaron lejos de los complejos QRS, en el puesto 30 y 29, respectivamente; mientras que en el espacio ESC se ubicaron muy cerca, en los puestos 4 y 3 de forma respectiva. Ambas funciones de escala madre son idénticas y presentan una cúspide muy perfilada con una curva similar a un *overshoot*, también muy perfilada. Además, presenta una relación de áreas más positiva. Estos elementos harían que ambas funciones de escala madre se ubiquen más cerca de la señal de ESC (más ancha) que de la señal BIH (más angosta). En general, se cree que la presencia de oscilaciones, de cúspides y la forma de éstas influyó en la distancia de las bases wavelet a los complejos QRS.

Las bases que se identificaron como *outliers* son Haar, Sym1, Bior1.1, Rbio1.1 y Dmey. En el anexo 9 se pueden apreciar sus funciones de escala madre. La función de escala madre de las 4 primeras bases nombradas son idénticas, y se justifica su posición como *outlier* por ser tan distinta a la señal objetivo. Por otra parte, Dmey tiene una forma de onda similar a una función *sync*, pero con una oscilación centrada muy alta en frecuencia y muy corta en tiempo. Los complejos QRS poseen una cúspide importante; sin embargo, no es tan fina y angosta como lo es Dmey. Por esto, se cree que se justifica su posición como *outlier*.

El último paso correspondió a la validación. Se aplicó *denoising* a cada uno de los registros EKG con cada base wavelet en cuestión, y luego se calculó el

Tabla 5: Valores de RMSE, y sus correspondientes promedios por base wavelet.

valor RMSE de cada uno de los resultados (respecto a la señal original). Los promedios de RMSE por cada base son los valores a comparar. La tabla 5 presenta los valores detallados por base y por registro EKG, entregando también sus promedios. Respecto a estos, Db3 y Sym3 tuvieron el mismo valor RMSE. Recordemos que sus funciones de escala madre son exactamente iguales, por lo que se esperaba que los resultados sean idénticos. No obstante, estas dos bases wavelet obtuvieron el menor RMSE, por lo que en teoría generaron el mejor resultado. Luego, le sigue en valor Db4, luego Db8, después Db9 y muy cerca, Haar. El valor más alejado del grupo lo tuvo Db44. En los anexos 10, 11 y 12 se visualiza una reconstrucción con cada base wavelet analizada en la etapa de validación. Haciendo un análisis visual, se confirma que Db3/Sym3, Db4 y Db8 generaron mejores resultados, desde el punto de vista de reconstruir la señal EKG. Se puede notar que en la reconstrucción de Db3/Sym3 los bordes de las ondas P, R y T son muy perfilados, mientras que las de Db4 y Db8 son mucho más suavizados. La reconstrucción de Haar, aunque posee un valor RMSE dentro de los esperado como base “óptima”, presenta una reconstrucción muy cuadrada, que visualmente no concuerda con un valor bajo de RMSE. Por último, por parte de Db44, la señal presenta oscilación excesiva respecto a la señal original. Esto estaría dado en gran parte por la forma de su función de escala madre.

Tabla 6:
Comparación de
resultados
parciales.

Base	Distancias sub espacio BIH	Distancias sub espacio ESC	RMSE Promedio
Db4	32.533,2863	13.438,15241	27,87696
Db9	5.888,984404	40.515,61003	31,69501
Haar	263.021.482,4	263.054.831,3	34,36017
Db3	33.655,51583	13.537,69738	26,78163
Db8	13.534,7933	32.127,16338	29,1403
Db44	1.009.400,527	1.039.186,536	44,90466
Sym3	33.655,51583	13.537,69738	26,78163

Para entregar un resultado final, se consideran tres resultados ya analizados. La tabla 6 presenta los valores obtenidos por el método de selección en ambos sub espacios, y el valor RMSE promedio proveniente del proceso de validación.

Las bases consideradas en la descripción del resultado final son aquellas indicadas como las más adecuadas para analizar señales EKG y aquellas que obtuvieron la menor distancia en cada grupo. Las bases que peor se desempeñaron fueron Haar y Db44. Haar respecto a distancias fue parte del grupo de los outliers, por tanto, la distancia a la que se ubica en cada sub espacio es la más grande. Su valor RMSE promedio es el segundo más grande, por lo que se puede descartar como una base bien adecuada, o similar. La base Db44 presentó distancias muy parecidas en ambos estudios, y corresponden a las más grandes después de Haar. Además, su valor RMSE promedio es el mayor de todos. Estos antecedentes descartan también a Db44 como una base adecuada para la aplicación en señales EKG. La base Db9, a pesar de poseer la menor distancia en el sub espacio BIH, en el sub espacio ESC su distancia fue

mucho mayor. Esta diferencia está dada, en gran parte por la diferencia de orígenes y de muestreo de los complejos QRS. Una propuesta para evitar que este factor influya es emplear grupos de complejos QRS homogéneos. Es decir, que si existe más de un origen de los complejos QRS, se mezclen de forma homogénea las señales. El valor de RMSE promedio confirma su descarte como la base más adecuada para registros EKG.

Las bases wavelet restantes son las que mejor se desempeñaron. Además, tienen valores muy similares entre sí. Respecto a las distancias, en el espacio BIH la menor de éstas (sin considerar a Db9) corresponde a Db8, mientras que en el espacio ESC, la menor distancia está asociada a Db4. Respecto al valor RMSE promedio, la diferencia entre una y otra base es muy pequeña (tomando como elemento de comparación el valor de Db4), siendo la base Db4 la poseedora del menor valor. Db3 y Sym3 poseen funciones de escala madre idénticas, que si bien sus distancias son similares a las de Db4 y Db8, en sus RMSE promedio obtuvieron los menores valores. Estos antecedentes nos dejan en la debacle de qué base es la más adecuada. Se podría utilizar un método numérico para dirimir entre las bases. Un ejemplo podría ser añadir un peso w a cada resultado, y luego sumar los valores de los resultados en base al aporte indicado por su peso. El problema se presenta en cómo distribuir esos pesos, pues se debe establecer por importancia dentro de la metodología y por impacto de su valor absoluto en la sumatoria. No es lo mismo un porcentaje del valor de la distancia presentada en la tabla 6 que el mismo porcentaje del valor de RMSE promedio. En este trabajo, considerando que el método de selección (resultado principal) identificó a dos bases, y que una de éstas se encuentra en este grupo de los mejores resultados, gracias a los valores de la validación (resultado secundario), se escoge a Db4 como la señal más similar a una señal EKG; es decir, es considerada como “mejor adaptada”.

7. Conclusiones

Si bien este estudio no es concluyente por varias razones que se describirán más adelante, se pueden concluir algunas cosas. La principal conclusión que se puede sacar es que efectivamente existe una relación de similitud entre la base wavelet y la señal. Mientras más similar sea la función de escala madre a la señal que se está analizando, más fácil será encontrar patrones. Esto puede ser de dos formas: que se encuentre la señal que se busca entre el ruido u otras señales, y que se encuentren patrones que no son parte de la señal, quedando expuesto en la descomposición piramidal.

