

Detección de sonidos torácicos aplicando detección de actividad de voz

Julio A. Valdez¹, Pedro Mayorga¹, Gilberto Chavez¹, Christopher Druzgalski²,
Vesna Zeljkovic³

¹ Tecnológico Nacional de México/I.T. Mexicali, Mexicali, México

² California State University, Electrical Engineering Department, California, USA

³ The Lincoln University, School of Science and Technology, Pennsylvania, USA
julito_valdez@hotmail.com

Abstract. La evaluación de los sonidos pulmonares (LS) y los sonidos cardíacos (HS) y la documentación de sus características forman parte de los procedimientos rutinarios de diagnóstico cardiopulmonar. Los componentes acústicos peculiares asociados con las fases inspiratoria, espiratoria, S1 y S2 de la función cardiopulmonar, como por ejemplo las sibilancias asmáticas, sirven como indicadores de diagnóstico útiles. Complementar la auscultación tradicional con la detección automática y la categorización de las señales características proporciona oportunidades de monitoreo mejorado a largo plazo que es indispensable en la EPOC u otras anomalías de origen cardíaco. Además, debido a la edad u otras limitaciones sensoriales de un profesional de la salud, la detección automática de eventos de LS-HS podría mejorar significativamente estos aspectos del diagnóstico médico. En particular, el método novedoso propuesto se basa en los elementos de Detección de Actividad de Voz (VAD), asociados con componentes de frecuencia LS-HS similares, y la utilización de Modelos Mezclados Gaussianos (GMM). La detección automática con GMM está acompañada por los Coeficientes Cepstrales de Frecuencia de Mel (MFCC) y la transformada de Hilbert. Esta técnica permite detectar y clasificar eventos peculiares de LS y HS, y obtener tasas de clasificación más altas en comparación con otras formas de VAD. La eficiencia de clasificación obtenida con Modelos Ocultos de Markov (HMM) fue del 95% para LS y del 92% para HS, documentando su viabilidad clínica.

Palabras clave: Diagnóstico cardiopulmonar, Modelos Mezclados Gaussianos (GMM), Modelos Ocultos de Markov (HMM), transformada de Hilbert, Detección de Actividad de Voz (VAD).

Thoracic Sound Detection Using Voice Activity Detection

Abstract. The lung and heart sounds (LS, HS) assessment and characteristics documentation are part in common procedures to cardiopulmonary diagnostic. Peculiar acoustic components associated to inspiratory and expiratory phases, S1 and S2 of cardiopulmonary function, as in sibilance are useful indicators in diagnostics for asthma. To complete traditional auscultation with automatic detection and categorization signals, give better opportunity of monitoring to long term, even this is indispensable for Chronic Obstructive Pulmonary Disease (COPD) and other anomalies of cardiac origin. Moreover, due to sensorial limitations and age of health professionals, automatic detection could significantly

improve these medical aspects. Particularly, the novel method proposed here, is based on Voice Activity Detections principles (VAD), but they are associated with LS-HS frequency components, which are similarities with voice, and strengthened with Gaussian Mixture Models (GMM). Automatic detection is here reinforced with Mel Frequency Cepstral Coefficients (MFCC) and Hilbert transform. This technique allowed to detect and classify peculiar events in LS and HS, and to obtain higher classification rates compared with other VAD techniques. The classification efficiency reached with Hidden Markov Models (HMM) were 95% for LS and 92% for HS, evidencing clinical viability for our proposed method.

Keywords: Cardiopulmonary diagnosis, Gaussian Mixed Models (GMM), Hidden Markov Models (HMM), Hilbert transform, Voice Activity Detection (VAD).

1. Introducción

La auscultación representa uno de los procedimientos de diagnóstico más comunes y bien establecidos [1]. Sin embargo, todavía se continúa confiando en criterios descriptivos subjetivos más que en la caracterización cuantitativa. Los sonidos cardíacos y pulmonares proporcionan información de diagnóstico muy útil en particular cuando su origen específico se extrae de forma independiente de los sonidos torácicos generales registrados.

