

# Análisis Acústico sobre Señales de Auscultación Digital para la Detección de Soplos Cardíacos

## Acoustic Analysis Oriented to the Cardiac Murmur Detection over Digital Auscultation Signals

Andrés M. Castaño, Ing.<sup>1</sup>, Edilson Delgado T., MSc.<sup>2</sup>, Juan I. Godino, PhD.<sup>2</sup>, Germán Castellanos, PhD.<sup>2</sup>

<sup>1</sup>Grupo de Control y Procesamiento Digital de Señales, Universidad Nacional de Colombia

<sup>2</sup>Grupo de Bioingeniería y Optoelectrónica, Universidad Politécnica de Madrid, España  
{amcastanos, edelgadol, cgcastellanosd}@unal.edu.co, igodino@ics.upm.es

Recibido para revisión 26 de Marzo de 2007, Aceptado 30 de Noviembre de 2007, Versión final 9 de Diciembre de 2007

**Resumen**—Se presenta la metodología basada en el análisis acústico de señales fonocardiográficas (FCG) para detectar soplos cardíacos. En primer lugar se desarrolla un sistema de filtración basado en la transformada wavelet para reducir las perturbaciones que usualmente se presentan en la etapa de adquisición, ajustando la calidad del sonido de acuerdo a los requerimientos clínicos y validado por especialistas en semiología. Se propone un algoritmo de segmentación basado en la energía promedio normalizada de Shannon y la transformada wavelet. Sobre los segmentos se extraen características derivadas del análisis acústico y espectral. La efectividad de las características son evaluadas mediante un modelo en cascada de clasificadores del tipo máquina de soporte vectorial para separar 3 clases: normal, soplo y otros. La base de registros FCG utilizada pertenece a la Universidad Nacional de Colombia; de esta base de datos etiquetada se usaron 111 registros distribuidos así: 37 registros con etiqueta normal, 24 registros con etiqueta de soplo y 50 con etiqueta de otras anomalías. Se obtienen resultados de precisión de clasificación para los casos en que se caracterizan las señales filtradas y cuando se toman las señales originales sin filtrar, encontrando que el proceso de filtración incrementa considerablemente el acierto de clasificación hasta un 96%.

**Palabras Clave**—Aprendizaje de Máquinas, Análisis Acústico, Detección de Patologías, Soplos Cardíacos y Fonocardiografía.

**Abstract**—A methodology based on acoustic analysis of digitized phonocardiographic signals (PCG) is presented, oriented to detection of cardiac murmurs originated by valvular pathologies. Initially, a filtration system based on the wavelet transform is developed to reduce the disturbances that usually appear in the acquisition stage, adjusting the sound quality according to the clinical requirements and validated for specialists in semiology. A segmentation algorithm based on the normalized average Shannon energy and wavelet transform is

proposed. Features derived from the acoustic analysis are extracted on the segments. Feature effectiveness is evaluated by a support vector machine in cascade configuration for separating the classes: normal, murmur and other. The used database of phonocardiographic records belongs to the National University of Colombia, having 111 records as follows: 37 records labeled as “normal”, 24 labeled as “murmur” and 50 labeled as “other” abnormalities. The classification results are obtained with the original signals and when the signals have been filtered. The filtering stage increases the classification accuracy to 96%.

**Keywords**—Machine Learning, Acoustic Analysis, Pathologic Detection, Cardiac Murmurs and Phonocardiography.

### I. INTRODUCCIÓN

Las enfermedades cardíacas son una de las principales causas de mortalidad y morbilidad en la sociedad moderna [1], por lo tanto, cualquier avance en las técnicas de diagnóstico y tratamiento de enfermedades cardíacas, contribuye no solo al apoyo en el diagnóstico y tratamiento de estas enfermedades, sino también al entendimiento de la dinámica cardíaca y sus fenómenos subyacentes. La auscultación médica es un procedimiento clínico donde se escuchan los sonidos emitidos por las cavidades abdominal y torácica con el fin de emitir o aproximar un diagnóstico del estado de los órganos que se encuentran dentro de estas cavidades. En el caso particular de la auscultación cardíaca, se escuchan los sonidos que emite el corazón para evaluar el estado de las válvulas cardíacas, es decir, de las auriculoventriculares (mitral y tricúspide) que separan las aurículas de los ventrículos, y de las sigmoideas (pulmonar y

aórtica) que separan el corazón de las grandes arterias [2]. Las señales digitalizadas de los sonidos del corazón se conocen con el nombre de fonocardiogramas o señales fonocardiográficas (FCG), las cuales son fuentes de información de la actividad mecánica cardíaca de adquisición no invasiva y bajo costo económico. Las restricciones inherentes al método de auscultación tradicional, tales como: la incapacidad del oído humano para percibir todas las frecuencias sonoras, la subjetividad del examinador y las dificultades que se presentan en la descripción de hallazgos en la historia clínica, pueden ser significativamente disminuidas con el actual desarrollo de herramientas de procesamiento computacional, modelado matemático y análisis digital de señales FCG.

La literatura reporta el uso frecuente de la transformada *wavelet* (WT) para el procesamiento digital de señales FCG por la naturaleza altamente no estacionaria que se presenta en este tipo de señales; sin embargo, la mayoría de las *wavelets* evidencian errores relacionados al análisis de distribución de energía (ED), provocando que los resultados encontrados para la discriminación de señales FCG usando la transformada WT no sean los mejores [3]. En [4] se afirma que es importante mantener la longitud de la ventana de análisis tan corta como sea posible para garantizar la hipótesis de estacionariedad sobre los pequeños segmentos analizados. No obstante, esto puede reducir la resolución en frecuencia del espectrograma obtenido, pero si se ajusta el tiempo de deslizamiento de la ventana, se puede obtener un resultado aceptable. Para este caso, el espectrograma se calcula, en primera instancia, por medio de una ventana de corta longitud para generar una representación temporal del FCG, luego se aplica una ventana de mayor longitud para generar una representación espectral de la energía del FCG. La resolución en tales representaciones depende directamente de la longitud de la ventana. De esta manera, la representación temporal permite medir la duración de los ruidos y ciclos cardíacos, mientras que el espectro, asumiendo una buena resolución en frecuencia, permite la caracterización espectral de los sonidos cardíacos. En [5], se aplica la Transformada de Fourier de tiempo corto (STFT) para caracterizar señales FCG y obtener representaciones de parámetros tales como, localización en el tiempo de los ruidos cardíacos y sus componentes, contenidos de frecuencia, localización del ciclo cardíaco, y la forma de las envolventes en la caracterización de los soplos cardíacos. Adicionalmente, los espectrogramas han demostrado ser de uso clínico en la identificación de soplos analizando el contorno, a pesar de los problemas que se generan a causa de la resolución [6].

Debido a que las diferentes condiciones patológicas del corazón causan soplos y anormalidades en los sonidos cardíacos mucho antes de reflejarse en otros síntomas, la modificación de los sonidos del corazón podría ser obtenida mediante la auscultación cardíaca [7]. Particularmente, el retardo de algunos sonidos del corazón mejora la capacidad

de los mismos en el reconocimiento y la discriminación de características que resultan de diversos desórdenes cardíacos. Así por ejemplo, en [8] se emplean dos técnicas para analizar la separación de los sonidos por retraso: análisis WSOLA (*Waveform Similarity Overlap-and-Add*) y MTSM (*Multiresolutional Time Scale Modification*). En este proceso de modificación, se preserva la calidad de los ruidos cardíacos, y su espectro de frecuencia no es desplazado. Para evaluar el desempeño del método propuesto, se aplica un test de evaluación auditiva subjetiva por parte de 20 médicos expertos. Se demostró que el retraso de tiempo de algunos ruidos cardíacos mejora la capacidad de los médicos para reconocer y discriminar eventos que resultan de desórdenes cardíacos, por lo que se concluye que la modificación de los ruidos cardíacos facilita la interpretación de los sonidos obtenidos en el proceso de auscultación.

