

# ANÁLISIS DE CLASIFICADORES ESTOCÁSTICOS PARA LA DETECCIÓN EFECTIVA DE SOPLOS CARDÍACOS A PARTIR DE SEÑALES FONOCARDIOGRÁFICAS

## ANALYSIS OF STOCHASTIC CLASSIFIERS FOR EFFECTIVE DETECTION OF HEART MURMURS FROM PHONOCARDIOGRAPHIC SIGNALS

Miguel A. Becerra Botero<sup>1\*</sup>; Cristian Mejía Arboleda<sup>2</sup>; Milton Sarria Paja<sup>2</sup>; Leonardo Duque Muñoz<sup>2</sup>

---

<sup>1</sup> Grupo de Electrónica y Automática (GEA), Institución Universitaria Salazar y Herrera - Colombia. migb2b@gmail.com

<sup>2</sup> Grupo SINERGIA, Instituto Tecnológico Metropolitano – Colombia. medixina@gmail.com, miltonsarria@gmail.com, leonardoduque@itm.edu.co.

---

Recibido: Marzo 5 de 2012

Aceptado: Mayo 7 de 2012

\*Correspondencia del autor. Institución Universitaria Salazar y Herrera - Colombia. Cra. 70 No 52-49 Medellín-Colombia. e-mail: migb2b@gmail.com

### RESUMEN

Los sistemas de diagnóstico automatizados para la detección de soplos cardíacos, descritos en la literatura, registran en intervalos cortos de tiempo la dinámica fisiológica a partir de bases de datos de señales fonocardiográficas segmentadas, obteniendo una multiplicidad de muestras del mismo paciente sin tener en cuenta la dinámica cardíaca de cada individuo, lo cual no garantiza una adecuada identificación de la anomalía cardíaca. El sistema de asistencia de diagnóstico propuesto inicia con el estudio de 1060 señales tomadas de los cuatro focos de auscultación cardíaca de 144 pacientes, agrupados principalmente en dos clases: normales y patológicos. Los registros se agruparon por su posición en el periodo cardíaco en 4 tipos: normales, soplo sistólico, diastólico y sistodiastólico, obteniendo una base de datos de señales fonocardiográficas, a la cual se le realiza un preprocesamiento, sin embargo, no se segmenta para conservar la dinámica cardíaca de cada individuo en el estudio. Luego se genera un espacio de representación a partir de los PLP y los coeficientes cepstrales calculados a partir de la FFT y la STFT, desarrollando un análisis con los clasificadores estocásticos HMM y GMM y técnicas de selección de características para la extracción de información relevante de la dinámica fisiológica, que permitan un adecuado y eficiente entrenamiento del sistema con una tasa adecuada de clasificación para soporte de diagnóstico clínico.

**Palabras Clave:** Procesamiento de señales, Modelos de Mezclas Gaussianas, Modelos Ocultos de Markov, Soplos Cardíacos, Detección de Patologías

**ABSTRACT**

Automated diagnostic systems for the detection of cardiac murmurs, described in the literature, recorded in short time intervals, the physiological dynamic from databases of segmented phonocardiography signals, obtaining a multiplicity of samples from the same patient regardless the cardiac dynamics of each individual, which does not guarantee proper identification of cardiac abnormality.

The diagnostic system aided proposed begins with the study of 1060 signals taken from the four areas of cardiac auscultation in 144 patients, mainly grouped into two classes: normal and pathological. The records are grouped by their position in the cardiac period in four types: normal, systolic murmur, diastolic and systolic-diastolic, obtaining a database phonocardiographic signals, it is preprocessing, however, it isn't segmented for to conserve cardiac dynamics of each individual in the study. Then generate a representation space from the PLP and the cepstral coefficients calculated from the FFT and STFT, developing a stochastic analysis with HMM and GMM classifiers and feature selection techniques to extract relevant information from the physiological dynamics, to permit adequate and efficient training of the system with an adequate rate of classification for clinical diagnostic support.

**Key words:** Signal processing, Gaussian Mixture Models, Hidden Markov Model, Heart murmur, Pathology detection

**INTRODUCCIÓN**

La cardiología es una de las subespecialidades de la medicina que ha facilitado el abordaje de diferentes patologías del corazón, haciendo más objetivo su diagnóstico. Parte fundamental del método clínico utilizado consiste en observar, palpar, percudir y auscultar, como herramientas para acceder directamente al paciente y cuyo éxito depende de los conocimientos, la experiencia y la agudeza de los sentidos de quien evalúa, razón por la cual, toda anomalía cardíaca como la estenosis o las insuficiencias valvulares, que involucran sonidos denominados soplos, demanda esas características del evaluador para poder acercarse a un diagnóstico certero. (1,2).