Este hecho abre un campo a desarrollar en este tema. Poniendo un ejemplo: hay una señal de EKG y hay un patrón anómalo que se repite sólo algunas veces en la señal. Se podría aplicar la transformada de wavelet con la base más similar a EKG para localizar ese patrón anómalo como una singularidad, como un elemento extraño en los coeficientes de descomposición. La detección de patrones es un tema bastante avanzado en wavelet (Tang 2009).

La posibilidad de utilizar un vector de características, o descriptor como se le llamó, y utilizarlo para caracterizar una señal es posible, es viable. Lo

importante en este trabajo es que, si bien las características que se emplearon se recolectaron de una sola lista (que es una recopilación en sí (Phinyomark 2012, 2013)), permitió que señales que son similares se agrupen. El trabajo futuro en este ámbito es realizar un estudio sobre qué tipos de características deben ser incluidas y si el número de ellas debe modificarse. En el presente trabajo se emplearon sólo 3 características del dominio de la frecuencia. Existen algunas otras características, como los Coeficientes Cepstrales en la Frecuencia de Mel (MFCC, por sus siglas en inglés). En este grupo, existe una gran variedad de características que se relacionan con el espectro en frecuencia. Trabajan en base a la percepción auditiva humana (Han 2011).

Otro punto importante es la generación del sub espacio de características con PCA. Si bien entrega la posibilidad de reducir dimensiones, es necesario utilizar muchos datos cada vez que se desee crear un nuevo sub espacio. La idea inicial en este trabajo fue lograr crear un descriptor tal que se calcule una vez para cada base wavelet y se almacene como una base de datos. Así, si alguien desea saber qué función de escala madre es similar a su señal, basta con calcular el descriptor de la señal objetivo y compararlo con la base de datos existente de descriptores de bases wavelet. De esa forma, el criterio “a priori” toma sentido.

Por último, por un método de selección se identificaron dos bases wavelet como las más similares a sus conjuntos de señales (Db4 y Db9) y en la etapa de validación se mostraron muy cercanas a ser las más similares. Estos resultados son satisfactorios, dadas las posibilidades de mejora de cada sección de la metodología. No cabe duda que este método es viable, es posible de llevar a cabo. Como se dijo en un principio, esta investigación fue pensada para ser el primer paso, para llevar a cabo una idea y corroborar con datos si la metodología tenía validez. Desde ese punto de vista, los resultados son satisfactorios.

Referencias Bibliográficas

- Bedoya, J.V., Maria, J.M. (2006) *La Transformada Wavelet y su Aplicación en Ingeniería Biomédica*. Programa de Ingeniería Biomédica EIA-CES.
- Burrus, C.S., Gopinath, R., Guo, H. (1998) *Introduction to Wavelet and Wavelet Transforms*. Prentice Hall, New Jersey.
- Chapa, J., Rao, R. (2000). *Algorithms for Designing Wavelets to Match a Specified Signal*. IEEE Transactions on Signal Processing 48 (12).
- Chui, C.K. (1997) *Wavelet: A Mathematical Tool for Signal Processing*. SIAM, Filadelfia.
- Daubechies, I. (1992). *Ten lectures on wavelets* (Vol. 61, pp. 198-202). Philadelphia: Society for industrial and applied mathematics.
- Daubechies, I. (2008) *Orthonormal Bases of Compactly Supported Wavelets*. AT&T Laboratories.
- De Chazal, P., O'Dwyer, M., Reilly, R. (2004) *Automatic Classification of Heartbeats Using ECG Morphology and Heartbeat Interval Features*. IEEE Transactions on Biomedical Engineering 51 (7).1196-1206.
- De Sobral Cintra, R.J., Tchervensky, I.V., Dimitrov, V.S., Mintchev, M.P. (2004) *Optimal Wavelets for Electrogastrography*. The 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society; San Francisco, CA, USA.
- Deng, J., Simmermacher, C., Cranefield, S. (2008) *A Study on Feature Analysis for Musical Instrument Classification*. IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics. 38 (2), 429-438.
- Faundez, P., Fuentes, A. (2001) *Procesamiento Digital de Señales Acústicas utilizando Wavelets*. Instituto de Matemáticas UACH, Chile.
- Goldberger A.L., Amaral L.A.N., Glass L., Hausdorff J.M., Ivanov P.C.H., Mark R.G., Mietus J.E., Moody G.B., Peng C-K., Stanley H.E. (2000). PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. *Circulation*101(23):e215-e220 [Circulation Electronic Pages; <http://circ.ahajournals.org/cgi/content/full/101/23/e215>]; (June 13).
- Galindo, C. (2009) General [Power Point]. Ingeniería Civil Biomédica, Universidad de Valparaíso, Valparaíso, Chile.
- Haidekker, M.A. (2011) *Advanced Biomedical Image Analysis*. John Wiley & Sons, INC, Décima versión.

- Hamed, M., Salleh, S., Noor, A., Swee, T., Afizam, I. (2012). *Comparison of Different Time-Domain Feature Extraction Methods on facial Gestures' EMGs*. Progress in Electromagnetics Research Symposium Proceedings, KL, Malaysia. Marzo 27-30.
- Han, Z., Wang, J., Wang, X., Lun, S. (2011). *Robust Feature Extraction for Speech Recognition Based on Perceptually Motivated MUSIC and CCBC*. Chinese Journal of Electronics 20 (1).
- Huang, G., Zhang, Z., Zhang, D., Zhu, X. (2013). *Spatio-spectral filters for low-density surface electromyographic signal classification*. Medical and Biological Engineering and Computing 51, 547-555.
- Jolliffe, I. (2002) *Principal Component Analysis*. John Wiley & Sons, Ltd.
- Kania, M., Fereniec, M., Maniewski, R. (2007). *Wavelet Denoising for Multi-lead High Resolution ECG Signals*. Measurement Science Review 7 (4), 30-33.
- Laguna P., Mark R.G., Goldberger A.L., Moody G.B. (1997). *A Database for Evaluation of Algorithms for Measurement of QT and Other Waveform Intervals in the ECG*. Computers in Cardiology 24:673-676.
- Mallat, S. (1999). *A Wavelet Tour of Signal Processing, (Wavelet Analysis & Its Applications)*.
- Meyer, Y. (1992) *Wavelets and Operators*. Cambridge University Press, Primera Versión.
- Nagar, B., Hashmi, F., Dhakad, P. (2011) *Comparative Analysis of Fast Wavelet Transform for Image Compression for optimal Image Quality and Higher Compression Ratio*. IJEST, Vol 3 (5), pp 4014-4019.
- Oskoei, M., Hu, H. (2007). *Myoelectric control systems – A survey*. Biomedical Signal Processing and Control 2(4): 275-294.
- Peters, T; Williams, J. (1998) *The Fourier Transform in Biomedical Engineering*. Birkhäuser Boston.
- Phinyomark, A., Phukpattaranont, P., Limsakul, C. (2012). *Feature reduction and selection for EMG signal classification*. Expert Systems with Applications 39: 7420-7431.
- Phinyomark, A., Quaine, F., Charbonnier, S., Serviere, C., Tarpin-Bernard, F., Laurillau, Y. (2013). *EMG feature evaluation for improving myoelectric pattern recognition robustness*. Expert Systems with Applications 40: 4832-4840.
- Rafiee J., Rafiee, M.A., Prause, N., Schoen, M.P. (2011). *Wavelet basis functions in biomedical signal processing*. Expert Systems with Applications (38), 6190:6201.
- Samanta, B., Al-Balushi, K. (2003). *Artificial Neural Network Based Fault Diagnostics of Rolling Element Bearings using Time-Domain Features*. Mechanical Systems and Signal Processing 17(2), 317:328.

