

Detección y extracción automática de eventos S1, S2, S3 y S4 en sonidos del corazón

P. Mayorga¹, G. Chavez¹, V. Arguelles¹, C. Druzgalski², V. Zeljkovic³

¹ Instituto Tecnológico de Mexicali, Depto. de Posgrado, Mexicali, México

² California State University, Elec. Eng. Dept., Long Beach, CA, USA

³ The Lincoln University, CPES Dept., PA, USA

christopher.druzgalski@csulb.edu

Resumen. Debido a la limitación de los rangos de frecuencia en la percepción humana, las enfermedades cardiacas son difícilmente detectadas solo por auscultación con estetoscopio clásico. En muchos de los casos resulta en un diagnóstico tardío, y además de esto las técnicas actuales no logran ser suficientes o los equipos son sofisticados y costosos. Diversas propuestas han surgido en la literatura para otros eventos acústicos como la voz, estos se apoyan en transformadas de Hilbert-Huang, y detección de eventos. Por lo que en el presente artículo se propone un método novedoso de detección automática de eventos en señales HS. La propuesta en este trabajo está basada en Detección de Actividad de Voz, Modelos Mezclados Gaussianos (VAD-GMM), y la transformada de Hilbert (HT). Los resultados son alentadores logrando hasta un 97% de eficiencia en la clasificación mediante Modelos Ocultos de Markov (HMM), para eventos de S1, S2, S3 y S4 en sonidos del corazón (HS).

Palabras clave: Sonidos del corazón (HS), detección de actividad de voz (VAD), modelos mezclados gaussianos (GMM), transformada de Hilbert (HT).

Detection and Automatic Extraction of Events S1, S2, S3 and S4 in Heart Sounds

Abstract. Because the human hear has limitations in some perception ranges, the cardiac diseases are hardly detected by traditional methods of auscultation. In many cases it results in a late diagnosis, and also the current techniques can not be enough, even sophisticated equipment's are too expensive to be used in simple medical offices. Some proposals have arisen in the literature for acoustic events such as the voice, these are based on Hilbert-Huang transform, and event

detection. Therefore, in the present paper we propose a novel method of automatic detection of events in Heart Sounds (HS), signals. Our proposal is based on Voice Activity Detection (VAD), implemented with Mixed Gaussian Models (VAD-GMM) and the Hilbert Transform (HT). The results are motivated, because the efficiency in classification using Hidden Markov Models (HMM), for S1, S2, S3 and S4 events in heart sounds has reached 97%.

Keywords: Gaussian mixed models (GMM), heart sounds (HS), Hilbert transform (HT), voice activity detection (VAD).

1. Introducción

El uso de la tecnología ha facilitado y mejorado los procesos en la medicina. Tal es el caso de la auscultación cardiaca y respiratoria, la cual por medio del avance de la tecnología ha logrado mejorar las técnicas de auscultación aplicando el estetoscopio, y de esta manera tener una mejor percepción de los sonidos cardiopulmonares. Una alternativa en la reducción de ruido es el filtrado (pasa banda, pasa baja, pasa altas), que permite eliminar frecuencias no necesarias y que afectan la detección de los eventos S1 y S2 [1]. Sin embargo, las limitaciones con las que se cuenta en los centros de atención médica han originado propuestas como grabaciones de sonidos cardiacos utilizando micrófonos con condensador, es decir, convirtiendo los estetoscopios tradicionales en estetoscopios digitales, con ello se facilita el acceso a consultorios y centros de investigación que cuenten con recursos limitados [1]. En la actualidad existen distintos niveles de atención médica [2], algunos básicos y otros más especializados; los niveles de atención médica básicos cuentan con cierta funcionalidad, implicando al menos equipo como un estetoscopio. Además, por normatividad se debe contar con una computadora para el expediente electrónico del paciente [3, 4], lo cual es asequible.