Actualmente, la identificación de anomalías cardíacas se ha facilitado con la utilización de aparatos tecnológicos y por el desarrollo en las imágenes diagnósticas y de laboratorio, pero el acceso a este tipo de exámenes sigue siendo limitado o nulo por sus altos costos o por lo apartado de algunas zonas geográficas del país, especialmente, en las áreas rurales.

Por lo anterior, se hace relevante retomar la clasificación de sonidos cardíacos de manera automática por medio de técnicas de reconocimiento de patrones (3) y la estandarización de algunos signos clínicos como los ruidos cardíacos por medio de la fonocardiografía, técnica en la que poco ha intervenido la electrónica (4) a pesar

de que se han desarrollado prototipos con DSP (Digital Signal Processing) para el monitoreo de sonidos cardíacos en tiempo real, y su clasificación como normales y patológicos, obteniendo rendimientos adecuados pero no óptimos (5, 6), lo que a su vez, genera la necesidad de un mayor nivel de avance en las investigaciones en esta área para brindar al médico una herramienta asequible, confiable y segura y, a los pacientes, un mayor beneficio reflejado en diagnósticos más acertados y rápidos, haciendo que su pronóstico pueda cambiar en relación con el abordaje de su enfermedad.

La literatura reporta, un uso frecuente de cuatro etapas para la clasificación de soplos cardíacos, las cuales son adquisición, procesamiento, caracterización y clasificación, donde la mayoría de estudios usan bases de datos de pocos individuos segmentando las señales para obtener multiplicidad de muestras, lo que mejora la tasa de rendimiento de los clasificadores pero se desprecia la dinámica del individuo lo cual es de alta relevancia en este tipo de estudio (7). Para la eliminación de ruido usualmente se utilizan filtros digitales convencionales aunque también se han realizado investigaciones combinando diferentes técnicas como la Descomposición de modo empírico (EMD), descomposición de valores singulares (SVD) y análisis de componentes independientes (ICA) incrementado sustancialmente la relación señal a ruido de la señal FCG (fonocardiográfica) (8). La segmentación también permite incrementar la se-

parabilidad de las clases, en la cual se han aplicado diferentes técnicas como los HMM (Modelos Ocultos de Markov) en combinación con la energía de Shannon y los MFCC (Coeficientes Cepstrales de las Frecuencias de Mel) para la caracterización alcanzando una tasa de acierto del 93% para la segmentación de los sonidos cardiacos S1 y S2. (9), pero se obtiene una mayor tasa de acierto cuando se toma un menor número de latidos en la clasificación (10), de igual forma se ha obtenido la CW(characteristic waveform) y el CMW(Characteristic moment waveform) a partir de la señal cardiaca y expresado mediante el método de la integral de Viola, obteniendo una segmentación automática más rápida que la obtenida por medio de la wavelet (11), en varias técnicas de segmentación de señales FCG se utiliza una señal auxiliar como la ECG (Electrocardiografía), pero al realizar la segmentación automática de la señal mediante la detección del primer y segundo sonido cardiaco sin el uso de señal auxiliar se obtiene una exactitud superior al 97.47% (12). Posterior al preprocesamiento viene la etapa de caracterización encontrando que las técnicas más utilizadas para la caracterización de señales de audio tanto de voz como de FCG son los MFCC (3,13,14,15,16), los cuales son lineales en la región de las frecuencias de los sonidos cardiacos pero a pesar de ello, se obtienen resultados inferiores con los coeficientes cepstrales convencionales (17). Finalmente, viene la clasificación de la señal utilizando técnicas estocásticas, las cuales resultan más adecuadas para la clasificación del audio debido a su alta aleatoriedad, siendo aún mayor en las señales FCG, por lo que los clasificadores basados en HMM y GMM (Modelos de Mezclas Gaussianas) son considerados una buena opción por diferentes autores, mejorando su rendimiento al trabajar con otras técnicas como: GMM-SVM (Maquinas de Soporte Vectorial) obteniendo una tasa de acierto del 99.67% comparada al 95.82% para los GMM en el reconocimiento de voz (14, 17). GMM, HMM y otras variaciones de los HMM como los FDHMM (Fuzzy Discrete Hidden Markov Models) obtienen una adecuada clasificación alcanzando tasas de acierto hasta del 100% a partir de la extracción de características por medio de la STFT (Transformada de Fourier en Tiempo Corto), wavelet, entre otras, para generar los espacios de representación con diferentes técnicas como los MFCC, los LPC (Coeficientes de Predicción Lineal), los CMN (Media de los Coeficientes cepstrales normalizados), entre otros (15,16,18,19). De igual forma la clasificación basada GMM-4 y análisis cepstral muestra que la señal debe ser procesada en segmentos de 0.5 segundos usando una resolución completa en el dominio de la fre-