Saridis, G., Gootee, T. (1982). *EMG pattern analysis and classification for a prosthetic arm*. IEEE Transactions on Biomedical Engineering BME-29 (6), 403:412.

Sede Bogota Dnsav, Universidad Nacional de Colombia – Sede Bogotá. Recuperado el 10 de Enero de 2011 de http://www.virtual.unal.edu.co/cursos/ciencias/2001004/lecciones_html/cap8/cap8s3.html.

Singh, B.N., Tiwari, A.K. (2006) *Optimal Selection of Wavelet Basis Function applied to ECG Signal Denoising*. Digital Signal Processing (16), pp 275:287.

Strang, G., Nguyen, T. (1996) *Wavelets and Filter Banks*. Wellesley-Cambridge Press, Segunda edición. ISBN 0-9614088-7-1.

Takla, G., Nair, B., Loparo, K. (2006). *Matching a Wavelet to ECG Signal*. Proceedings of the 28th IEEE EMBS Annual International Conference New York City, USA, Aug 30-Sept 3.

Tang, Y. (2009). *Wavelet Theory and Its Application to Pattern Recognition*. Series in Machine Perception Artificial Intelligence 74. Segunda Edición. World Scientific.

Tzanetakis, G., Cook, P. (2002). *Musical Genre Classification of Audio Signals*. IEEE Transactions on Speech and Audio Processing 10 (5), 293:302.

Anexos

Anexo N°1: Código de característica creada para el estudio: POSNEG.

```
function S=posneg(x,opt)

% POSNEG(X,OPT) entrega la relacion de áreas positivas y negativas
% de
% la señal X.
%
% OPT = 1, indica que se realizará de la forma INICIAL.
%           |sum(pos)| { si POS>NEG, 1 <posneg<Inf
%   posneg(x)= ----- { si POS=NEG, posneg = 1
%           |sum(neg)| { si POS<NEG, 0 <posneg< 1
%
% OPT = 2, indica que se hará de la forma FINAL.
%           |sum(pos)|-|sum(neg)| { si POS>NEG, 0 <posneg< 1
%   posneg(x)= ----- { si POS=NEG, posneg = 0
%                   sum(x)      { si POS<NEG, -1 <posneg< 0
%
% OPT = 3, indica que se hará de la forma TESIS.
%           |sum(pos)|-|sum(neg)| { si POS>NEG, 0 <posneg< 1
%   posneg(x)= ----- { si POS=NEG, posneg = 0
%           |sum(pos)|+|sum(neg)| { si POS<NEG, -1 <posneg< 0
%
switch opt
case 1
flag=1;
case 2
flag=2;
case 3
flag=3;
otherwise
flag=1;
end

P=abs(x.*(x>0));
N=abs(x.*(x<0));

if flag==1
S=sum(P)/sum(N);
elseif flag==2
S=(sum(P)-sum(N))/sum(x);
else
S=(sum(P)-sum(N))/(sum(P)+sum(N));
end
```

```

function v=vfeat(sig,varargin)

if numel(varargin)==0
fs=1;
else
fs=varargin{1};
end

[fy,fx]=pwelch(sig,[],[],round(fs));
[~,ind]=max(fy);

aux1=mean(sig); % MM
aux2=var(sig,0,2); % VAR
aux3=skewness(sig); % SK
aux4=kurtosis(sig); % KU
aux5=moment(sig,5)/(std(sig)^5); % TM5
aux6=moment(sig,6)/(std(sig)^6); % TM6
aux7=moment(sig,7)/(std(sig)^7); % TM7
aux8=moment(sig,8)/(std(sig)^8); % TM8
aux9=moment(sig,9)/(std(sig)^9); % TM9
aux10=moment(sig,10)/(std(sig)^10); % TM10
aux11=sum([0 sig].*[sig 0])<0); % ZC1 (phinyomark)
aux12=sum(abs([sig 0]-[0 sig]))/2/numel(sig); % ZC2 (deng)
aux13=0.5*sum(abs([sig 0]>0)-([0 sig]>0))); % ZC3 (tzanetakis)
aux14=sum(abs([sig 0]-[0 sig]))-abs(sig(end)); % WL
aux15=sqrt(sum(sig.*sig)/numel(sig)); % RMS
aux16=(1/numel(sig))*sum(abs(sig)); % MAV
aux17=posneg(sig,3); % AR
aux18=m2freq(sig,fs); % MedFre
aux19=(sum(fx.*fy))/(sum(fy)); % MeanFre
aux20=fx(ind); % PeaFre

v=[aux1 aux2 aux3 aux4 aux5 aux6 aux7 aux8 aux9 aux10 ...
    aux11 aux12 aux13 aux14 aux15 aux16 aux17 aux18 aux19 aux20];

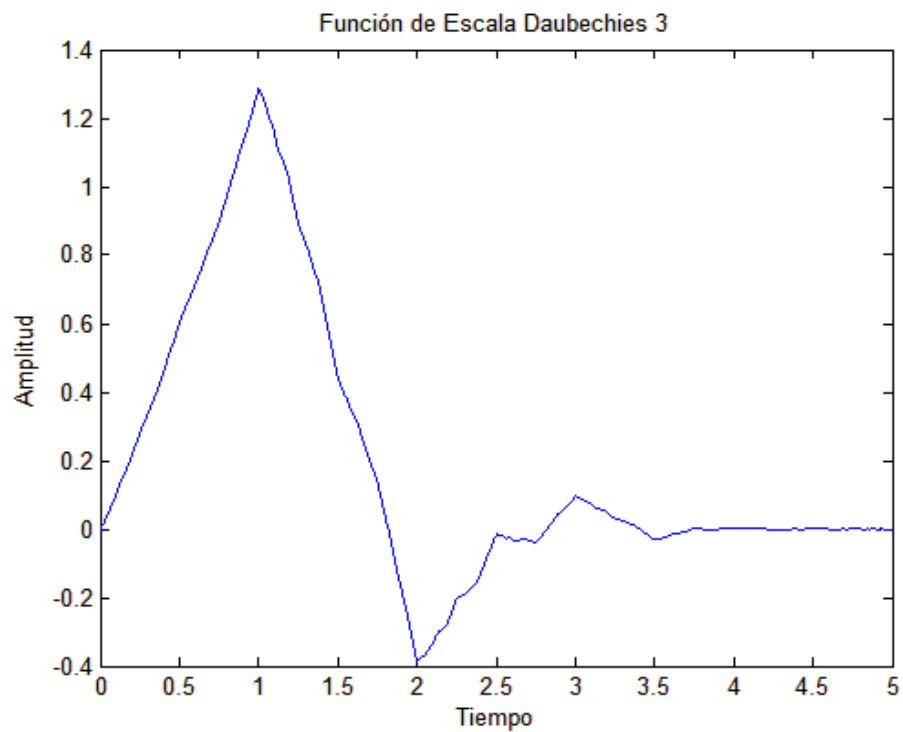
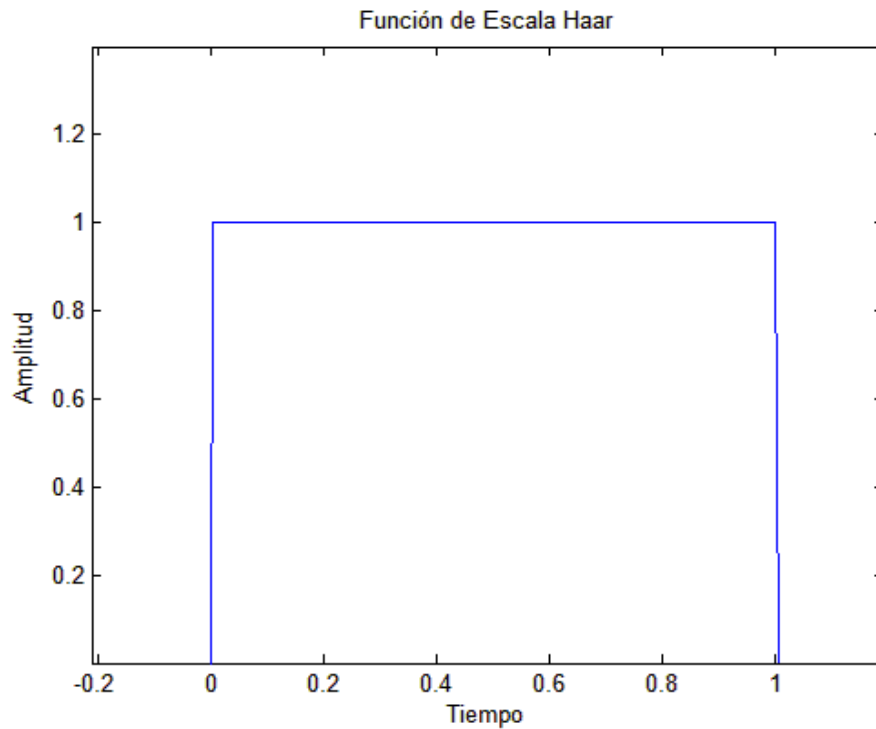
```