Desafortunadamente, el estetoscopio presenta varios retos, como el ruido ambiental y el traslape de los Sonidos del Corazón (HS), con los Sonidos del Pulmón (LS). Debido a esto, la percepción de sonidos cardiacos queda restringida a la capacidad y experiencia del médico, y puesto que hay frecuencias fuera de los rangos auditivos humanos se dificulta diagnosticar con certeza la existencia de alguna enfermedad [5]. Debido a lo anterior, es necesario un sistema que no dependa del oído humano, que pueda detectar y clasificar sonidos cardiopulmonares por métodos automatizados y computarizados. Algunas aproximaciones son dirigidas a enfermedades endémicas, en donde se utilizan las características acústicas de la tos y crepitaciones para reforzar vectores de Coeficientes Cepstrales en Frecuencia Mel (por sus siglas en inglés, MFCC), aplicando ondículas (Wavelets), [6]. Otros trabajos se destinan al análisis y monitoreo de las ondas sonoras del corazón [7]. Otros autores mencionan que, si bien el uso de estetoscopio es una herramienta de bajo costo, los movimientos de los pacientes sobre todo en niños contaminan los registros de los sonidos [8]. En [9], se propone un algoritmo basado en un método de doble umbral para una detección robusta de los sonidos cardiacos S1 y S2.

La señal original del sonido del corazón (HS), es filtrada aplicando una ventana de Hamming. La envolvente de sonido cardiaco se extrae mediante la Transformada de

Tabla 1. Características principales en sonidos HS.

Sonido	Punto de auscultación	Frecuencia	Características	Duración	Forma de auscultación
S1	Mitral con mayor intensidad que el tricúspide	30- 120 Hz	Causado por la sístole	0.08 – 0.16seg. (0.14seg)	Diafragma del estetoscopio
S2	Mitral	70-150 HZ	Por el cierre valvular aórtico	0.06 – 0.12seg. (0.11seg)	Diafragma del estetoscopio
S3	Mitral	27-70 Hz	Diástole por disfunción ventricular	0.04 - 0.08seg	Campana del estetoscopio
S4	Mitra 1	10-50 Hz	Ruido auricular por tensión en válvulas	0.06 – 0.08seg	Campana del estetoscopio

Hilbert-Huang (HHT), y se segmenta el sonido cardiaco por el método de doble umbral [9]. En [10], se presenta un método de baja complejidad para la detección del primer y segundo sonidos del corazón (S1 y S2), y los períodos de sístole y de diástole sin necesidad de utilizar una referencia electrocardiográfica.

El algoritmo utiliza el Modo de Descomposición Empírica (EMD, de sus siglas en inglés), que produce envolventes de intensidad de los principales sonidos del corazón en el dominio del tiempo [10]. En otro estudio [11], se sugiere un método de localización para S1 y S2, basado en un algoritmo que implica el filtrado en frecuencia, detección de energía, y la duración de intervalo. La exactitud de la localización se evaluó comparando el algoritmo con el método de localización basado en transformar de Hilbert tradicional [11].

En [12], se propone un método automático para segmentación y análisis de detección de pico en patrones de Sonidos Cardíaco (HS), con especial atención a las características de las envolventes de HS y teniendo en cuenta las propiedades de la Transformada de Hilbert (HT). Con esto se aplica la Transformada de Hilbert Modificada en Tiempo Corto (STMHT), para segmentar y localizar automáticamente los puntos pico para HS mediante cruce por cero de la STMHT [12].

A través del uso de Coeficientes Cepstrales en Frecuencia Mel (MFCC), se obtienen las características más importantes de los eventos, así como la Detección de Actividad de Voz (VAD), la cual es una herramienta que puede facilitar la detección de actividad y silencio [13]. Por otra parte algunos autores proponen la extracción de características principales por medio de la Transformada Rápida de Fourier (FFT), para llevar a cabo la clasificación [14].

Los sonidos HS patológicos interfieren de manera más significativa en LS que un HS normal y son más difíciles de percibir mediante el oído. Por lo cual, aquí se propone la detección de eventos de S1, S2 en señales HS con presencia de S3 y S4. Por medio de la transformada de Hilbert que facilita y permite la detección de puntos extremos (máximos y mínimos). Además, con el apoyo de las técnicas de VAD que, a su vez

están basadas en modelos GMM se efectúa la extracción automática, lo cual se mostrará a lo largo de este trabajo.