cuencia requiriendo el uso de coeficientes cepstrales de alto orden para la identificación de personas, notando que son más útiles los LFBC(Linear Frequency Bands Cepstra) que los MFCC. (13), donde la mayor parte de estudios sobre señales FCG, desprecian la dinámica del individuo para medir el rendimiento de los sistemas de clasificación.

Por lo anterior es de interés realizar estudios que por medio de técnicas estocásticas como HMM y GMM en combinación con una adecuada generación del espacio de representación teniendo en cuenta la dinámica total de cada individuo, para mejorar el rendimiento de los clasificadores de señales FCG como normales y con soplos de diferentes tipos, para generar un adecuado diagnóstico.

En este artículo se desarrolla una metodología para clasificar señales FCG capturando la dinámica total del individuo, evitando que este sea utilizado en la etapa de entrenamiento y validación al mismo tiempo, clasificando las señales en 3 tipos de soplos y normales, generando un espacio de representación mediante los MFCC calculado a partir de la STFT, los LFCC (Coeficientes Cepstrales de Frecuencias Lineales) y los PLP (coeficientes de predicción lineal perceptual), para la caracterización de la señal y se realiza un análisis de rendimiento con un clasificador basado en HMM ergódico y un clasificador basado en GMM, para el cual se realizan pruebas de rendimiento con y sin reducción de la dimensionalidad del espacio de representación mediante PCA (Análisis de Componentes Principales).

## MATERIALES Y METODOS

La metodología desarrollada se describe en la figura 1.

### Base de datos:

La base de datos de sonidos fonocardiográficos con la que se realizó la investigación pertenece a Universidad Nacional de Colombia Seccional Manizales y a Telesalud de la Universidad de Caldas, la cual consta de 905 señales obtenidas a partir de los cuatro focos de auscultación cardiaca en inspiración y expiración de 146 sujetos.

### Etiquetado de las señales

El médico experto Cristian Mejía, del grupo de investigación SINERGIA del ITM; analizó la base de datos y seleccionó 316 señales con soplos y 516 normales, perteneciente a 49 y 65 sujetos diferentes respectivamente,



**Figura1.** Esquema Metodológico

obteniendo una clasificación de las señales con soplos así: 33 diastólico, 1 holodiastólico, 71 holosistólico, 6 protodiastólico, 76 protosistólico, 26 sistodiastólico, 91 sistólico, 8 soplo eyectivo y 4 telesistólico. Debido a la diversidad de clases, se agruparon las señales con soplos por posición en el periodo cardiaco etiquetándolas como: 26 sistodiastólico, 40 diastólico, 242 sistólico y 526 normales.

### Preprocesamiento

Las señales fueron adquiridas con una frecuencia de muestreo de 44100Hz y remuestreadas a 4410Hz y normalizadas pero no filtradas, esto con la finalidad de analizar el rendimiento de los sistemas de clasificación sin aplicar filtrado de la señal, teniendo como punto de comparación la clasificación realizada por el médico, la cual se limitó a la audición y visualización de la señal.

### Extracción de características

Se generaron diferentes espacios de representación para pruebas a partir de: los MFCC, la derivada de los MFCC, los LFCC y los PLP. Los coeficientes cepstrales fueron calculados como se ilustra en la figura 2.