El análisis acústico tiene la ventaja de convertir el procedimiento subjetivo de las apreciaciones médicas en valores cuantitativos mediante técnicas digitales de procesamiento de señales, lo cual puede aprovecharse para la detección automática de soplos cardíacos. La caracterización de estados funcionales correspondientes a casos normales y con evidencias de soplos cardíacos usando procedimientos basados en el análisis acústico, permite recuperar la fonocardiografía como un procedimiento de bajo costo económico, válido e importante para el apoyo diagnóstico y evaluación de valvulopatías cardíacas.

## II. CONSIDERACIONES FISIOLÓGICAS

### A. Ciclo cardíaco

Los hechos que ocurren desde el inicio de un latido hasta el inicio del próximo se conocen como *ciclo cardíaco*. Cada ciclo cardíaco consta de un período de relajación, llamado diástole, durante el cual el corazón se llena de sangre, seguido de un período de contracción llamado sístole. El corazón cuenta con cuatro válvulas: las válvulas auriculoventriculares (A-V) y las sigmoideas. Las válvulas A-V (Mitral y Tricúspide) impiden el flujo retrógrado de la sangre de los ventrículos a las aurículas durante la sístole, y las válvulas sigmoideas (Aórtica y Pulmonar) impiden que la sangre de las arterias aorta y pulmonar regrese a los ventrículos durante la diástole [9].

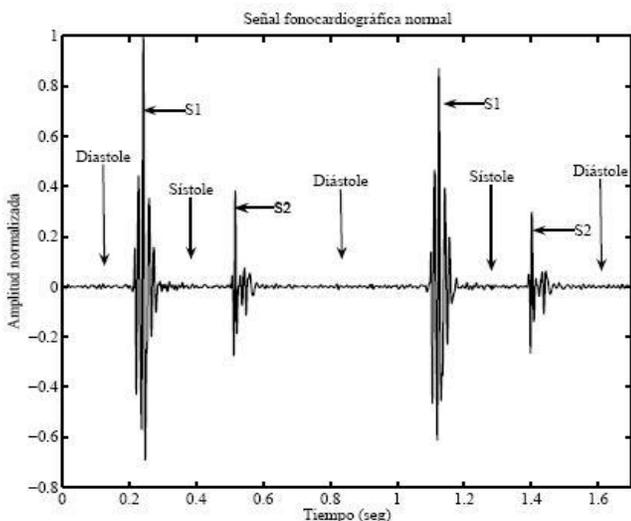
### B. Sonidos cardíacos

Reflejan el comportamiento mecánico de las válvulas que controlan la correcta dirección del flujo sanguíneo entre aurículas, ventrículos y conductos del corazón. El cierre o apertura de estas válvulas producen sonidos característicos que permiten deducir si el funcionamiento es adecuado, es decir, si se producen en el momento, manera e intensidad fisiológicamente normal. Cada latido tiene una duración aproximada de 0.8 *seg*. Durante la sístole ventricular, que dura 0.3 *seg*, los ventrículos se contraen para expulsar la

sangre hacia las arterias; en ese momento las aurículas están en la fase de diástole auricular (relajadas), y las válvulas mitral y tricúspide se cierran para evitar que la sangre retroceda hacia ellas, mientras que las aórtica y pulmonar se abren (primer sonido, *S1*). Durante la sístole auricular, que dura unos 0.15 *seg*, las dos aurículas se contraen para impulsar la sangre a los ventrículos, los cuales están en ese momento en la fase de diástole ventricular (relajados). En esos instantes las válvulas aórtica y pulmonar se cierran y las válvulas mitral y tricúspide se abren (segundo sonido, *S2*). Finalmente, se produce una fase de relajación de las aurículas y ventrículos (diástole general) que dura unos 0.4 *seg*, a partir de la cual se iniciará un nuevo ciclo. Las características en frecuencia correspondientes a cada uno de los sonidos cardíacos se exponen en la Tabla 1 y la forma de onda para una señal FCG donde pueden apreciarse de forma visual los sonidos *S1* y *S2* se muestra en la Figura. 1.

**Tabla. 1** Características de los sonidos cardíacos

Sonido Cardíaco	Duración (seg)	Rango de frecuencia (Hz)
<i>S1</i>	0.1 - 0.12	20 - 150
<i>S2</i>	0.08 - 0.1	50 - 60



**Figura 1.** Señal FCG normal

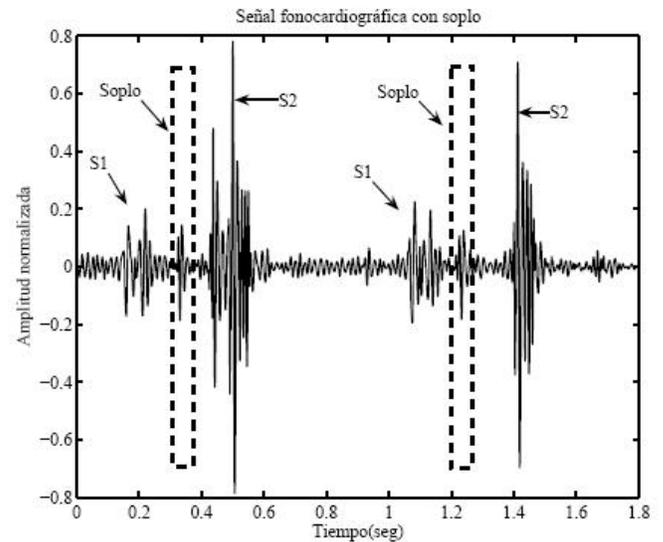
**C. Soplos cardíacos**

Son ruidos adicionales a *S1* y *S2* producidos por turbulencias en el flujo de la sangre a través de las distintas cámaras y grandes vasos. El flujo sanguíneo normal es laminar y por lo tanto es silencioso, mientras que cuando se hace turbulento se vuelve ruidoso y perceptible, dando origen al soplo. De acuerdo al instante en que se presenta el soplo puede considerarse sistólico o diastólico. En la Figura. 2 se aprecia una señal FCG que evidencia la presencia de un soplo cardíaco.

**III. TÉCNICAS DE PREPROCESAMIENTO Y PROCESAMIENTO DE LA SEÑAL**

**A. La transformada wavelet**

Esta técnica es una de las más sobresalientes en el análisis de señales no estacionarias. Las *wavelets* son familias de funciones que permiten el análisis de una señal para extraer características de espacio, tamaño y dirección. Una familia de *wavelets* se define como:



**Figura 2.** Señal FCG con evidencia de soplo cardíaco

$$h_{a,b} = \frac{h\left(\frac{x-b}{\alpha}\right)}{\sqrt{|\alpha|}} \tag{1}$$

que es generada a partir de funciones madre  $h(x)$ , agregándole la escala  $\alpha$  la cual permite que la onda se dilate o se contraiga, y la variable de traslación  $b$ , permitiendo que la señal se desplace en el tiempo.

La transformada *wavelet* permite obtener información en tiempo-frecuencia de señales no estacionarias. En el caso de la transformada *wavelet* continua CWT, se define como:

$$CWT_x^{\Psi}(\tau, \alpha) = \frac{1}{\sqrt{|\alpha|}} \int x(t) \Psi^*\left(\frac{t-\tau}{\alpha}\right) dt \tag{2}$$

donde la escala de la *wavelet*  $\alpha$  se puede considerar como el inverso de la frecuencia. El proceso de cálculo de la CWT se hace de manera similar al de la STFT. La CWT revela muchos detalles de la señal, debido a que realiza el proceso de cálculo en forma iterativa hasta que toda la señal es procesada cubriendo todas las escalas. No obstante, el costo computacional es mayor.