Esta es una tarea desafiante, ya que los componentes de frecuencia del corazón y los sonidos pulmonares se superponen, aunque obviamente, sus características periódicas son distintas debido a las diferencias en las frecuencias cardíaca y respiratoria. El S1 y el S2, como dos componentes principales de los sonidos cardíacos, suelen incluir componentes dominantes de frecuencia por debajo de 150 Hz y se pueden correlacionar con el Electrocardiograma (ECG), aunque por razones prácticas rara vez se realiza en un entorno clínico. Los sonidos pulmonares que reflejan las fases inspiratoria y espiratoria incluyen también componentes de mayor frecuencia especialmente cuando, por ejemplo, hay sibilancias. Por las mismas razones prácticas, se registran como parte de la auscultación general, ya que se usa un estetoscopio electrónico como unidad independiente. Por lo tanto, para minimizar el efecto de correlación en el análisis computacional, una separación de los sonidos del corazón de los sonidos del pulmón es útil [1]. En ambos casos, un algoritmo eficiente también puede reducir el ruido.

Diferentes autores [2-6] proponen diversos métodos de extracción de características de LS y HS, que permiten extraer los componentes primordiales de los sonidos cardíacos y respiratorios. Algunos de estos métodos incluyen el Análisis Espectral Singular (SSA) para separar los componentes importantes HS de las señales LS [2]. La extracción HS se mejoró, mediante la localización de los componentes en función de la envolvente de la señal aplicando la transformada de Hilbert. Además, dado que en la envolvente de señal se puede usar la fórmula de Herón para obtener las componentes S1 y S2, por lo cual se puede obtener una mejor eficiencia en detección que con SSA [6]. Debido a que los eventos LS están asociados con características relacionadas con la duración, frecuencia y amplitud, algunos autores [3, 4], aplicaron métodos de Detección de Actividad de Voz (VAD) y Coeficientes Cepstrales de Frecuencias Mel (MFCC),

que podrían ser útiles para extraer eventos importantes de LS. Otros autores han propuesto explorar la realidad virtual, monitorear y detectar la inhalación y la exhalación [4]. En general, los métodos tienen como inconveniente el ruido (ruido ambiental, ruido cardíaco o ruido pulmonar dependiendo del interés en el origen particular de los sonidos). Además, las características principales de extracción todavía se ven obstaculizadas por las condiciones de registro de las señales [3]. Debido a que la detección de actividad de señales LS es bastante compleja, aquí presentamos una técnica centrada en un método novedoso que abarca también la detección automática de eventos acústicos relacionados con la inhalación y la exhalación basados en modelos mezclados gaussianos (GMM). Los modelos VAD normalmente son basados en principios de energía y cruce por cero, pero no en nuestro caso.

La sección 2 explica el conjunto de técnicas aplicadas para llevar a cabo los experimentos, así como los eventos que componen el sonido y la base de datos utilizada. La sección 3 muestra los experimentos llevados a cabo, incluyendo figuras y tablas obtenidas a partir de los resultados. Finalmente, la sección 4 presenta la conclusión de los experimentos y los resultados.

2. Metodología

Esta sección proporciona los fundamentos para la representación, el preprocesamiento y el modelado de eventos que abarcan LS en nuestros experimentos.

2.1. Vectores de Coeficientes Cepstrales en Frecuencia Mel (MFCC)

En MFCC, los sonidos se parametrizan implementando un énfasis previo con filtros FIR, seguidos de una ventana de Hamming aplicada a cada cuadro analizado [7-10]. En este proyecto, los experimentos se llevaron a cabo utilizando ventanas de Hamming de 50 ms (LS) y 130 ms (HS) con un 50% de traslape en ambos casos, a las que se aplicó la Transformada Rápida de Fourier (FFT); posteriormente, se obtuvo el módulo y luego se multiplicó por un banco de filtros cuyo rango de frecuencias y frecuencias centrales se distribuyeron por la escala Mel o Bark. Esto fue seguido por una etapa de cálculo de los valores obtenidos de las salidas de cada filtro y, posteriormente, la Transformada de Fourier Inversa. Como resultado, se obtuvo un vector de características llamado MFCC [4, 11, 12].