2. Detección de S1, S2, S3 y S4

Los sonidos cardiacos se componen de dos sonidos principales S1 y S2, y en ocasiones especiales se presentan dos ruidos simultáneamente denominados ruidos cardiacos patológicos S3 y S4. El primer sonido S1 y el segundo sonido S2, son producidos por la abertura de las válvulas atrio ventriculares y el cierre de la válvula semilunar, respectivamente y viceversa. Los sonidos S3 y S4 ocurren al final de S2 debido a la vibración del flujo sanguíneo dentro de los ventrículos, el cuarto sonido S4 se encuentra justo antes de S1 debido a la contracción de la aurícula [15]. La **Tabla 1**, contiene los datos más relevantes en cuanto a duración, frecuencia y otras características de HS:

3. Métodos y materiales

En esta sección se explican las técnicas y señales empleadas para llevar a cabo la detección de eventos de S1, S2, S3 y S4 en señales HS; particularmente, al final se describe la técnica VAD-GMM que se está proponiendo en este trabajo.

3.1. Vectores acústicos MFCC, cuartiles y PCA

En MFCC, los sonidos son parametrizados, haciendo un preénfasis con filtros FIR, seguido por una ventana Hamming aplicada a cada trama [16-19]. En este trabajo, se aplicó una tasa de 120 tramas por segundo con el 50% de traslape en señales HS, a las cuales se aplica la Transformada Rápida de Fourier (FFT); posteriormente, se obtiene el módulo y se multiplica por un banco de filtros donde sus rangos de frecuencia y frecuencias centrales están distribuidos en la escala de Mel. A esto le sigue una etapa de logaritmo de la energía obtenida de cada filtro y posteriormente la transformada inversa de Fourier. El resultado final es un vector de características llamado MFCC [13, 20, 21].

El Cuartil q_p , de una variable aleatoria está definido como el número q más pequeño, tal que la función de distribución acumulativa es mayor o igual a una probabilidad p , donde p se encuentra entre $0 < p < 1$. Esto se puede definir con la función de densidad de probabilidad continua $f(x)$ a través de (1):

$$p \int_{-\infty}^q f(x) dx . \quad (1)$$

La estacionariedad está relacionada con la duración de los eventos, por lo cual el tamaño del vector se seleccionó de una duración menor al evento [14].

En el cálculo de los cuartiles, el primer paso es la lectura de la señal, partiendo de archivos *.wav; posteriormente, se aplica la FFT por cada trama. Cumpliendo con un principio básico para una función de densidad de probabilidad, la distribución espectral se normaliza como en (2):

$$F_N(f) = \int_{-\infty}^{\infty} \frac{f(t)e^{-j2\pi ft} dt}{\text{area}(F(f))}, \quad (2)$$

Un ejemplo particular de Cuantiles son los Cuartiles, calculados aquí mediante (3), cuyos valores frecuenciales $f_{0.25}, \dots, f_{0.75}$ corresponden a cada uno de los respectivos coeficientes cuartílicos [22]:

$$A_{0.25} = \int_{-\infty}^{f_{0.25}} F_N(f) df, \dots, A_{0.75} = \int_{-\infty}^{f_{0.75}} F_N(f) df, \quad (3)$$

Uno de los propósitos del análisis de componentes principales (PCA), es reducir la dimensionalidad de p a d , en donde $d < p$, y al mismo tiempo conservar la mayor cantidad posible de varianza de los datos originales [23, 24]. En este trabajo, la idea es aplicar PCA para obtener una representación disminuida en dimensión de los vectores de características, pero sin menoscabo de la eficiencia. Al utilizar PCA, se transforman los datos a un nuevo grupo de coordenadas o variables, las cuales son una combinación lineal de las variables originales. Además, las observaciones en el nuevo espacio de componentes principales no están correlacionadas. Esto se hace con el propósito de obtener información útil para comprender los datos u observaciones en el nuevo espacio [25].

3.2. Modelos mezclados Gaussianos (GMM) y ocultos de Markov (HMM)

Un modelo GMM es una tripleta Λ compuesta por las medias, covarianzas y ponderaciones. El modelado GMM se sirve del algoritmo EM para calcular las tripletas $\Lambda_i = \{m_i, \vec{\mu}_i, \Sigma_i\}$. La media $\vec{\mu}_i$, representa el promedio de todos los vectores, mientras que la matriz de covarianza Σ_i , modela la variabilidad de las características en una clase acústica [26]:

$$p(\vec{x}|\Lambda) = \sum_{i=1}^M m_i b_i(\vec{x}). \quad (4)$$

En la ecuación (4), \vec{x} es un vector MFCC o Cuartílico, $b_i, \forall i = 1, \dots, M$, son las densidades componentes y $m_i, \forall i = 1, \dots, M$ son las ponderaciones de cada densidad en el modelo. Cada densidad componente es una función Gaussiana D-dimensional [17, 20]. Cada densidad Gaussiana contiene los parámetros representados como en (5). Específicamente, aquí se calculó un modelo GMM para cada evento y/o ruido-silencio:

$$b_i(\vec{x}) = \frac{1}{(2\pi)^{D/2} |\Sigma_i|^{1/2}} \exp \left[-\frac{1}{2} (\vec{x} - \vec{\mu}_i)^T \Sigma_i^{-1} (\vec{x} - \vec{\mu}_i) \right]. \quad (5)$$