La transformada *wavelet* discreta permite realizar el mismo procedimiento de la CWT en menor tiempo computacional,

ya que calcula los coeficientes *wavelet* en intervalos discretos de tiempo y escala. Para el cálculo de la *DWT*, la *wavelet*  $\Psi$  es dilatada y trasladada en valores discretos. A menudo se usan potencias de 2 dilataciones (llamadas *diádicas*) [10]:

$$\psi(2^k t + l), \tag{3}$$

Por lo tanto la *DWT* de  $f$  es una función de escala  $2^k$  y tiempo  $l$ .

$$DWT(f)(2^k, l) = \int f(t)\psi(2^k t + l) dt \tag{3}$$

**B. Parámetros relacionados a la reducción de ruido**

La reducción de interferencias acústicas presentes en los registros de FCG, usando la transformada *wavelet*, requiere del ajuste de algunos parámetros, como son: el umbral, el nivel de descomposición y el reescalamiento.

1) *Métodos de umbralización*. Se analizan dos métodos de umbralización: fuerte (*hard thresholding*), en la cual los coeficientes de valor absoluto menor al umbral se hacen 0 y los demás se dejan igual:

$$x_u = \begin{cases} x, & |x| > x_0 \\ 0, & |x| \leq x_0 \end{cases} \tag{5}$$

El segundo método corresponde a la umbralización *suave* o *soft thresholding*, en la cual los coeficientes cuyo valor absoluto se encuentran por debajo del umbral también se hacen 0, pero los coeficientes restantes son reducidos a un valor cercano a 0:

$$x_u = \begin{cases} \text{sgn}(x)(|x| - x_0), & |x| > x_0 \\ 0, & |x| \leq x_0 \end{cases} \tag{6}$$

donde  $x_0$  es el umbral,  $x$  es la señal original y  $x_u$  es la señal umbralizada [11].

2) *Reglas para la selección de umbral*. Se presentan cuatro opciones para escoger la regla de selección del umbral más apropiada, las cuales se pueden observar en la Tabla 2.

**Tabla.2** Reglas de selección del umbral.

Regla	Descripción
Rigrsure	Selección usando el principio de "Stein's Unbiased Risk Estimate (SURE)"
Sqtwolog	Forma fija del umbral que es igual a la raíz cuadrada de dos veces el logaritmo de la longitud de la señal.
Heursure	Selección usando una combinación de las dos opciones antes mencionadas.
Minimaxi	Selección del umbral usando el principio del <i>minimax</i> .

Estas reglas utilizan regresiones estadísticas sobre el tiempo de los coeficientes correspondientes a las perturbaciones para obtener estimaciones no paramétricas de la señal reconstruida. Para la primera regla mencionada se realiza la selección basándose en el principio de *Stein's*

*Unbiased Risk Estimate (SURE)*, el cual calcula una estimación del riesgo para un valor particular de umbral  $x_0$ , para lo cual se requiere una selección del valor de umbral. El principio de *SURE* está dado por la ecuación (7) en función de un parámetro de suavizado  $\lambda$ , particularmente un umbral.

$$SURE(\lambda) = F(\lambda) - \sigma^2 + 2\sigma^2 \frac{N - N_0}{N} \tag{7}$$

donde  $F(\lambda)$  es el dato de entrada,  $\sigma^2$  es la varianza del ruido,  $N$  es el número de puntos de la señal, y  $N_0$  es el número de coeficientes *wavelet* correspondientes al ruido ubicados bajo el umbral. Este método, a diferencia de la umbralización universal, depende directamente de la señal de entrada, y no sólo de una estimación basada en los los datos de la varianza del ruido  $\sigma^2$  [11]. En la segunda regla de selección, se usa una forma fija de umbral que resulta en la acción del *minimax* multiplicado por un factor pequeño proporcional al logaritmo de la longitud de la señal  $I(x)$ .

$$sqt = \sqrt{2 \log(I(F))} \tag{8}$$

En la tercera regla se emplea una combinación entre las dos anteriores, donde, si la relación *señal-ruido* es muy pequeña, la estimación con *SURE* resulta muy ruidosa. Por lo tanto, se utiliza la forma fija del umbral. En la cuarta regla el umbral es escogido con base en el principio del *minimax* para el error cuadrático medio. Este principio es utilizado en la Estadística para diseñar estimadores porque obtiene el mínimo del máximo error cuadrático medio. Para realizar la selección del umbral, minimiza el riesgo máximo relativo con respecto a la selección óptima del riesgo. Sostiene que esta rata de riesgo también se comporta como  $2 \log N$  y el mismo umbral con *minimax* es asintótico ( $\square$ ) [11].

$$\lambda_{mM} \square \sqrt{2 \log N} \sigma \tag{9}$$

donde  $\lambda_{mM}$  representa la selección del umbral con *minimax*,  $N$  el número de puntos de la señal, y  $\sigma$  es la desviación estándar del ruido.

3) *Niveles de descomposición*. El proceso de descomposición puede ser iterativo, con aproximaciones sucesivas siendo descompuestas cada vez, de manera que la señal quede reducida en componentes de más baja resolución. Para este proceso se usa un método denominado *árbol de descomposición de wavelet*. En la práctica se selecciona un nivel de descomposición adecuado según la naturaleza de la señal [12].

4) *Métodos de reescalamiento*. En la Tabla 3 se listan los métodos disponibles en el toolbox de *wavelet* de Matlab. En el primer método se sigue el modelo básico del ruido, y no realiza reescalamiento. En el segundo se sigue el modelo básico de ruido pero con ruido no escalado, y realiza el

reescalamiento usando solo una estimación simple del nivel de ruido, el cual se calcula con base en los coeficientes del primer nivel de descomposición. El tercero corresponde al modelo básico de ruido con ruido no blanco, y el reescalamiento lo realiza basado en una estimación que depende del nivel de ruido en este nivel de descomposición [13].

**Tabla.3** Métodos de reescalamiento.

Método	Descripción
One	Sigue el modelo básico del ruido.
Sln	Modelo básico del ruido con ruido no escalado.
Mln	Modelo básico del ruido con ruido no blanco

5) *Waveles utilizadas*: Los algoritmos se implementan con *wavelets* ortogonales, pues permiten una mejor reconstrucción de la señal original y su proceso de cálculo es relativamente sencillo [14],[13]. Esto se debe a que cada coeficiente *wavelet* se calcula únicamente con un producto escalar entre la señal original y la *wavelet*. Su tiempo de cómputo es corto y permiten hallar la transformada con el mismo número de muestras de la señal original [15]. Las *wavelets* que satisfacen estas condiciones son las *Haar*, *Daubechies*, *Symlets* y *Coiflets* [13].

### C. Envolvente de Energía

El proceso de segmentación de señales FCG puede ser basado sobre la envolvente de la señal [16], calculada por medio de la energía de Shannon:  $E = -x^2 \log x^2$ . Este es el método más adecuado para el cálculo de la envolvente ya que es bondadoso en la intensidad media y atenúa el efecto de intensidad baja mucho más que el de intensidad alta en la señal; además hace más fácil encontrar sonidos con baja intensidad debido a la poca diferencia existente entre estos y los de alta. La energía promedio de Shannon esta definida como los promedios de energía de ventanas sucesivas traslapadas de una señal, esto es:

$$E_s(t) = -\frac{1}{N} \sum_{i=1}^N x^2(i) \log x^2(i) \quad (10)$$

donde,  $x$  es el valor de la señal y  $N$  el número de datos por ventana de la señal, y la energía promedio normalizada de Shannon es calculada como:

$$P_a(t) = \frac{E_s(t) - \bar{E}_s(t)}{S(E_s(t))} \quad (11)$$

donde  $\bar{E}_s(t)$  es el valor medio de  $E_s(t)$  y  $S(E_s(t))$  es la desviación estándar de  $E_s(t)$ .