Anexo N°3: Tabla de distancias ordenadas de mayor a menor en sub espacio BIH.

N°	Wavelet	Distancia	N°	Wavelet	Distancia
1	haar	263021482,4	43	coif3	84305,24955
2	sym1	263021482,4	44	rbio6.8	67659,08724
3	bior1.1	263021482,4	45	rbio3.9	46737,45949
4	rbio1.1	263021482,4	46	db12	41605,00319
5	dmey	77927182,02	47	rbio3.1	34684,94496
6	bior3.1	17364601,38	48	lem1	34659,40228
7	sym30	9206636,747	49	db2	34523,75815
8	sym27	5790887,783	50	sym2	34523,75815
9	sym26	5260632,458	51	rbio1.3	34297,8963
10	sym25	4554504,308	52	bior1.3	34177,16306
11	sym22	2772788,182	53	rbio2.2	33932,97603
12	sym20	1899219,059	54	db3	33655,51583
13	sym19	1493552,923	55	sym3	33655,51583
14	db45	1051199,882	56	rbio3.3	33632,52166
15	db44	1009400,527	57	db4	32533,2863
16	db42	929964,7573	58	bior2.6	32489,73624
17	db41	890307,7323	59	coif1	30762,51494
18	bior2.2	886349,0695	60	rbio1.5	28942,12349
19	db39	812978,1084	61	sym8	28632,9724
20	sym16	786523,7163	62	bior3.3	28451,76979
21	coif5	708452,4964	63	lem2	28157,93245
22	db36	700562,0931	64	sym5	28153,14814
23	db34	628261,7811	65	bior1.5	27996,14509
24	db33	592969,1968	66	rbio3.5	27279,90015
25	db30	490851,4899	67	db6	26321,88752
26	sym14	455426,7994	68	rbio4.4	26127,21802
27	db28	426120,7452	69	rbio2.4	25529,72915
28	db27	394854,5481	70	bior3.5	24791,16175
29	sym13	324142,3685	71	meyr	19497,09962
30	db24	305792,5646	72	bior3.7	17584,35864
31	coif4	292164,4787	73	rbio5.5	17252,59665
32	db21	224662,7667	74	rbio2.6	17082,96484
33	db20	199559,4158	75	bior4.4	16412,7988
34	db18	152580,0248	76	db8	13534,7933
35	db17	130807,0778	77	bior5.5	11957,7816
36	sym11	110915,4804	78	coif2	9680,561452
37	rbio2.8	99486,01921	79	sym7	9236,0034
38	sym10	98900,09298	80	rbio3.7	8791,228028
39	bior3.9	93870,80511	81	bior2.4	7672,491244
40	bior2.8	93642,38227	82	lem3	7074,295214
41	db15	91087,25734	83	db9	5888,984404
42	bior6.8	87746,85998			