Un HMM es un autómata finito basado en estados que no son directamente observados. En esta metodología, cada estado está constituido por un GMM, el cual

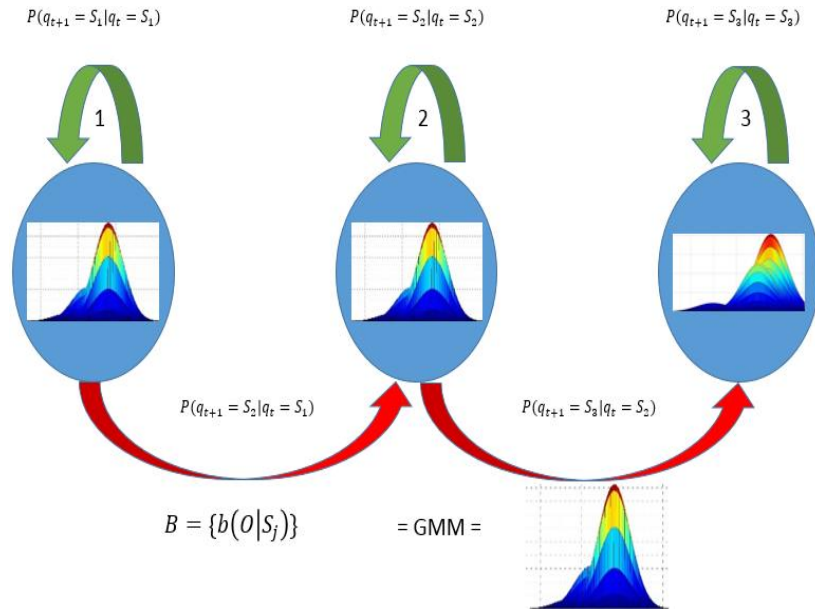


Fig. 1. Modelo HMM.

modela las observaciones correspondientes a ese estado. Un HMM, también es una tripleta $\lambda = (A, B, \pi)$, donde A es la matriz que regula las transiciones de estado a estado, B (un GMM), es la función de probabilidad que determina si un vector pertenece a un estado determinado y π da la probabilidad de iniciar en alguno de los estados. Formalmente, HMM está definido con más precisión en [27, 28], y es entrenado con el algoritmo EM. En el caso de los experimentos de este estudio, las observaciones pueden ser vectores acústicos MFCC, Cuartiles.

Al igual que en el caso de los modelos GMM, es convencional expresar los modelos HMM como tripletas (A, B, π) . Un ejemplo de modelo HMM para señales acústicas se muestra en la Fig. 1.

El entrenamiento o aprendizaje de los parámetros HMM, dado un conjunto o secuencia de observaciones $\{O_i\}$, es típicamente efectuado aplicando el algoritmo Baum-Welch [28], el cual determina los parámetros maximizando la *verosimilitud* o probabilidad $P(O_i | \lambda)$. En la etapa de evaluación, se requiere calcular $P(O | \lambda)$, dado el modelo λ y una secuencia O de observaciones; aquí se aplicó algoritmo de forward-backward [28].

La arquitectura HMM fue de tipo izquierda-derecha (*Bakis*), como lo muestra la Figura 1. Aquí, el vector π denota las probabilidades iniciales (*a priori*), de estar en alguno de los estados q ; los valores $a_{i,j}$ son las probabilidades de transición entre estados, mientras que $b_i(O)$, es la probabilidad de que la observación O (vector acústico), haya sido emitido en el estado q_i (para este caso un GMM). Se puede destacar que las probabilidades de transición y de estado inicial, fueron inicializadas aleatoriamente.