### D. Análisis acústico

Este tipo de análisis hace referencia a la cuantificación de parámetros que describen comportamientos sonoros. En particular, el análisis sobre señales de auscultación digital requiere que el oído médico capture rasgos útiles en el descubrimiento de sonidos asociados a disfuncionalidades cardíacas. La automatización de estos procedimientos médicos se obtiene mediante la realización de algoritmos que obtienen medidas relacionadas a las cualidades sonoras y espectrales, formando el vector de características con las cuales se busca obtener patrones discriminantes entre las clases normal y patológica.

Un sonido puede ser considerado como la vibración del aire a diferentes frecuencias, de tal manera que una señal se puede descomponer como la suma de componentes de amplitud y frecuencia definidas [17]. En la dinámica cardíaca, los continuos movimientos de apertura y cierre de las válvulas producen sonidos que pasan a través de la piel y son capturados por un estetoscopio digital (señal FCG).

1) *Coefficientes cepstrales en la escala de frecuencia Mel (MFCC)*: Estudios fisiológicos han demostrado que la percepción humana del contenido espectral acústico no sigue una escala lineal, realmente, la percepción en las bajas frecuencias sigue una escala lineal y en frecuencias altas una escala logarítmica [18]. Los MFCC son una representación definida como los cepstrum reales de una señal ventaneada a corto tiempo obtenida mediante la transformada rápida de Fourier (FFT). La diferencia de los MFCC con los cepstrum reales es que se usa una escala no lineal de frecuencia. Así, los MFCC pueden ser considerados como una aproximación al comportamiento del sistema auditivo humano [19], partiendo de que el sistema auditivo funciona como un banco de filtros en las que el módulo de amplitud decae exponencialmente, y las frecuencias centrales de cada filtro están espaciadas de forma creciente y logarítmica. Los MFCC intentan emular este comportamiento mediante un banco de filtros que usa la escala Mel para la determinación de las frecuencias centrales de cada filtro. La energía que pasa por cada filtro corresponde a cada MFCC de la señal.

2) *Jitter*: Es una medida que toma tramos de la señal que contienen al menos un ciclo de la frecuencia fundamental  $f_0$  y determina la variación de la  $f_0$  de cada tramo con respecto a la  $f_0$  de la señal [19].

En la frecuencia existen otras medidas que ayudan a la representación de dinámicas sonoras como es el ancho de banda efectivo y envolventes de energía [20].

### E. Máquinas de soporte vectorial

La técnica de clasificación basada en las máquinas de soporte vectorial es uno de los métodos de aprendizaje

generalmente usado en el reconocimiento de patrones [21]. Este método consiste en buscar un hiperplano de máxima separación entre las clases, para lo que se requiere hacer una transformación no lineal de los datos de entrada (datos de entrenamiento) hacia un espacio de características de dimensión mayor donde las clases tengan máxima separación. En el hiper-espacio resultante se procede a realizar la búsqueda del hiperplano que divide las clases.

Sea el conjunto de datos de entrenamiento:

$$(x_1, y_1), (x_2, y_2), \dots, (x_r, y_r) \in X \times \{\pm 1\}$$

donde cada  $x_i \in \mathbb{R}^n$  es un elemento del conjunto  $X$  y  $y_i$  es la etiqueta de  $x_i$ . A este conjunto  $X$  se aplica una transformación no lineal de los datos, a través de un Kernel:

$$k: X \times X \rightarrow \mathbb{R}$$

$$(x, x') \rightarrow k(x, x') := (x, x')$$

En el espacio resultante se conservan las mismas etiquetas  $y_i$  para cada  $x_i$  y se hallan los elementos promedio de cada clase,

$$m_{k_1} = \frac{1}{r_{k_1}} \sum_{\{i, y_i=1\}} x_i \quad (12)$$

$$m_{k_2} = \frac{1}{r_{k_2}} \sum_{\{i, y_i=-1\}} x_i \quad (13)$$

donde  $r_{k_1}$  y  $r_{k_2}$  es el número de elementos que corresponden a la etiqueta  $y_i=1$  y  $y_i=-1$  respectivamente, de manera análoga para los promedios de clase  $m_{k_1}$  y  $m_{k_2}$ . De acuerdo a esto, se obtiene el vector  $\mathbf{w}$  cuyos extremos se ubican en los dos valores medios  $m_{k_1}$  y  $m_{k_2}$ , donde la magnitud de  $\mathbf{w}$  informa acerca de la distancia que existe entre las medias de las clases. Al tomar una muestra de prueba  $\mathbf{d}$  se procede a hallar el vector  $\mathbf{q}$  que relaciona la distancia entre  $\mathbf{d}$  y el punto medio ubicado entre las medias de las clases,  $\mathbf{q} = \mathbf{d} - \frac{m_{k_1} + m_{k_2}}{2}$ . Se analiza el ángulo  $\theta$

formado por  $\mathbf{q}$  y  $\mathbf{w}$ , de tal manera que si  $\frac{\pi}{2} > \theta$  y  $\theta < -\frac{\pi}{2}$

la muestra  $\mathbf{d}$  pertenece a una clase, y si  $\frac{\pi}{2} < \theta$  y  $\theta > -\frac{\pi}{2}$  pertenece a la otra clase. Por lo tanto, la predicción de la etiqueta de la muestra de prueba es:

$$y = \text{sgn} \left\{ \left( \mathbf{d} - \frac{m_{k_1} + m_{k_2}}{2} \right) \cdot m_{k_1} - m_{k_2} \right\} \quad (14)$$

$$= \text{sgn} \left\{ (\mathbf{d} \cdot m_{k_1}) - (\mathbf{d} \cdot m_{k_2}) + b \right\} \quad (15)$$

donde,

$$b = \frac{1}{2} \left( \|m_{k_1}\|^2 - \|m_{k_2}\|^2 \right) \quad (16)$$

Por último se reescribe la ecuación (15) basados en (12) y (13)

$$y = \text{sgn} \left\{ \frac{1}{r_{k_1}} \sum_{\{i, y_i=1\}} (\mathbf{d} \cdot x_i) - \frac{1}{r_{k_2}} \sum_{\{i, y_i=-1\}} (\mathbf{d} \cdot x_i) + b \right\}$$

de lo que se obtiene que:

$$y = \text{sgn} \left\{ \frac{1}{r_{k_1}} \sum_{\{i, y_i=1\}} k(x, x_i) - \frac{1}{r_{k_2}} \sum_{\{i, y_i=-1\}} k(x, x_i) + b \right\}$$

y  $b$  puede ser considerado como:

$$b := \frac{1}{2} \left\{ \frac{1}{r_{k_1}^2} \sum_{\alpha} k(x_i, x_j) - \frac{1}{r_{k_2}^2} \sum_{\beta} k(x_i, x_j) \right\} \quad (17)$$

donde  $\alpha$  es el conjunto de índices  $\{i, j: y_i = y_j = 1\}$  y  $\beta$  el conjunto de índices  $\{i, j: y_i = y_j = -1\}$ . Se puede observar entonces que las máquinas de soporte vectorial son útiles en la clasificación de patrones, solo basta con hacer una elección adecuada del kernel de transformación.

#### IV. MARCO EXPERIMENTAL

##### A. Base de Datos

La base de datos con la que se realiza el trabajo pertenece a la Universidad Nacional de Colombia y está conformada por registros FCG tomados a sujetos normales y a pacientes que evidencian la presencia de soplos cardíacos causados por valvulopatías cardíacas (estenosis aórtica, regurgitación mitral, etc). A cada paciente se le toman 8 registros correspondientes a los cuatro focos de auscultación (mitral, tricuspídeo, aórtico y pulmonar) en fase de apnea post-espíroratoria y post-inspiratoria, cada uno con una duración de 12 segundos y en posición decúbito dorsal. El tiempo no se pudo extender más debido a que los pacientes con problemas cardíacos son incapaces de sostener la apnea post-inspiratoria y post-espíroratoria por más tiempo. Las muestras fueron adquiridas a 50 sujetos normales y 98 pacientes con evidencia de soplo cardíaco con un estetoscopio electrónico marca WelchAllyn® modelo Meditrón. La toma de los registros se realiza sin limitaciones de edad y sexo de las personas que colaboraron en este procedimiento, siempre y cuando se haya verificado su mayoría de edad. Todos los registros normales se adquieren dentro de la cabina acústica del Hospital Infantil

Universitario “Rafael Henao Toro”, después de haber valorado a los sujetos mediante un chequeo médico para comprobar mediante la auscultación cardíaca el estado funcional del corazón e indicarles el propósito del estudio. De esta base de datos etiquetada se usaron 111 registros distribuidos así: 37 registros con etiqueta normal, 24 registros con etiqueta de soplo y 50 con etiqueta de otras anomalías.