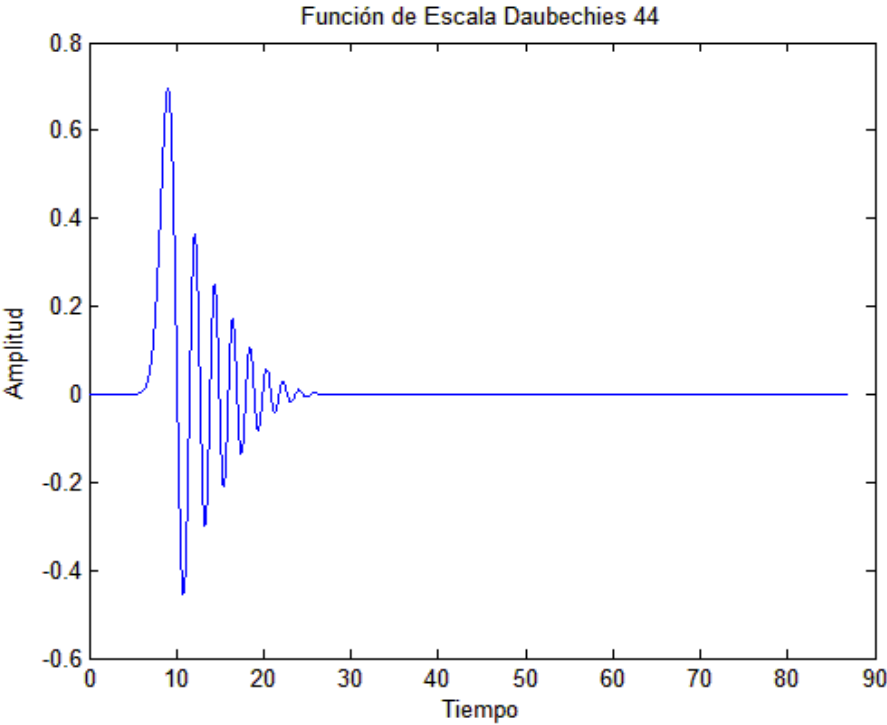
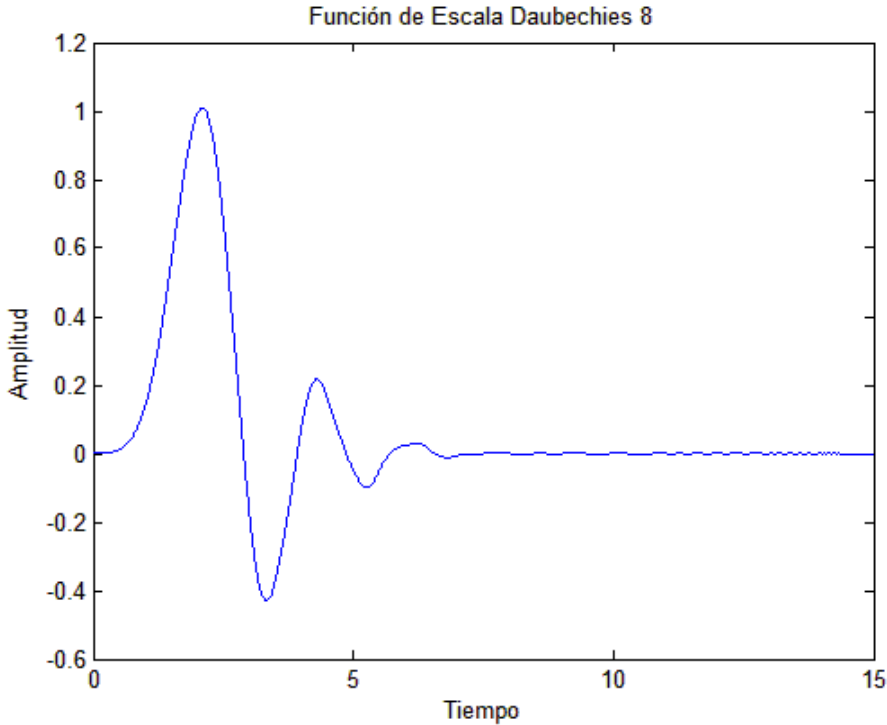
Anexo N°4: Tabla de distancias ordenadas de mayor a menor en sub espacio ESC.

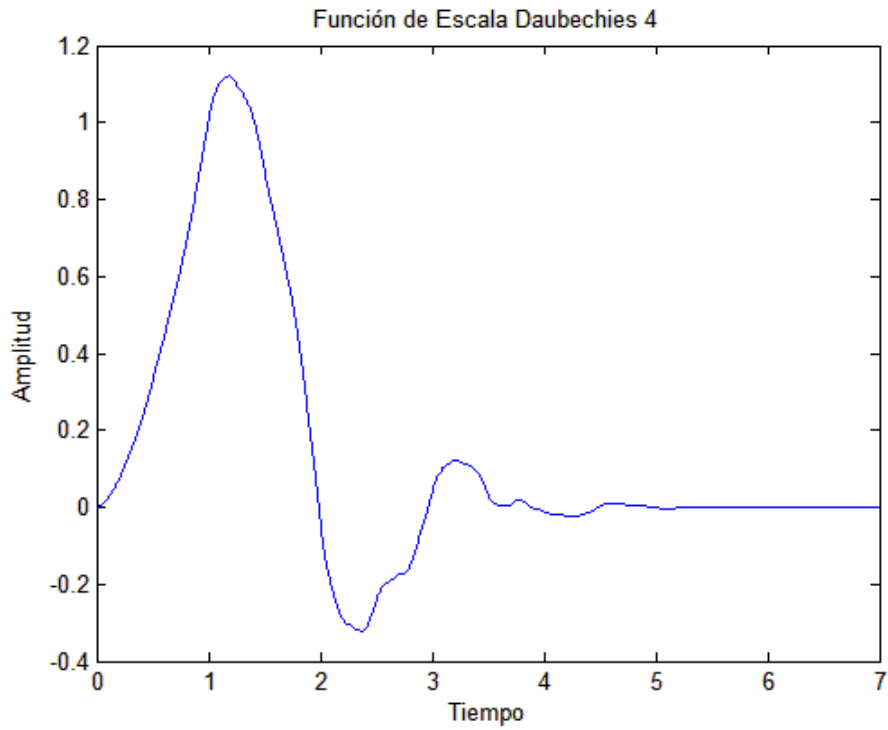
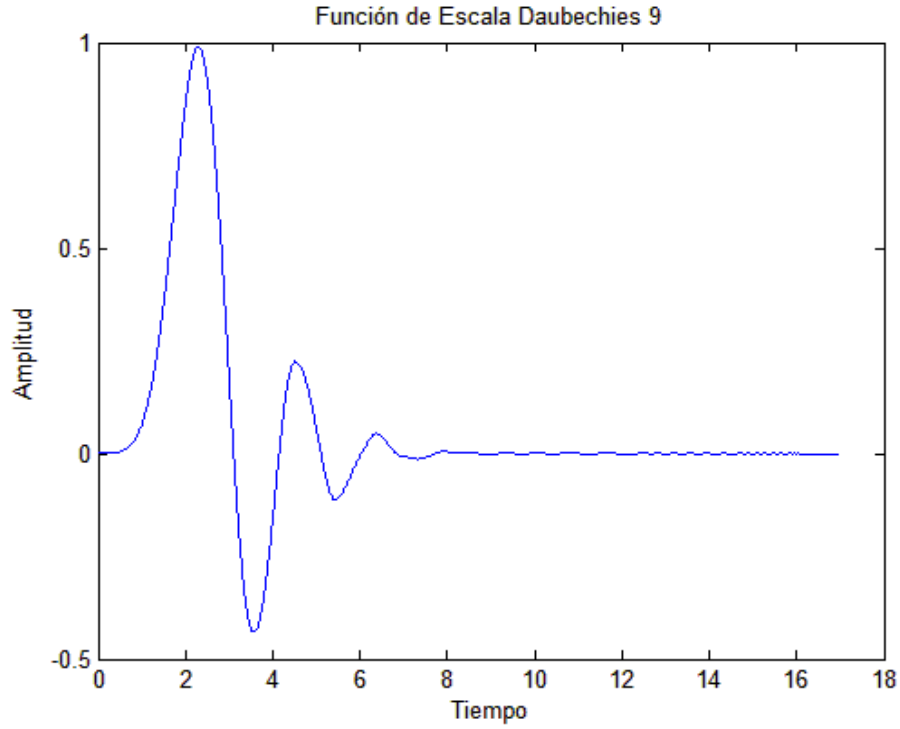
1	haar	263054831,3	43	coif3	118322,0607
2	sym1	263054831,3	44	rbio6.8	101755,8399
3	bior1.1	263054831,3	45	rbio3.9	80943,54815
4	rbio1.1	263054831,3	46	db12	75875,6704
5	dmey	77956463,71	47	bior2.6	66803,18787
6	bior3.1	17403179,57	48	sym8	62973,10343
7	sym30	9244321,242	49	meyr	53914,63771
8	sym27	5827195,499	50	bior3.7	52042,58912
9	sym26	5296685,243	51	rbio2.6	51565,41902
10	sym25	4590116,737	52	bior2.4	42312,73628
11	sym22	2805220,503	53	db9	40515,61003
12	sym20	1928962,522	54	lem3	38745,18708
13	sym19	1521795,174	55	rbio3.7	37213,68794
14	db45	1080853,651	56	sym7	36636,79009
15	db44	1039186,536	57	coif2	36160,36487
16	db42	960011,2264	58	bior5.5	33823,41558
17	db41	920489,8331	59	db8	32127,16338
18	bior2.2	918842,2127	60	bior4.4	29175,01404
19	db39	843430,6792	61	rbio5.5	28349,872
20	sym16	816371,8518	62	bior3.5	20546,16467
21	coif5	738656,6511	63	rbio2.4	20057,60089
22	db36	731436,0534	64	rbio4.4	19444,54932
23	db34	659421,9755	65	db6	19191,42885
24	db33	624274,2073	66	rbio3.5	18360,14838
25	db30	522608,5323	67	bior1.5	17631,54527
26	sym14	486909,0356	68	sym5	17522,96941
27	db28	458175,9386	69	lem2	17510,40209
28	db27	427064,2546	70	bior3.3	17257,90828
29	sym13	356390,1212	71	rbio1.5	16733,08861
30	db24	338469,5492	72	coif1	15152,30955
31	coif4	324618,0324	73	rbio3.1	14424,27982
32	db21	257826,0266	74	lem1	14404,45191
33	db20	232882,0117	75	sym2	14293,92377
34	db18	186230,0367	76	db2	14293,92377
35	db17	164619,854	77	rbio1.3	14104,01124
36	sym11	144742,1329	78	bior1.3	14000,79532
37	rbio2.8	133417,4529	79	rbio2.2	13778,12629
38	sym10	132827,1637	80	db3	13537,69738
39	bior3.9	127837,9908	81	sym3	13537,69738
40	bior2.8	127608,4492	82	rbio3.3	13522,67242
41	db15	125096,7522	83	db4	13438,15241
42	bior6.8	121743,4113			