3.3. Transformada de Hilbert

La transformada de Hilbert de una señal produce un adelanto de fase en $\pi/2$, radianes. Cuando una señal es causal en un dominio, ya sea tiempo o frecuencia, la parte real y la imaginaria en el otro dominio están vinculadas por la transformada de Hilbert [29]. Se define la transformada de Hilbert como la convolución de $f(t)$, con la función $-1/\pi t$:

$$HT\{f(t)\} = f(t) * \frac{-1}{\pi t} = \frac{-1}{\pi} \int_{-\infty}^{\infty} \frac{f(\tau)}{t-\tau} d\tau. \quad (6)$$

Convolucionar en tiempo por $-1/\pi t$, es equivalente a multiplicar en frecuencia por $i \cdot \text{sign}(w)$, es decir no se modifica el espectro en amplitud, solo se efectúa un corrimiento $\pi/2$ para frecuencias positivas y de $-\pi/2$ para frecuencias negativas. Si escribimos la función exponencial compleja de la siguiente forma:

$$e^{i\omega t} = \cos \omega t + i \sin \omega t = \cos \omega t - iHT\{\cos \omega t\}. \quad (7)$$

Podemos generalizar esta idea y crear una función compleja a partir de una función real, cuya parte imaginaria tenga un retardo en fase de 90° respecto de su parte real, es decir:

$$g(t) = f(t) - iHT\{f(t)\}. \quad (8)$$

La función $g(t)$ es conocida como función analítica asociada a $f(t)$. Dada una función en tiempo cuya parte imaginaria sea igual a menos la transformada de Hilbert de su parte real, su transformada de Fourier será causal. Análogamente, si una función temporal es causal, la parte real e imaginaria de su transformada de Fourier estarán vinculadas por la transformada de Hilbert. De aquí, se define a la envolvente $E(t)$ de una función $f(t)$, como el módulo de su función analítica:

$$E(t) = |g(t)| = \sqrt{f(t)^2 + HT\{f(t)\}^2}. \quad (9)$$

3.4. VAD-GMM

Los detectores de actividad de voz (VAD), son fundamentales en el uso eficiente de ancho de banda [30], por lo cual resulta útil en señales HS y LS. Aquí se propone una versión de VAD basada en Modelos Mezclados Gaussianos (GMM), a diferencia de los modelos de VAD tradicionales, aquí se toman como referencia los sonidos cardiacos. Estos permiten detectar los segmentos activos de interés en las señales relacionadas con patologías cardiacas. Mediante VAD y cálculo de umbrales se extrajo automáticamente cada uno de los eventos de las señales HS.

En la primera etapa del experimento se utiliza un corpus de señales HS a las cuales se les aplica blanqueado, centrado y filtrado. Específicamente, se aplicó un filtro pasa bajas Butterworth de orden 7 y frecuencia de corte 150Hz, con el fin de suprimir ruido. A partir de estas señales se efectuaron recortes manuales obteniendo un nuevo corpus de señales que contienen un solo evento (S1, S2, S3 o S4 según sea el caso).

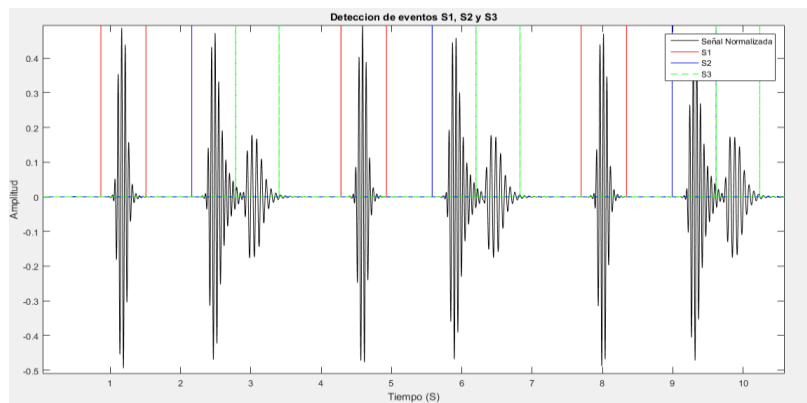


Fig. 2. Detección de eventos S1, S2 y S3 en señal Hs.

Con este nuevo corpus se calcularon los vectores acústicos para calcular modelos GMM correspondientes a cada evento, estos modelos GMM fueron utilizados para la implementación del método VAD propuesto. VAD se aplica para determinar zonas de actividad (correspondientes al evento), y zonas de no actividad (ruido o silencio).