### B. Descripción del Procedimiento

En la Figura. 3 se muestra el diagrama esquemático del procedimiento propuesto.

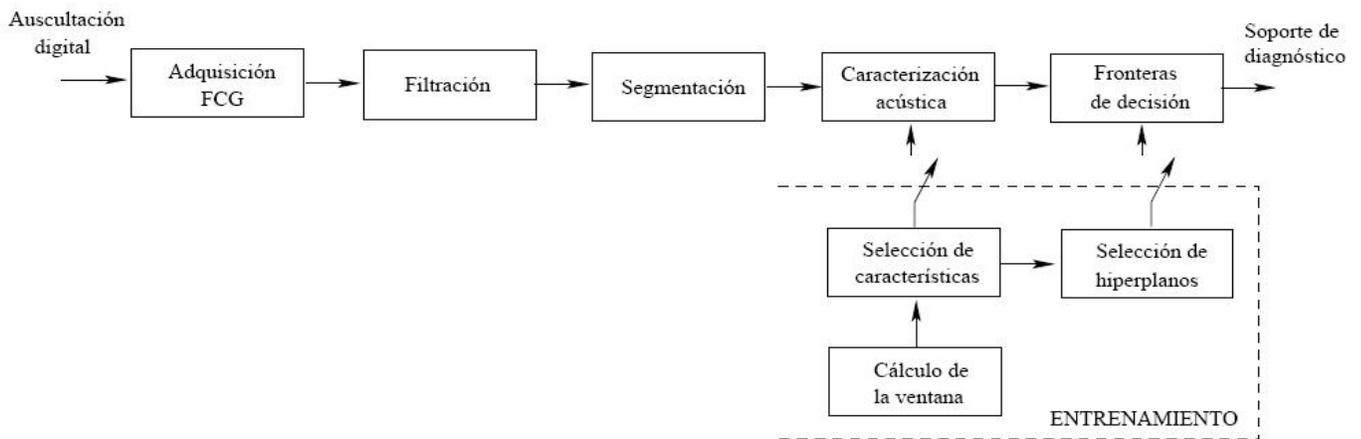


Figura 3. Diagrama de bloques del procedimiento propuesto

1) *Filtración*: Antes de iniciar cualquier tipo de procesamiento de las señales FCG es importante realizar una etapa de filtración con el objetivo de reducir interferencias acústicas que son involucradas en la etapa de adquisición, pero de ninguna manera se pueden modificar los atributos sonoros de la dinámica cardíaca. El procedimiento para lograr la reducción de las perturbaciones en los fonocardiogramas consistió en diseñar algoritmos del tipo *denoising*, usando cada una de las reglas de selección del umbral con los métodos de reescalamiento. Estas combinaciones se realizaron variando el nivel de descomposición desde 1 hasta 10 para diferentes familias *wavelets* (ver Sec. III-B5). Posteriormente, un grupo de tres médicos expertos en auscultación cardíaca escucharon los registros originales y los filtrados mediante *denoising*. A partir de los conceptos emitidos por el grupo de especialistas, se descartaron los algoritmos que alteraron de forma inadecuada las características sonoras de la señal y se eligió el que de forma más clara permitió percibir tanto los sonidos cardíacos como los soplos.

El método de validación matemática utilizado para este análisis es el estudio de la *relación señal-ruido SNR*, la cual

se obtiene mediante la siguiente expresión,

$$SNR = 10 \log \left( \frac{\sum_{n=1}^N (x(n) - \bar{x})^2}{\sum_{n=1}^N (x(n) - \bar{x}(n))^2} \right) \quad (18)$$

en donde  $x(n)$  es la señal original,  $\bar{x}(n)$  la señal reconstruida,  $\bar{x}$  la media de la señal y  $N$  el tamaño de la ventana sobre la cual se calcula. La obtención del valor de la *SNR* se realizó sobre las señales procesadas con los algoritmos que, a criterio de los médicos expertos en auscultación, presentaron los mejores resultados de filtración.

2) *Segmentación*: Se divide en las siguientes etapas:

--*Selección de picos*. Teniendo la envolvente de energía promedio normalizada de *Shannon*, usando ventanas de 0.02 *seg* traslapadas 0.01 *seg*, se realiza la identificación de los sonidos *S1* y *S2*. Se define un umbral para eliminar el efecto del ruido y las señales de muy baja intensidad, el cual puede ser determinado como sigue:

$$Th = kP_{\max} \quad (19)$$

donde  $P_{\max}$ , es el valor máximo de la envolvente y  $k$  es la constante que define cuanto porcentaje de amplitud de la envolvente se toma para determinar el umbral. En este trabajo el valor de  $k$ , escogido empíricamente, es 0.08. Los picos de la envolvente que exceden el umbral son escogidos, y se asumen temporalmente como sonidos pertenecientes al ciclo cardíaco para las señales que contiene soplos.

--*Remoción de picos extras*. Pueden existir problemas al escoger los picos correspondientes a los sonidos cardíacos, debido a que se pueden presentar en la envolvente de uno de los sonidos, picos locales, correspondientes a la existencia de un sonido dividido. Para eliminar los picos sobrantes se calculan todos los intervalos de tiempo entre picos adyacentes

( $t_i$ ), y seguidamente se estima un valor alto ( $I_a$ ) y uno bajo ( $I_b$ ), basados en la media ( $\bar{t}$ ) y la desviación estándar  $\sigma_t$  de los intervalos  $t_i$ , dados por:

$$I_b = \bar{t} - \sigma_t \quad (20)$$

$$I_a = \bar{t} + \sigma_t \quad (21)$$

Cuando un intervalo entre dos picos adyacentes es menor que el límite de tiempo bajo, existe un pico extra que debe ser removido, teniendo en cuenta:

- a) Si en un intervalo menor a 50 ms (tiempo de duración de un sonido cuando esta dividido) existen dos picos, se escoge el primer pico si su valor de amplitud es comparable con la del segundo; de lo contrario el primer pico es tomado como ruido y se escoge el segundo.
- b) Si el intervalo es mayor a 50 ms se escoge el pico con mayor valor.

--*Rescate de picos perdidos.* Algunos sonidos cardíacos  $S1$  ó  $S2$ , presentan menor energía que la del umbral escogido. Para rescatar lo picos perdidos se examinan los intervalos que sobrepasen el límite de tiempo alto, asumiéndose que un pico ha sido perdido, por tanto el umbral será disminuido tomando un factor  $K$  igual a 0.005 para encontrar dichos picos, y nuevamente se aplica el criterio para eliminar picos extra.

--*Identificación de  $S1$  y  $S2$ .* Las perturbaciones por artefactos pueden presentar picos de energía similares a los picos de los sonidos  $S1$  y  $S2$  que no pueden ser descartados usando los métodos antes mencionados. Se debe identificar cuál es  $S1$  y cuál es  $S2$ , teniendo en cuenta lo siguiente:

- a) El período diastólico es el intervalo más largo de la señal.
- b) El período sistólico es constante comparado con el período diastólico.