Anexo N°6: Funciones de escala madre de las funciones consideradas "óptimas": Daubechies 8 y Daubechies 44.

Inspección de un
Criterio y
Metodología de
Selección de
Bases Wavelet
para Aplicación
en señales de
EKG

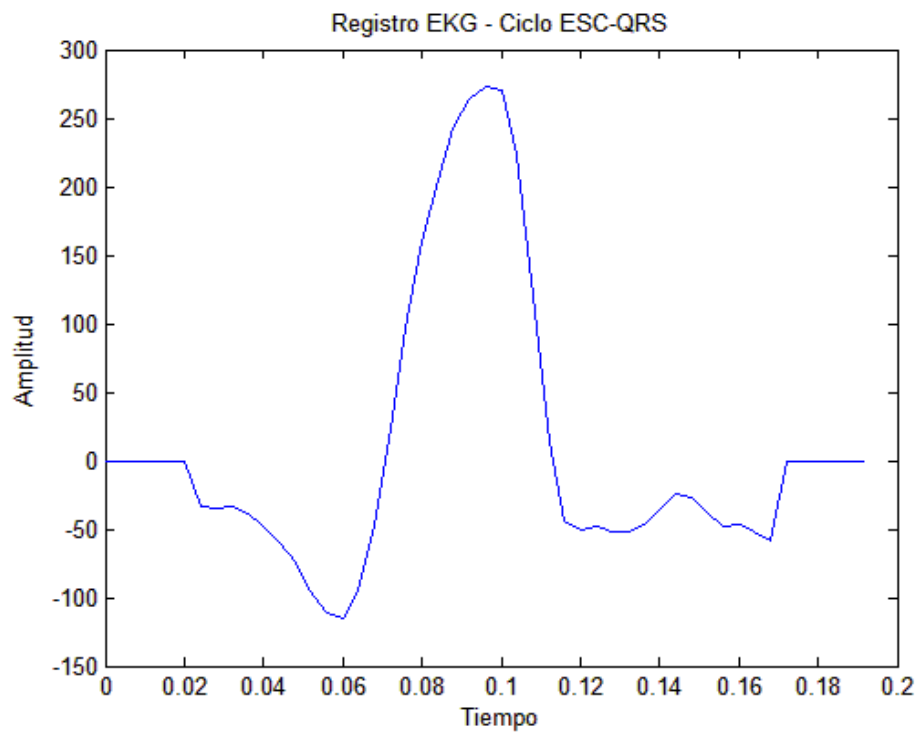
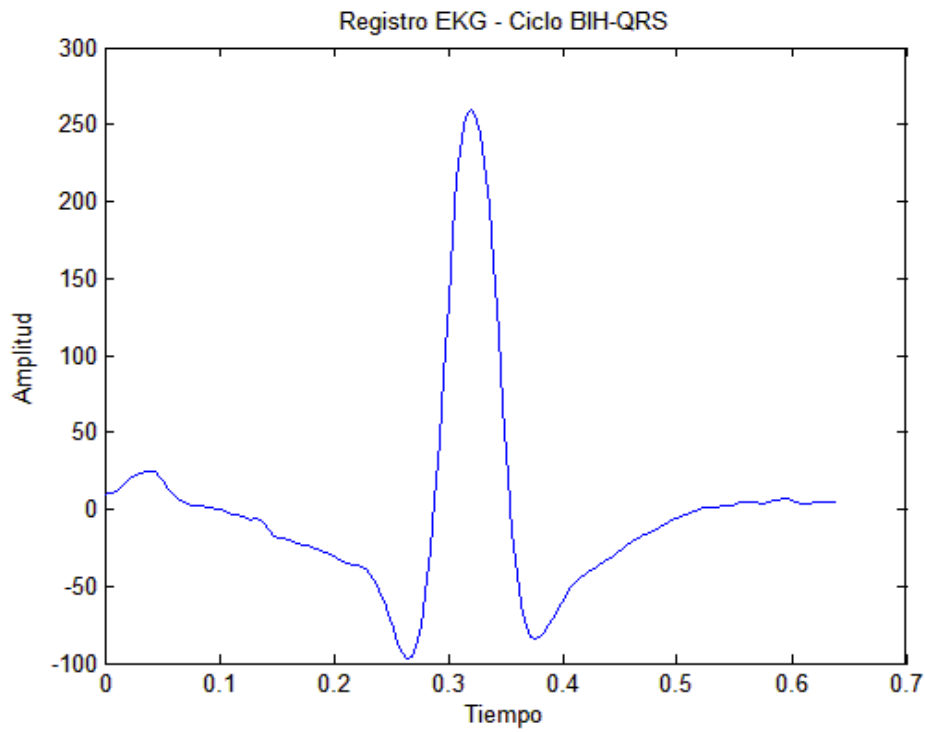