MFCC: estos coeficientes fueron calculados realizando un ventaneo hamming de 50 ms con un traslapamiento de 25 ms entre ventanas, seguidamente la señal en el tiempo fué representada mediante la STFT (Short Time Fourier Transform) (20) establecida en (1).

$$T_{STFT}(t, f) = \int_{-\infty}^{\infty} x(T)H(T - t)e^{-j2\pi Tf} dt \quad (1)$$

A la representación obtenida, se le aplicó un banco de filtros distribuidos en la escala de frecuencias Mel, que en el caso de estudio son 20, luego se calculó la energía de cada uno de los filtros y finalmente se aplicó el logaritmo y la Transformada Coseno (DCT), obteniendo así los coeficientes MFCC y la derivada del mismo, generando el vector de características de cada señal, que para este caso son 20 coeficientes MFCC y 20 que corresponden a su derivada.

LFCC: estos coeficientes fueron calculados de igual manera que los MFCC, pero los filtros están distribuidos de forma lineal.

PLP (Predicción Lineal Perceptual): estos presentan

una mayor inmunidad al ruido que los coeficientes cepstrales (7). En el caso de estudio, se obtuvieron 128 coeficientes a partir de cada ventana la cual es del mismo tamaño que la utilizada para el cálculo de los coeficientes cepstrales.

### Reducción del espacio de representación

Se realizó una reducción del espacio de representación por medio de PCA (21), con la finalidad de analizar el rendimiento de los clasificadores con el espacio reducido y sin reducir.

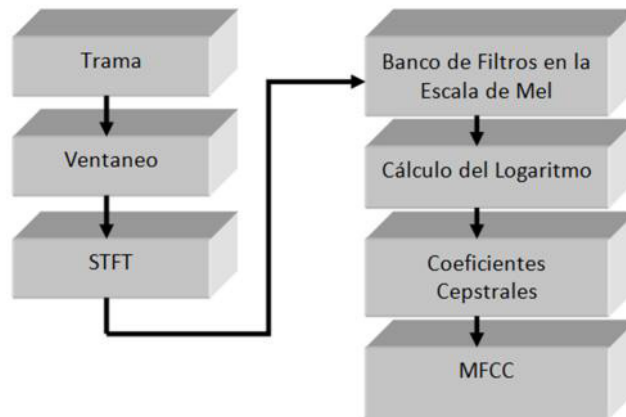


Figura2. Procedimiento para la obtención de los MFCC

### Entrenamiento y Reconocimiento: GMM- HMM

Se implementaron dos clasificadores estocásticos uno basado en GMM y otro basado en HMM, ambos emplean el algoritmo de agrupamiento de k-medias.

El clasificador basado en GMM utiliza el algoritmo de optimización EM (Máxima esperanza) con un máximo número de iteraciones de 1000, el número de gaussianas  $M=16$ .

Los HMM son considerados un caso particular de los modelos de Markov, donde los eventos ya no son observables permaneciendo ocultos y su observación corresponde a una función probabilística del estado (22, 23).

El clasificador desarrollado basado en HMM ergódico, se inicializó estimando los parámetros aleatoriamente pero cumpliendo restricciones estadísticas y el número de estados y gaussianas fueron determinados aleatoriamente ejecutando pruebas con 5 estados y 16 gaussianas como lo recomiendan para este tipo de señal (24), ejecutando el algoritmo para una tolerancia de 10-5 y un máximo de 1000 iteraciones.

### Validación

Se utilizó validación cruzada 80 - 20, con 11 fold por cada tipo de soplo con respecto a todas las demás señales. Y se calculó la tasa de reconocimiento correcto, la sensibilidad y especificidad para la evaluación del sistema.

### RESULTADOS Y DISCUSIÓN

Con la base de datos etiquetada de señales FCG, se realizaron 11 pruebas por cada tipo de validación para determinar el comportamiento en cuanto a especificidad y sensibilidad del clasificador GMM y del clasificador HMM ergódico continuo con los coeficientes MFCC, LFCC y PLP generados a partir de la STFT, con los cuales se realizaron pruebas con y sin reducción del espacio de representación y mediante PCA. En la tabla 1 se presenta el promedio de los resultados obtenidos de la ejecución del entrenamiento y validación con estos clasificadores.

Los resultados arrojados muestran que el clasificador basado en HMM presenta una tasa de acierto de 86% para soplo sistólico, 86% para soplo diastólico, 88% para soplo sistodiastólico y 88% para señales normales, mientras que el clasificador basado en GMM presenta una tasa de acierto de 65,2%, 78%, 85% y 82% respectivamente, ambos clasificadores obtuvieron las tasas más altas para la caracterización de la señal utilizando los MFCC, LFCC y PLP conjuntamente sin aplicar reducción en el espacio de representación.