Para corroborar que la segmentación tenga una distribución sucesiva de sonidos  $S1$  y  $S2$ , se busca si existen dos sonidos del mismo tipo y en este caso es necesario encontrar el sonido faltante, por medio de la ubicación del máximo pico. Finalmente con la ubicación temporal de los sonidos cardíacos, es posible determinar el inicio y final para cada uno de los periodos sistólicos y diastólicos del registro FCG.

3) *Extracción de características acústicas:* La etapa de caracterización consiste en hallar un espacio coordinado  $p$ -dimensional que represente adecuadamente la señal FCG, donde  $p$  es el número de características. Así, en este nuevo espacio se busca una frontera de decisión que separe las clases a discriminar. Por lo tanto, a partir de la señal se extrae la información acústica o espectral relevante que permita distinguir entre un FCG normal y uno que presente

evidencias de soplo cardíaco. La caracterización acústica se realizó usando un ventaneo de longitud  $l$  y desplazamiento  $m$ , de manera que el valor estimado de cada característica para todo el registro se hizo igual al promedio de los valores obtenidos para cada característica en todas las ventanas. los valores de  $m$  y  $l$  fueron estimados en la etapa de entrenamiento de los algoritmos.

---

**Algoritmo 1** Procedimiento de diagnóstico usando las fronteras de decisión

---

**Require:**  $\xi = \{\xi_1, \dots, \xi_q\}$ ,  $\Psi_1$ ,  $\Psi_2$ ,  $\gamma_i$ .  $\xi$  es el vector de entrada,  $\Psi_1$  es la superficie de decisión obtenida para el primer clasificador,  $\Psi_2$  es la superficie de decisión obtenida para el segundo clasificador y  $\gamma_i$  es un escalar que permite sintonizar el resultado  $y$ .

1.  $y = \text{sign}\{(\xi)(\Psi_1) + \gamma_1\}$   $y$  puede tomar dos valores:  $\pm 1$ .
2. **if**  $y = 1$  **then**  
     Diagnóstico: NORMAL.  
     **Salir.**
3. **else**  
      $y = \text{sign}\{(\xi)(\Psi_2) + \gamma_2\}$   $y = 1 \rightarrow$  diagnóstico: SOPLO.  
      $y = -1 \rightarrow$  diagnóstico: OTROS.
4. **end if**

**Output:** Diagnóstico  $\rightarrow$  NORMAL, SOPLO u OTROS.

---

--*Estimación de los MFCC:* En este trabajo se calculan 14 MFCC para cada segmento intra-latido de la señal FCG ( $S1$ , *sístole*,  $S2$ , *diástole*). Se diseñan 14 filtros triangulares cuyas frecuencias centrales van desde 0 hasta  $f_s/2$  de forma logarítmica (ver Figura. 4), donde  $f_s$  es la frecuencia de muestreo de la señal ( $f_s = 3\text{kHz}$ ). Las amplitudes decaen exponencialmente hacia  $f_s/2$ . De esta manera, la representación de Fourier de cada segmento del latido,  $X$ , se multiplica por los coeficientes de cada filtro  $H_m$ :

$$X_F[m] = \ln \left( \sum_{k=0}^{N-1} |X[k]|^2 H_m[k] \right), 0 < m \leq M$$

Finalmente, se determinan los MFCC  $c[n]$  (1 por cada filtro) de la siguiente manera:

$$c[n] = \sum_{m=0}^{M-1} X_F[m] \cos(\pi n(m-1/2)/M), 0 \leq n \leq M$$

Siendo  $M = 14$  el número de coeficientes MFCC.

También se estimó el MFCC relativo, tomado como el valor máximo entre el cuarto MFCC sistólico y el cuarto MFCC diastólico, debido a que la frecuencia central de cuarto MFCC coincide con la mayor componente en frecuencia de los soplos cardíacos.

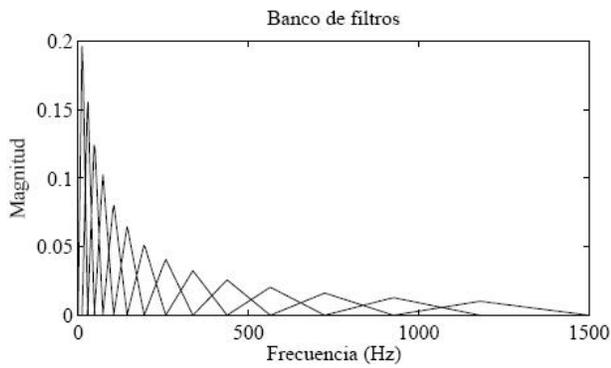


Figura 4. Banco de 14 filtros para obtener los 14 MFCC

--*Estimación del Jitter*. Inicialmente, el procedimiento requiere determinar la longitud de ventana que contenga al menos un ciclo de la componente de menor frecuencia con magnitud significativa en el espectro promedio de las señales FCG. Este valor se estimó en 20 Hz por lo tanto la longitud de la ventana se tomó de 5 ms. Luego, se hace un ventaneo de los segmentos ( $S1$ , *sístole*,  $S2$ , *diástole*), hallando por medio de la transformada rápida de Fourier la frecuencia fundamental  $f_0$  de cada tramo, desplazando la ventana con 80% de solapamiento. Finalmente, se obtiene el vector de frecuencias fundamentales de dimensión  $N$  para cada segmento, y se estima el Jitter, como:

$$\text{Jitter} = \frac{\left( \sum_{i=1}^N \left| \frac{f_0(i-1) + f_0(i) + f_0(i+1)}{3} - f_0(i) \right| \right)}{\sum_{i=1}^N f_0(i)}$$

Así, las medidas acústicas tomadas en cuenta para este trabajo fueron: los 14 MFCC para cada segmento intra-latido, el MFCC relativo y el Jitter. El conjunto total de características estimadas es de 58 y queda conformada una matriz de entrenamiento  $\mathbf{X}$  de dimensiones  $111 \times 58$ .

4) *Fronteras de decisión*: En esta etapa se desarrolla el soporte de diagnóstico automatizado que consta de 3 clases:  $k1$  - señal normal,  $k2$  - registro con evidencia de soplos y  $k3$  - señal con otro tipo de anomalías. Para este fin se construye un modelo en cascada de dos clasificadores (del tipo máquina de soporte vectorial). Para cada clasificador se obtiene una superficie  $n$ -dimensional que son usadas como fronteras de decisión. El primer clasificador distingue entre la clase  $k1$  y las clases  $k2$  y  $k3$ , quedando la tarea del segundo clasificador de discriminar entre las clases  $k2$  y  $k3$ . De esta manera, si a la entrada se tiene un vector con  $q$  elementos, este vector se toma como un punto en el espacio  $q$ -dimensional y de acuerdo a su ubicación con respecto a las fronteras de decisión, al vector de entrada se le asigna la etiqueta de cualquiera de las tres clases. En el Algoritmo 1 se describe el procedimiento de diagnóstico.