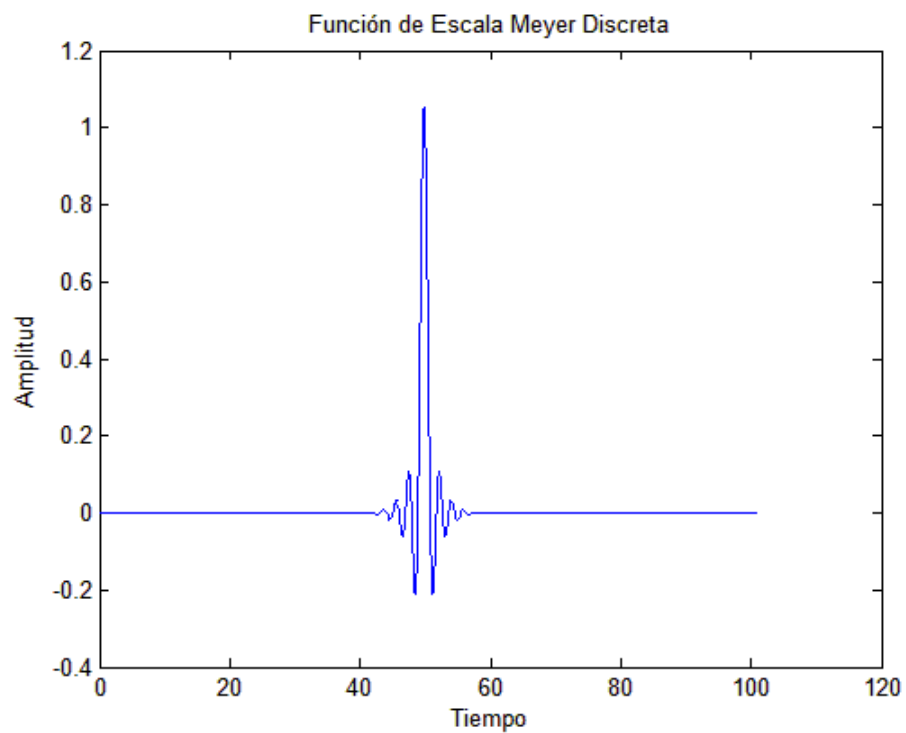
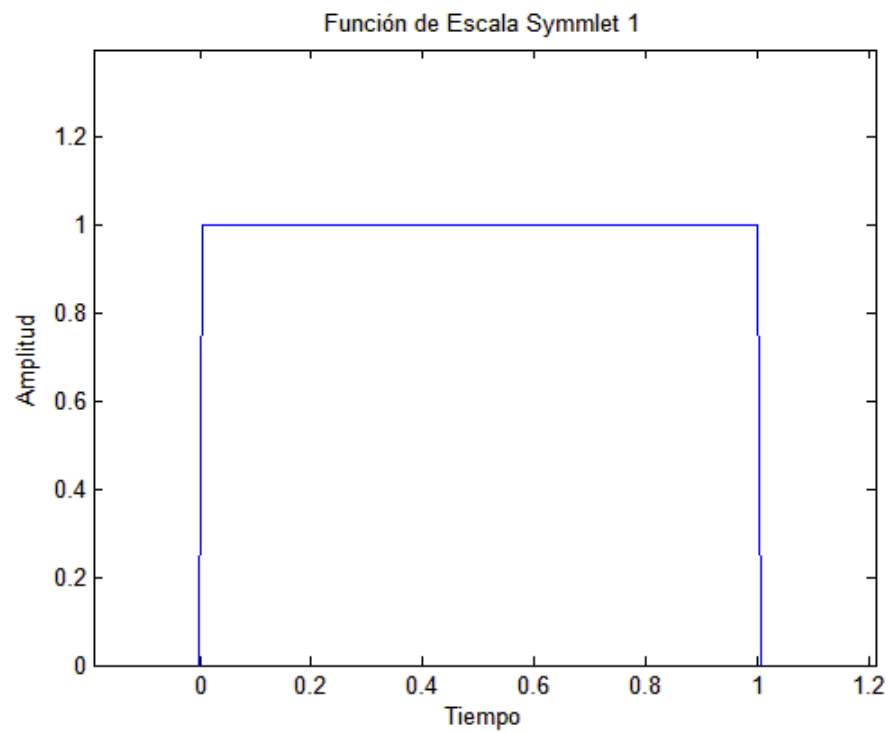


Anexo N°8: Ejemplo de complejos QRS de ambos bloques: BIH-QRS y ESC-QRS.

Inspección de un
Criterio y
Metodología de
Selección de
Bases Wavelet
para Aplicación
en señales de
EKG

41

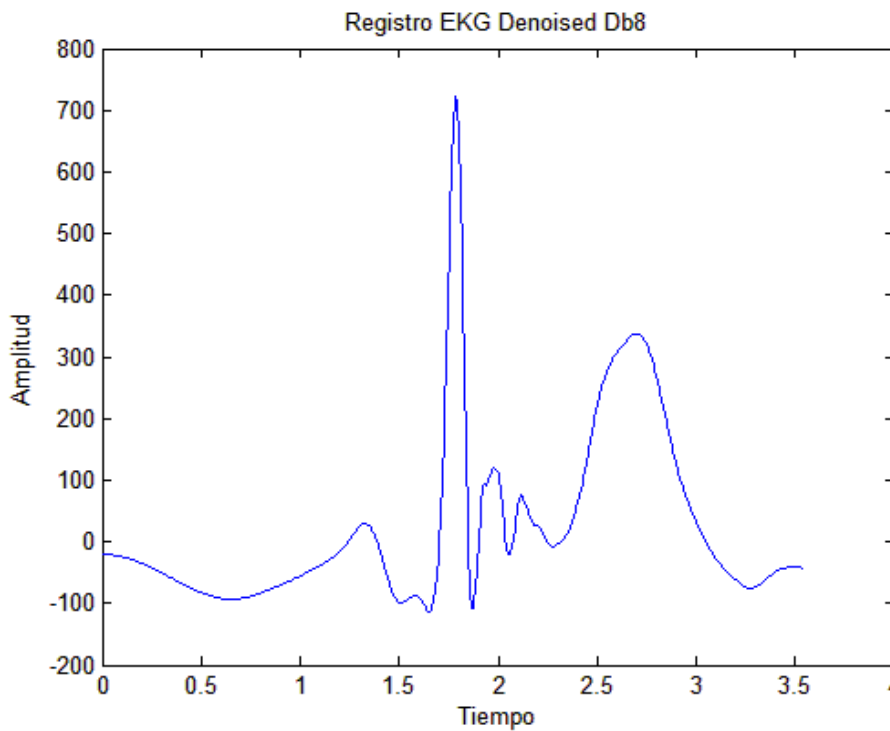
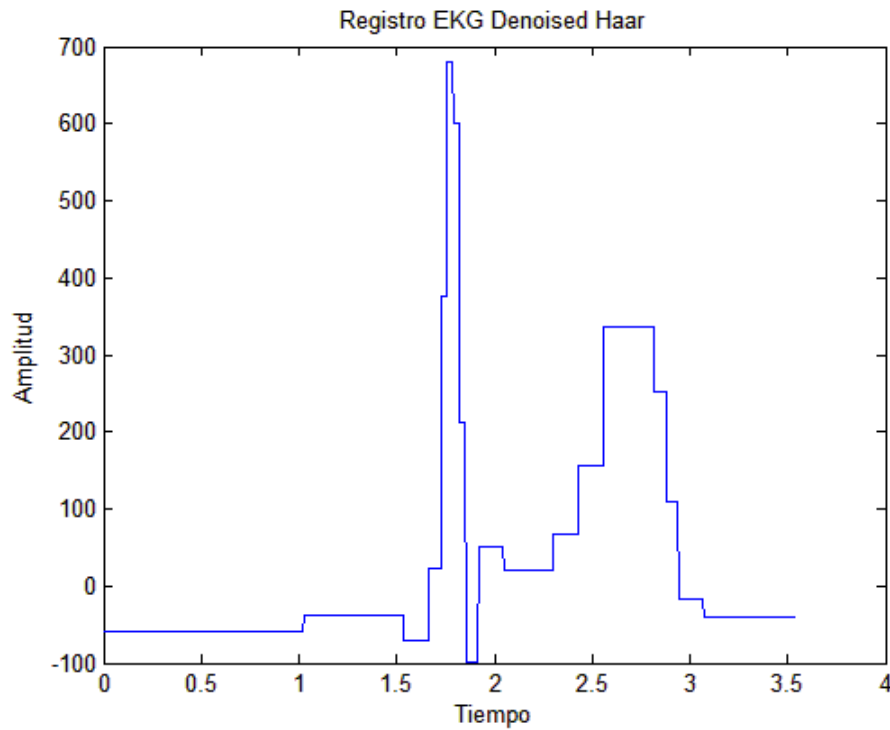