Los vectores acústicos calculados por clase fueron MFCC (en S1, S2, S3 y S4), otra clase que se define aquí es la compuesta por ruido o silencio. Para MFCC se aplicó una tasa de 120 tramas por segundo, 13 coeficientes por vector. Posteriormente, se le aplicó PCA para obtener una representación mejorada, de la cual se utilizan únicamente los primeros 4 coeficientes.

Aplicando los modelos GMM sobre la señal se determina si el vector corresponde a actividad (asignando un 1), o a silencio (asignando un 0). Una vez que se determina la zona de actividad y no actividad esta se multiplica por la señal original obteniendo una nueva señal con solo actividad. A esta nueva señal se le extrae la envolvente por medio de la transformada de Hilbert, la cual es suavizada con un filtro Butterworth de orden 5 y frecuencia de corte de 8 Hz.

Para determinar S1, S2 y S3 o S4 es necesario calcular umbrales, los cuales son obtenidos con base en las amplitudes de los eventos correspondientes.

4. Resultados

Los resultados fueron obtenidos con particiones logradas con validación cruzada. Para medir la eficiencia de la propuesta se efectuó clasificación con los eventos con nuevos corpus de señales obtenidos por detección automática. Para lo anterior se efectuaron experimentos de clasificación aplicando Modelos Ocultos de Markov (HMM), sobre el nuevo corpus de señales. Aplicando detección automática VAD-GMM se obtiene la Fig. 2 y Fig. 3, así mismo esto arroja los índices de inicio y fin de cada evento para su extracción. Para el primer y segundo experimento de clasificación se utilizó una base de datos de 20 señales de S1, S2 y S3 a su vez 20 señales de S1, S2 y S4, las cuales fueron extraídas automáticamente con VAD-GMM.

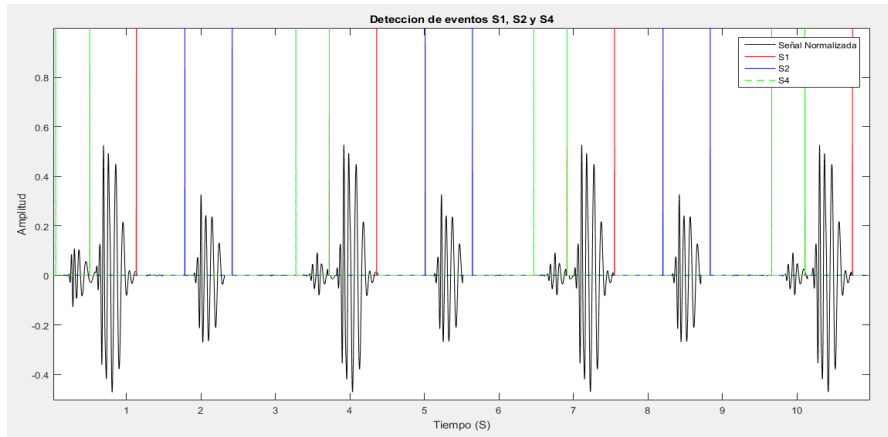


Fig. 3. Detección de eventos S1, S2 y S4 en señal Hs.

Tabla 2. Eficiencia de clasificación con VAD aplicando detección automática en señales HS con S1, S2 y S3.

# Estados	# Gaussianas	Vectores Acústicos	Eficiencia de clasificación
3	3	Cuartiles	90.4762
3	3	MFCC	95.3968
2	3	Cuartiles	85.0794
2	3	MFCC	96.9841

Tabla 3. Eficiencia de clasificación con VAD aplicando detección automática en señales HS con S1, S2 y S4.

# Estados	# Gaussianas	Vectores acústicos	Eficiencia de clasificación
3	3	Cuartiles	90.2143
3	3	MFCC	96.1111
2	3	Cuartiles	91.3456
2	3	MFCC	97.2222

Para evaluar la eficiencia en este proceso, se experimentó con varias configuraciones de arquitectura en los modelos HMM, así como 2 tipos de vectores acústicos (cuartiles y MFCC). Donde los mejores resultados de la eficiencia de clasificación se muestran en la **Tabla** , siendo para una arquitectura compuesta por 2 estados y 3 Gaussianas obteniendo hasta un 96.98% de eficiencia de clasificación.

Para el segundo experimento se utilizó el mismo conjunto de señales que en el experimento anterior (esta vez para S4), obteniendo mejores resultados de clasificación con una composición de 2 estados y 3 Gaussianas como se muestra en la **Tabla** .