5) *Entrenamiento*: En el entrenamiento se ajustan los

parámetros de operación que rigen el sistema de detección.

a) *Longitud de la ventana de análisis*. En el procesamiento digital de señales el cálculo del ancho de la ventana de análisis es uno de los problemas más inmediatos a resolver y consiste en encontrar la longitud de la señal donde el comportamiento es cuasi-estacionario, de manera que las características estadísticas del segmento son independientes del tiempo en el que se tome. Para este trabajo el cálculo del ancho de la ventana se hizo mediante un procedimiento donde se estimaron los valores para las características variando dos parámetros de la ventana sobre la señal:  $l$  y  $m$ .  $l$  es la longitud la ventana en número de latidos y  $m$  es el desplazamiento de la ventana por latidos o por el porcentaje de la longitud promedio de un latido. La bondad de la ventana se estimó mediante la precisión de clasificación (PC) usando el conjunto de características a la entrada de los dos clasificadores en cascada. Los resultados se exponen en la Tabla IV.

b) *Selección de características*. Este procedimiento es usado para determinar el subconjunto de variables  $\zeta$  del espacio inicial de características  $\xi$  que contribuyen de manera efectiva en la representación. A este conjunto de variables se le conoce con el nombre de *características relevantes*. En general, el problema de selección de características consiste en encontrar un subconjunto  $\hat{\xi} \subset \xi$ , tal que, si el cardinal de  $\hat{\xi}_j$  es  $q$ , y en  $\tilde{\xi} = \{\hat{\xi}_j \subset \xi : j=1, \dots, M\}$  están todos los subconjuntos de cardinal  $q$ , se procura hallar el subconjunto  $\zeta = \hat{\xi}_j$ , para el cual se optimice la función de evaluación  $f_{\xi}$ , dado un conjunto de etiquetas de clase  $\mathbf{k}$  [22],

$$f_{\xi}(\mathbf{k}, \zeta) = \max_{\hat{\xi} \subset \xi} f_{\hat{\xi}}(\mathbf{k}, \hat{\xi}_j)$$

La literatura reseña rutinas *wrapper* (cuando  $f_{\xi}$  usa información del procedimiento de clasificación buscando minimizar el error del clasificador) que realizan la búsqueda en un número considerablemente menor de subconjuntos, entre las cuales se encuentra la selección flotante hacia adelante (SFFS) [23]. En esta técnica, cada paso incluye una nueva variable por medio de un procedimiento secuencial hacia adelante, pero luego realiza la exclusión de las variables menos significativas, una por una, hasta que la tasa de error de clasificación correcta disminuya. Una vez que ya no se puede seguir excluyendo variables se hace otro paso hacia adelante para incluir otra variable y nuevamente se realiza la exclusión de variables, si es posible. El proceso es recurrente hasta que ya no se

puedan efectuar más pasos hacia adelante debido a que la precisión de clasificación ya no se incremente.

c) *Selección de hiperplanos.* En esta etapa se tienen las matrices correspondientes a todas las pruebas con diferentes ventanas, se emplea entonces una SVM para generar los planos de decisión. Se escogen las fronteras que minimizan el error de validación, usando validación simple.

**V.RESULTADOS**

Los resultados obtenidos de la etapa de filtración se muestran en las Figuras. 5, 6, 7 y 8 donde se aprecian los diagramas tiempo-frecuencia de las señales normales y patológicas originales y filtradas. El mejor algoritmo de filtración fue seleccionado usando criterios de auscultación y apreciaciones auditivas de médicos especialistas.

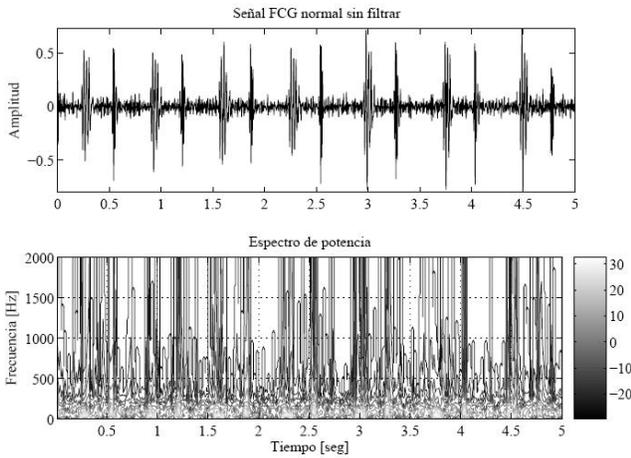


Figura 5. Respuesta tiempo-frecuencia de una señal FCG normal sin filtrar.

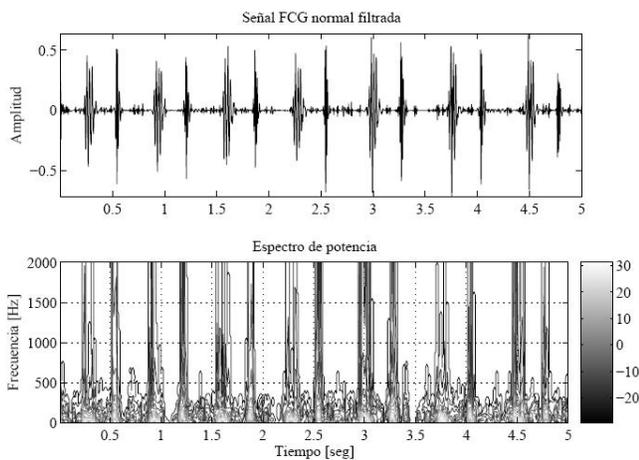


Figura 6. Respuesta tiempo-frecuencia de una señal FCG normal filtrada.

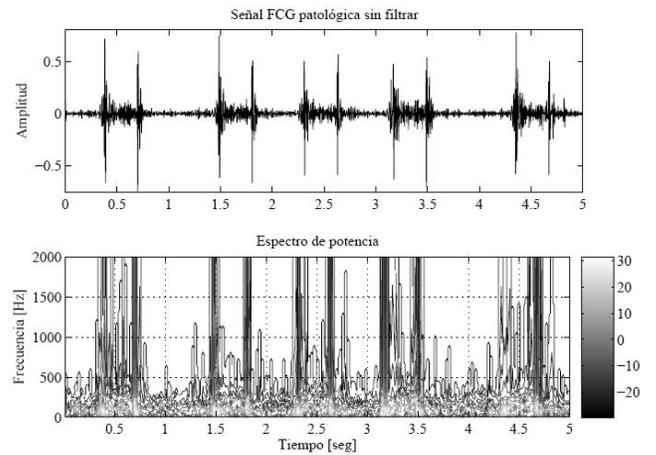


Figura 7. Respuesta tiempo-frecuencia de una señal FCG con soplo sin filtrar.

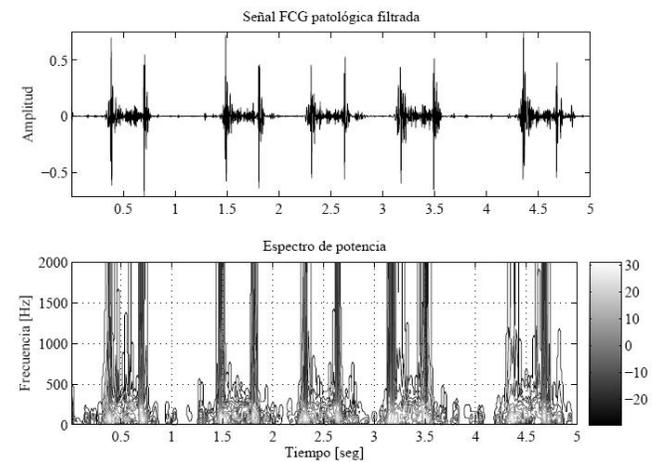


Figura 8. Respuesta tiempo-frecuencia de una señal FCG con soplo filtrada.

Los resultados de precisión de clasificación (PC) fueron usados en la determinación de la longitud de la ventana de análisis y se exponen en la Tabla 4. Se observa que los mejores resultados de clasificación obtuvieron cuando se tomó la ventana de longitud igual a un latido y el desplazamiento igualmente de un latido.

$$\text{Especificidad} = \frac{TN}{TN + FP} \times 100\% = 94,4\%$$

Tabla. 4 Bondad de clasificación para diferentes longitudes de la ventana de análisis

<i>I</i> (# latidos)	<i>m</i> (# latidos o % de un latido)	PC (%)
1	1	96
1	30%	62
1	60%	60
2	1	62
2	30%	60
2	60%	62
2	90%	58
3	1	62
3	2	62
3	30%	67
3	60%	67
3	90%	22

Estos valores significan que pocos latidos anómalos serán considerados como normales, lo cual es deseable. El hecho de que latidos normales sean considerados como posibles patológicos implica que el cardiólogo deberá comprobarlos de forma visual, pero esto no se considera un problema grave. De otra parte, se calcularon los resultados de clasificación usando las características extraídas de los registros sin filtrar y la precisión de clasificación fue considerablemente afectada, reduciéndose a un 61 %.