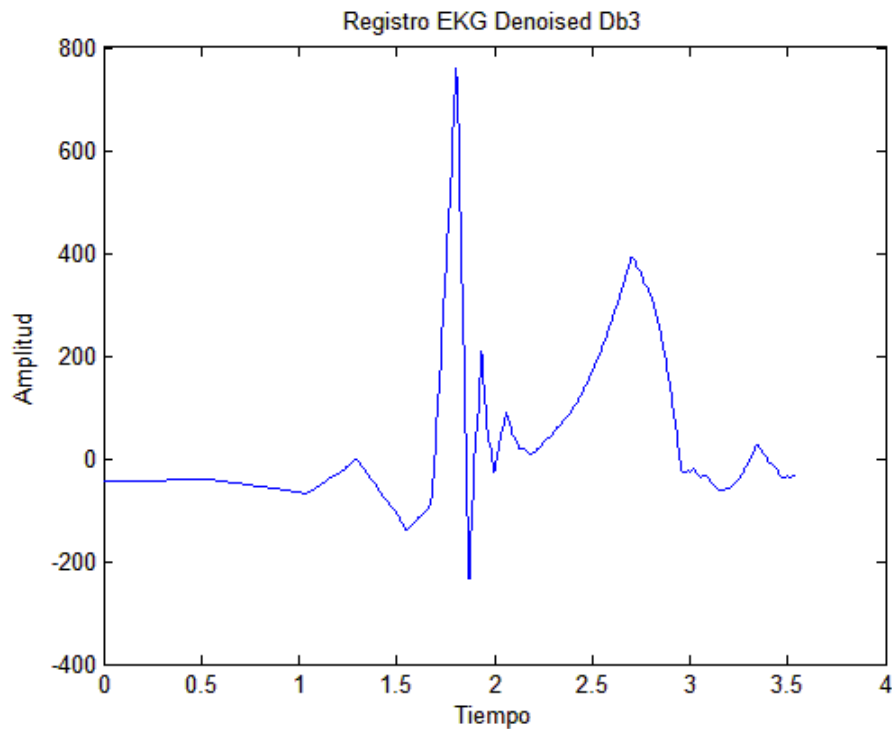
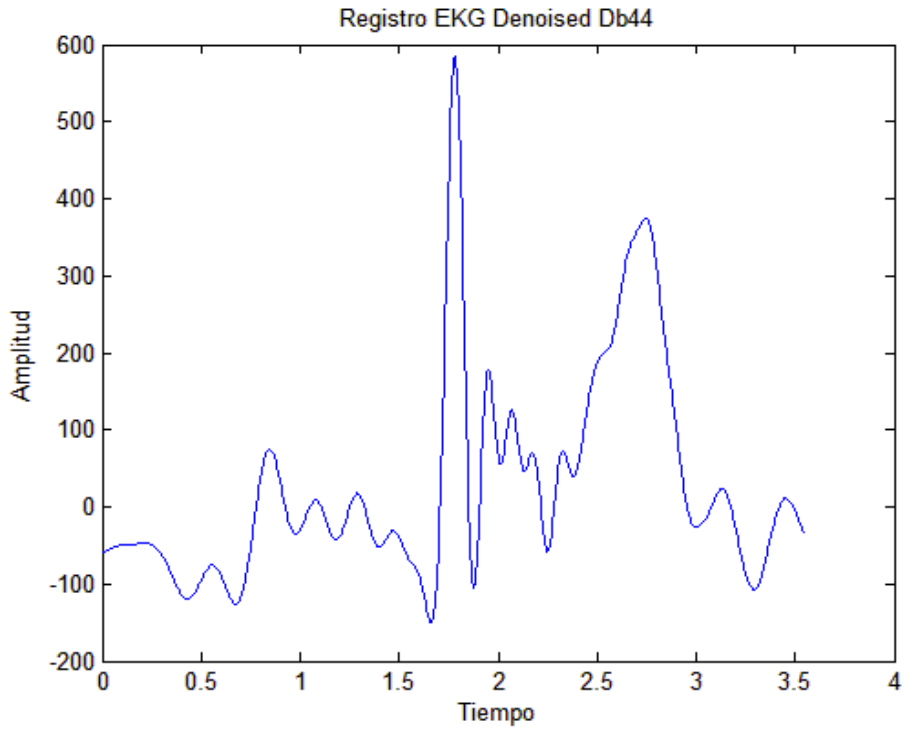
Anexo N°10: Registros EKG filtrados de ruido, reconstruidos con la transformada inversa de wavelet (Haar, Db8)

Inspección de un
Criterio y
Metodología de
Selección de
Bases Wavelet
para Aplicación
en señales de
EKG

43



Anexo N°11: Registros EKG filtrados de ruido, reconstruidos con la transformada inversa de wavelet (Db44, Db3). La función de escala madre de Db3 es idéntica a la función de escala madre de Sym3, por lo que sólo se presenta una de ellas.



Anexo N°12: Registros EKG filtrados, reconstruidos con la transformada inversa de wavelet (Db4, Db9).

Inspección de un
Criterio y
Metodología de
Selección de
Bases Wavelet
para Aplicación
en señales de
EKG

45