Se observó que los clasificadores estocásticos basados en HMM y GMM presentan un menor desempeño con un espacio de representación reducido por medio de PCA. Los resultados arrojados muestran que el porcentaje de reconocimiento correcto es mayor para el clasi-

**Tabla 1.** Tasa de acierto clasificador GMM y HMM

Tipo de soplo	Clasificador	Reducción con PCA	Tasa de acierto	Especificidad	Sensibilidad
Sistólico	GMM	NO	65.2%	74%	45%
		SI	63%	49%	78%
	HMM	NO	86%	88%	82%
		SI	68,7%	74%	63%
Diastólico	GMM	NO	78%	79%	75%
		SI	56,8%	56%	63%
	HMM	NO	86%	87%	83%
		SI	0,78	78%	78%
Sistodiastólico	GMM	NO	85%	85%	65%
		SI	65%	66%	42%
	HMM	NO	85%	85%	81%
		SI	77%	77%	73%
Sin soplo	GMM	NO	82%	85%	81%
		SI	65%	62%	70%
	HMM	NO	88%	85%	90%
		SI	74%	64%	80%

**GMM:** Modelos de mezclas Gaussianas

**HMM:** Modelos ocultos de Markov

ficador basado en HMM que para el basado en GMM, pero a pesar de esto el desempeño de los HMM no es adecuado para diagnóstico.

## CONCLUSIONES

Se propone una metodología para la detección de soplos en señales FCG empleando un sistema de clasificación basado en HMM Ergódico continuo sin realizar segmentación, lo que permite reducir el costo computacional y discriminar la señal de cada individuo mediante la captura de la dinámica total de esta, evitando así que el sistema entrene y valide el mismo sujeto. La evaluación comparativa de los modelos estocásticos HMM y GMM muestra que los resultados obtenidos en cuanto a sensibilidad, especificidad y tasa de acierto, son bastante inferiores a los obtenidos en otras investigaciones (19), lo cual puede atribuirse a que no se realizó segmentación de la señal FCG (9), ni filtrado adicional al realizado por hardware, afectando de manera significativa la tasa de reconocimiento (8), a pesar de que los MFCC y los PLP

reducen el error y proporcionan características bastante robustas cuando la señal es afectada por el ruido (7), de igual forma es de notar que la reducción del espacio de representación mediante PCA afecta notoriamente el rendimiento de los clasificadores estocásticos y resulta bastante complejo clasificar adecuadamente soplos cardiacos cuando se realizan pruebas de rendimiento teniendo en cuenta la dinámica de cada individuo.

A pesar de que no se obtuvieron resultados satisfactorios en el desempeño de los clasificadores, se propone realizar pruebas con éstos, pero ejecutando el preprocesamiento de la señal con un filtro pasabanda para eliminar de mejor manera el ruido de alta y baja frecuencia, adicionalmente se propone generar un espacio de representación más amplio que permita modelar de mejor manera la dinámica cardiaca y teniendo en cuenta las características de cada individuo y finalmente se propone aplicar un algoritmo de segmentación automática para volver a realizar pruebas con los clasificadores estocásticos.

## Agradecimientos

Se agradece a la Universidad Nacional de Colombia sede Manizales por facilitar la base de datos de sonidos cardiacos. Este trabajo se enmarca dentro del proyecto titulado “Metodología Dinámica Sobre Espacios de Representación Abstracta Basada en Técnicas Estocásticas Orientada a la Detección de Soplos Cardiacos a partir de los Cuatro Focos de Auscultación Cardiaca”, que realiza el grupo GEA de la IUSH en colaboración con el Grupo MIRP y el grupo SIENERGIA del ITM. El proyecto se identifica con el código 0403 en la Institución Universitaria Salazar y Herrera y con código PM11117 en el ITM.