## VI. CONCLUSIONES

Los resultados obtenidos en la detección de soplos a partir del análisis de señales FCG indican que el conjunto de características propuesto ofrecen discriminación entre las clases.

La atenuación de perturbaciones acústicas con la técnica de filtración implementada en este trabajo, mediante descomposición y reconstrucción utilizando transformada wavelet discreta y umbralización adaptativa, mejora las condiciones de percepción auditiva de las señales FCG. La valoración fue analizada y determinada por médicos expertos en semiología cardíaca. Se obtuvo la atenuación de componentes acústicas que no correspondían a la información clínica, sin deteriorar las características auditivas de los sonidos cardíacos. Adicionalmente, la reducción de perturbaciones aplicado a señales FCG incrementa la separabilidad de las clases, permitiendo el uso del análisis acústico para el reconocimiento de eventos patológicos en los latidos cardíacos, ya que la filtración resaltó las componentes espectrales de los soplos reduciendo el ruido de envolvente. Cuando se hizo la caracterización de los registros sin filtrar, la precisión de clasificación se redujo sustancialmente.

El análisis relacionado con la longitud de la ventana es un paso fundamental para el análisis de señales no estacionarias, debido a que se minimiza la varianza de los valores estimados como características, de manera que al promediarlos se puede tomar la información como consistente y representativa sobre el comportamiento de la señal completa.

El sistema de detección construido resulta ser muy dependiente del desempeño del segmentador, es por esto que el máximo esfuerzo se concentra en el desarrollo y ajuste de esta etapa. Como trabajo futuro se propone el desarrollo de un segmentador más robusto ante condiciones considerables de ruido y de anomalías de la señal.

## AGRADECIMIENTOS

Este trabajo se encuentra en el marco de la investigación titulada: “*Técnicas de computación de alto rendimiento en la interpretación automatizada de imágenes médicas y bioseñales*” con código 20201004224, financiado por la DIMA y la vicerrectoría de Investigaciones de la Universidad Nacional de Colombia sede Manizales.

## REFERENCIAS

- [1] J. A. del Río, “Aspectos del perfil de morbimortalidad,” Hospital Universitario de Caldas, Manizales, Tech. Rep., 2002.
- [2] A. C. Guyton and J. E. Hall, *Tratado de Fisiología médica*, novena ed. Mc. Graw-Hill, 1997.
- [3] B. Ergen and Y. Tatar, “Optimal continuous wavelet analysis of periodograms signals,” in *IJCI Proceedings of International Conference on Signal Processing*, vol. 1, no. 2, September 2003.
- [4] A. Djebbari and F. B. Reguig, “Short-time fourier analysis of the phonocardiogram signal,” in *The 7th International conference on electronics, circuits and systems (ICECS’00)*, vol. 2, December 2000, pp. 844–847.H.
- [5] B. Ergen and Y. Tatar, “Time–frequency analysis of phonocardiogram,” in MEASUREMENT’03, *Proceedings of the 4th International Conference*, Smolenice Castle, Slovakia, June 2003.
- [6] M. A. Lamfri, C. Cravero, R. F. Sisteró, and C. M. Scavuzzo, “Análisis espectral de ruidos cardíacos adquiridos por un estetoscopio digital,” in *SIS’00, Simposio Argentino de Informática y Salud*, Tandil, Argentina, Septiembre 2000.
- [7] T. Modegi, “Xml transcription method for biomedical acoustic signals,” *Medinfo*, 2001.
- [8] Z. Telatar and O. Eroglu, “Heart sounds modification for the diagnosis of cardiac disorders,” in *IJCI Proceedings of International Conference on Signal Processing*, vol. 1, no. 2, September 2003, pp. 100–105.
- [9] F. Ganong, *Fisiología médica: El manual moderno*, 15th ed. México: Prentice Hall, 1996.
- [10] S. Sitharama Iyengarand, E. C. Cho, and V. V. Phoha, *Foundations of Wavelet Networks and Applications*. Chapman and Hall/CRC, 2002.
- [11] M. Jansen, *Noise Reduction by Wavelet Thresholding*. Springer-Verlag, 2001.
- [12] M. V. Wickerhauser, *Adapted Wavelet Analysis from Theory to Software*. New Jersey. USA: A K Peters, 1994.
- [13] S. Messer, J. Agzarian, and D. Abbott, “Optimal wavelet denoising for phonocardiograms,” *Microelectronics Journal*, no. 32, pp. 931–941, 2001.
- [14] M. Misiti, Y. Misiti, G. Oppenheim, and J. Poggi, *Wavelet Toolbox: For Use With MATLAB*. The Math Works Inc, 1996.
- [15] B. B. Hubbard, *The World According to Wavelets*. A K Peters, 1996.
- [16] S. Lukkarinen, H. Liang, and I. Hartimo, “A heart sound segmentation algorithm using wavelet decomposition and reconstruction,” in *19th Annual International Conference IEEE Engineering in Medicine and Biology Society*, Chicago, USA, 1997, pp. 1630–1633.
- [17] T. D. Rossing, F. R. Moore, and P. A. Wheeler, *The Science of Sound, 3rd ed. Addison Wesley*, 2001.
- [18] S. Molau, M. Pitz, R. Schluter, and H. Ney, “Computing mel-frequency cepstral coefficients on the power spectrum,” in *IEEE International Acoustics, Speech, and Signal Processing*, 2001, pp. 73–76.
- [19] A. Acero and H. W. Hon, *Spoken Language Processing*. Prentice Hall, 2001.
- [20] D. G. Childers, *Speech Processing and Synthesis Toolboxes*. USA: John Wiley & Sons, 2000.
- [21] C. J. C. Burges, “A tutorial on support vector machines for pattern recognition,” *Knowledge Discovery and Data Mining*, vol. 2, p. 22, 1998.
- [22] A. Jain and D. Zongker, “Feature selection: Evaluation, application, and small sample performance,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 2, pp. 153–158, 1997.
- [23] A. R. Webb, *Statistical Pattern Recognition*, 2nd ed. John Wiley & Sons Ltd, 2002.

**Andrés Marcelo Castaño.** Ingeniero electrónico. Egresado de la Universidad Nacional de Colombia en la promoción 2007. Su área de trabajo es el procesamiento de señales biomédicas y reconocimiento de patrones. Actualmente trabaja en el Grupo de Control y Procesamiento Digital de Señales.

**Edilson Delgado-Trejos.** Profesor catedrático. Departamento de Ingeniería Eléctrica, Electrónica y Computación en la Universidad Nacional de Colombia Sede Manizales. Recibió sus títulos como Ingeniero Electrónico y como M.Sc. en Automatización Industrial en los años 2000 y 2003, respectivamente. Actualmente cursa último semestre del Doctorado en Ingeniería en la Universidad Nacional de Colombia. Sus áreas de desempeño más importantes se ubican en el tratamiento de señales biomédicas, entrenamiento de máquinas inteligentes, reconocimiento de patrones y análisis de estructuras multivariadas.

**Juan Ignacio Godino.** Profesor asociado. Departamento de Ingeniería de Circuitos y Sistemas en la Universidad Politécnica de Madrid.

**Germán Castellanos-Domínguez.** Profesor Asociado. Departamento de Ingeniería Eléctrica, Electrónica y Computación en la Universidad Nacional de Colombia Sede Manizales. Pregrado en diseño y desarrollo de receptores MSK (profundización en análisis de aleatoriedad en señales y sistemas, ingeniería electrónica y proceso de señales) y Ph.D. en dispositivos y sistemas de proceso en el Instituto Nauchno Issledovatel'skiy de Rusia en los años 1985 y 1990, respectivamente. Coordinador del Grupo de Control y Procesamiento Digital de señales. Su área de trabajo es el proceso y análisis de señales aleatorias, usando herramientas de aprendizaje de máquina y reconocimiento de patrones.