Primero se obtuvo un corpus con segmentación manual de los eventos, este nuevo corpus fue de utilidad para calcular modelos GMM que fueron la base del método propuesto de VAD. La segmentación automática se logró aplicando el método de VAD propuesto, mientras que para evaluar la eficiencia se calcularon modelos HMM empleando el corpus obtenido con el VAD propuesto. Aunque aquí no se muestran resultados con segmentación manual, los resultados obtenidos con segmentación automática indican que puede sustituir la segmentación manual realizada por una persona. Además, este proceso resulta más objetivo y menos dependiente de las capacidades auditivas y visuales del profesional de la salud.

5. Conclusión

Para el mejoramiento de la detección automática de eventos en señales HS se propuso la detección aplicando transformada de Hilbert combinada con modelos GMM constituyendo una técnica de VAD.

Por medio de detección automática se obtuvo un corpus de 20 señales (para cada uno de los eventos), de S1, S2, S3 y S1, S2 y S4. Se efectuó la clasificación de señales aplicando dos corpus, uno obtenido mediante recortes manuales, y otro mediante la detección automática. La detección automática supero los resultados en clasificación con respecto al corpus obtenido por detección (recortes), manuales. Se llevó a cabo la clasificación por medio de HMM obteniendo hasta un 96.9% de eficiencia para los sonidos de S3 y un 97% para S4.

El diagnóstico mediante detección automática promete ser un método más eficaz y seguro que métodos tradicionales, los cuales quedan limitados por la capacidad auditiva de un médico. Resulta interesante y necesario extender estas metodologías para sonidos del pulmón (LS), y hacer diagnósticos cardiopulmonares.

Referencias

1. Azarbarzin, A., Moussavi, Z. M. K.: Automatic and Unsupervised Snore Sound Extraction From Respiratory Sound Signals. *IEEE Transactions on Biomedical Engineering*, 58, pp. 1156–1162 (2011)
2. Earis, J.: Lung sounds. *Thorax*, 47, pp. 671–672 (1992)
3. NOM-016-SSA3-2012: Características mínimas de infraestructura y equipamiento de hospitales y consultorios de atención médica especializada. *Diario Oficial* (2013)
4. Forgacs, P.: Lung sounds. *Br J Dis Chest*, 63, pp. 1–12 (1969)
5. Lozano, M., Fiz, J. A., Jan, R.: Automatic Differentiation of Normal and Continuous Adventitious Respiratory Sounds Using Ensemble Empirical Mode Decomposition and Instantaneous Frequency. *IEEE Journal of Biomedical and Health Informatics*, 20, pp. 486–497 (2016)
6. Kosasih, K., Abeyratne, U. R., Swarnkar, V., Triasih, R.: Wavelet Augmented Cough Analysis for Rapid Childhood Pneumonia Diagnosis. *IEEE Transactions on Biomedical Engineering*, 62, pp. 1185–1194 (2015)