## BIBLIOGRAFÍA

1. Cediell AR. Métodos Exploratorios: Semiología Médica. Quinta edición. Mexico: Editorial Celsus; 2002.
2. Mcphee SJ. Transtornos cardiovasculares: Fisiopatología de la Enfermedad. Sexta edición. Bogotá, Colombia: 2011.
3. Katarina K. Biological Signals In Medical Diagnostics. AIP Conference Proceedings, 2010; 1204: 147-150.
4. Sarbandi RR, Doyle JD, Navidbakhsh M, Hassani K, Torabiyani H. A color spectrographic phonocardiography (CSP) applied to the detection and characterization of heart murmurs: preliminary results. Biomedical Engineering OnLine, 2011; 10:42.
5. Balasubramaniam D, Nedumaran D. Design and development of digital signal processor. Based phonocardiogram system. International Conference on Signal Acquisition and Processing. IEEE Computer Society, 2010; 366-370.
6. Ari S, Sensharma K, Saha G. DSP implementation of a heart valve disorder detection system from a phonocardiogram signal. Journal of Medical Engineering & Technology, 2008; 32 (2): 122-132.
7. Becerra M, Sarria M. Metodología dinámica sobre espacios de representación abstracta basada en técnicas estocásticas orientadas a la detección de soplos cardiacos. XVI Simposio de tratamiento de señales, imágenes y visión artificial. Pontificia Universidad Javeriana, Cali-Colombia: 2011; 183-186.
8. Warbhe AD, Dharaskar RV, Kalambhe B. A Single channel Phonocardiograph Processing using EMD, SVD, and EFICA. ICETET '10 Proceedings of the 2010 3rd International Conference on Emerging Trends in Engineering and Technology, 2010; 578-581
9. Ricke AD, Povinelli RJ, Johnson MT. Automatic Segmentation of Heart Sound Signals Using Hidden Markov Models. Computers in Cardiology, 2005; 32: 953-956.
10. Castaño A, Delgado E, Godino J, Castellanos G. Análisis Acústico sobre Señales de Auscultación Digital para la Detección de Soplos Cardiacos. Revista avances en sistemas e informática, 2007; 4(3): 171-182.
11. Yan Z, Jiang Z, Miyamoto A, Wei Y. The moment segmentation analysis of heart sound pattern. Computer methods and programs in biomedicine, 2009. XXX. :1-11
12. Ari S, Kumar P, Saha G, A robust heart sound segmentation algorithm for commonly occurring heart valve diseases. Journal of Medical Engineering & Technology, 2008; 32 (6): 456-465
13. Phua K, Chen J, Dat T, Shue L. Heart Sound as a biometric. The journal of the Pattern Recognition society, 2007; 41: 906-919
14. Wang X, Zhang J, Yan Y. Discrimination between Pathological and Normal Voices using GMM-SVM approach. Journal of voice, 2009; 25 (1): 38-43.
15. Kandali B, Routray A, Kumar BT. Emotion recognition from assamese speeches using MFCC features and GMM classifier. TENCON 2008 - 2008 IEEE Region 10 Conference, 2008. 1-5
16. Wang L, Kitaoka N, Nakagawa Seiichi. Robust distant speaker recognition based on position-dependent CMN by combining speaker-specific GMM with speaker-adapted HMM. Speech Communication Journal, 2007; 49 (6): 501-513.

17. Cheang L, Subari K, Abdullah M, Ahmad N, Besar R. Comparison of MFCC and Cepstral Coefficients as a Feature Set for PCG Biometric Systems. *World Academy of Science, Engineering and Technology*, 2010; 68: 754-758.
18. Jancovic P, Kokuer M. Incorporating the voicing information into HMM-based automatic speech recognition in noisy environments. *Speech Communication*, 2009; 51: 438-451.
19. Uguz H, Arslan A, Saracoglu R, Turkoglu I. Detection of heart valve diseases by using fuzzy discrete Hidden Markov Model. *Expert Systems with Applications*, 2008; 34(4): 2799-2811.
20. Kao WC, Wei C, Liu J, Hsiao P. Automatic Heart Sound Analysis with Short-Time Fourier Transform and Support Vector Machines. *Circuits and Systems, Midwest Symposium on*, 2009; 188-191.
21. Duda R, Hart P, Stork D. *Pattern Classification*. Second Edition, New York: Editorial: John Wiley & Sons, Inc; 2000.
22. Sierra A. *Aprendizaje Automático: conceptos básicos y avanzados*, Departamento de Ciencias de la computación e Inteligencia Artificial. Universidad del País Vasco. Madrid, España: editorial Prentice Hall; 2006.
23. Rabiner, L. A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. *Proceedings of the IEEE*, 1989; 77 (2): 257-286.
24. Sunita C, Ping W, Chu S, Anantharaman V. A computer-aided MFCC-based HMM system for automatic auscultation. *Computers in Biology and Medicine*, 2007; 38: 221-233.