7. Herzig, J., Bickel, A., Eitan, A., Intrator, N.: Monitoring Cardiac Stress Using Features Extracted From S1 Heart Sounds. *IEEE Transactions on Biomedical Engineering*, 62, pp. 1169–1178 (2015)
8. Emmanouilidou, D., McCollum, E. D., Park, D. E., Elhilali, M.: Adaptive Noise Suppression of Pediatric Lung Auscultations With Real Applications to Noisy Clinical Settings in Developing Countries. *IEEE Transactions on Biomedical Engineering*, 62, pp. 2279–2288 (2015)
9. Chen, Jie., Hou, Hai-Liang., Luo, Liang-Cai., Yun, C.: Automatic Identification Method for the First and Second Heart Sound Based on Double-threshold. *Journal Computer Engineering*, 38, pp. 174–177 (2012)
10. Bajelani, K., Navidbakhsh, M., Behnam, H., Doyle, J. D., Hassani, K.: Detection and identification of first and second heart sounds using empirical mode decomposition. *Journal of Engineering in Medicine*, 227, pp. 976–987 (2013)
11. Dong, M. S., Hangsik, S.: A Localization Method for First and Second Heart Sounds Based on Energy Detection and Interval Regulation. *Journal of Electrical Engineering and Technology*, 10, pp. 2126–2134 (2015)
12. Sun, S., Jiang, Z., Wang, H., Fang, Y.: Automatic moment segmentation and peak detection analysis of heart sound pattern via short-time modified Hilbert transform. *Computing Methods Prog. Biomed*, 114, pp. 219–230 (2014)
13. Abushakra, A., Faezipour, M.: Acoustic Signal Classification of Breathing Movements to Virtually Aid Breath Regulation. *IEEE Journal of Biomedical and Health Informatics*, 17(2), pp. 493–500, DOI:10.1109/JBHI.2013.2244901 (2013)
14. Devangan, H., Jain, N.: A Review on Classification of Adventitious Lung Sounds. *International Journal of Engineering Research & Technology*, 4 (2015)
15. Ashok, M., Parthasarathi, B., Goutam, S.: An automated tool for localization of heart sound components S1, S2, S3 and S4 in pulmonary sounds using Hilbert transform and Heron's formula. Department of Electronics and Electrical Communication Engineering, Indian Institute of Technology, Kharagpur, Kharagpur, 721, pp. 302, DOI:10.1186/2193-1801-2-512 (2013)
16. Istrate, D. M.: *Detection et Reconnaissance des Sons pour la Surveillance Médicale*. These pour obtenir le grade de docteur de l'INPG: spécialité Signal, Image, Parole, Télécoms docteur Institut National Polytechnique de Grenoble (2003)
17. Mayorga, P., Druzgalski, C., Vidales, J.: Quantitative Models for Assessment of Respiratory Diseases. *Pan American Health Care Exchange (PAHCE)*, pp. 25–30 (2010)
18. Boston Children's Hospital: <http://www.childrenshospital.org/>
19. Pearce, D.: Developing the ETSI Aurora advanced distributed speech recognition front-end and what next?. *Automatic Speech Recognition and Understanding, (ASRU'01)*, IEEE Workshop, pp. 131–134 (2001)
20. Mayorga, P., Besacier, L., Lamy, R., Serignat, J. F.: Audio packet loss over IP and speech recognition. *IEEE Workshop on Automatic Speech Recognition and Understanding*, pp. 607–612 (2003)
21. Yoon, Jae Sam, Gil Ho Lee, Hong Kook Kim: A MFCC-Based CELP Speech Coder for Server-Based Speech Recognition in Network Environments. *IEICE Trans. Fundam. Electron. Commun. Comput. Sci*, E90-A, pp. 626–632 (2007)

22. Mayorga, P., Druzgalski, C., González, O. H., Lopez, O. H.: Modified classification of normal Lung Sounds applying Quantile Vectors. Engineering in Medicine and Biology Society (EMBC), Annual International Conference of the IEEE, pp. 4262–4265 (2012)
23. Martinez, W. L.: Computational Statistics Handbook with Matlab. Second ed. (2008)
24. Wendy, L., Martinez, A., Martinez, R., Solka, J. L.: Exploratory Data Analysis with MATLAB. Second Edition, (2011)
25. Mayorga-Ortiz, P., Druzgalski, C., Criollo-Arellano, M. A., González-Arriaga, O. H.: GMM y LDA aplicado a la detección de enfermedades pulmonares. Revista mexicana de ingeniería biomédica, 34, pp. 131–144 (2013)
26. Bimbot, F., Bonastre, J.-F., Fredouille, C., Gravier, G., Magrin-Chagnolleau, I., Meignier, S., et al.: A Tutorial on Text-Independent Speaker Verification. EURASIP Journal on Advances in Signal Processing, pp. 1–22 (2004)
27. Kannadaguli, P., Bhat, V.: A comparison of Gaussian Mixture Modeling (GMM) and Hidden Markov Modeling (HMM) based approaches for Automatic Phoneme Recognition in Kannada. Signal Processing and Communication (ICSC), International Conference on, pp. 257–260 (2005)
28. Rabiner, L. R., Juang, B. H.: Fundamentals of speech recognition. Englewood Cliffs, N.J., PTR Prentice Hall, (1993)
29. Mondal, A., Kumar, A. K., Bhattacharya, P. S., Saha, G.: Boundary estimation of cardiac events S1 and S2 based on Hilbert transform and adaptive thresholding approach. Medical Informatics and Telemedicine (ICMIT), Indian Conference on, pp. 43–47 (2013)
30. Jongseo-Sohn, N. S. K., Wonyong, S.: A Statistical Model-Based Voice Activity Detection. IEEE Signal Processing Letters, 6 (1999)