2.2. Vectores cuantílicos

El cuantil q_p de una variable aleatoria se define como el valor q más pequeño de tal forma que la función de distribución acumulativa sea mayor o igual que la probabilidad p , donde p es $0 < p < 1$. Esto se puede calcular a partir de una función de densidad de probabilidad continua (PDF) $f(x)$ a través de (1):

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

La duración de la fase de inhalación (~ 1.5 s.) y la fase de exhalación (~ 2.5 s.) para la mayoría de las señales LS utilizadas proviene de [13]. En este proceso, la primera etapa consiste en leer la señal, comenzando desde un archivo *.wav; posteriormente,

se aplica la FFT. Cumpliendo un principio básico para una función de distribución de probabilidad, la distribución espectral se tiene que normalizar (2):

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

Un ejemplo de cuantil es el uso de cuartiles, calculado por (3), donde cada valor de frecuencia $f_{0.25}, \dots, f_{0.75}$ corresponde a su coeficiente cuartil respectivo [13]:

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

2.3. Modelos Mezclados Gaussianos (GMM)

El modelo GMM es una tripleta compuesta por las medias, las covarianzas y las ponderaciones de cada densidad. El modelado GMM usa el algoritmo de maximización de la esperanza (EM) para entrenar los modelos $\Lambda_i = \{m_i, \vec{\mu}, \Sigma_i\}$. El promedio $\vec{\mu}$ representa el promedio de todos los vectores, mientras que la matriz de covarianza Σ_i modela la variabilidad de las características de una clase acústica [14]:

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

En la ecuación (4), \vec{x} es un MFCC o vector cuartil, $b_i, \forall_i = 1, \dots, M$ son las densidades componentes y $m_i, \forall_i = 1, \dots, M$ son las ponderaciones de cada densidad participante en el modelo. Cada densidad componente es una función Gaussiana D-dimensional [8, 11]. Cada densidad gaussiana contiene los parámetros representados en la Ecuación (5):

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

2.4. Detección de Actividad de Voz (VAD)

Los detectores de actividad de voz (VAD) son fundamentales para el uso eficiente del ancho de banda [3]. Han sido útiles en la transmisión y el procesamiento de audio digital, y ahora en señales LS-HS. En [3] VAD se aplica a una señal aleatoria con ruido de fondo estacionario, su duración es mayor que un fonema, y los eventos en LS tienen una duración promedio mayor que un fonema. Es por este motivo, que aquí se propone una versión de VAD basada en modelos de mezclados gaussianos (GMM). En primer lugar, un segmento de inhalación y exhalación asignado manualmente en muestras seleccionadas de señales LS, y se guardan en un formato wav, además, las zonas de silencio son segmentadas manualmente (que tienen ruido) y se guardan en formato wav. El siguiente paso es calcular los modelos GMM correspondiente a la inhalación y la exhalación. De la misma manera, se calcula otro modelo GMM para zonas de silencio-ruido. Con estos dos modelos se constituye el sistema VAD, que se aplica a todas las señales. Los vectores MFCC se calculan sobre las señales completas, y con VAD propuesto se detecta si cada MFCC corresponde a actividad (en este caso, inhalación o exhalación) o a una región de no actividad (silencio o ruido). Es importante tener en

cuenta que las señales y sus cortes se convirtieron en vectores MFCC para el cálculo de los modelos VAD-GMM. Eventualmente, los vectores MFCC se usarán para distinguir entre las fases de inhalación y exhalación, ya que la media de las sextas componentes de los vectores MFCC corresponde a un umbral para distinguir entre los segmentos de inhalación y espiración [4].

2.5. Análisis de Componentes Principales (PCA)

Uno de los propósitos del Análisis de Componentes Principales (PCA) es reducir la dimensionalidad de p a d , donde $d < p$, y al mismo tiempo preservar la mayor cantidad posible de varianza de los datos originales [15, 16]. En este proyecto, la idea es aplicar PCA para obtener una representación disminuida en la dimensión del vector de características, pero sin disminuir la eficiencia. Para usar PCA, se transforman los datos a un nuevo grupo de coordenadas o variables, que son una combinación lineal de las variables originales. Además, las observaciones en el nuevo espacio de componentes principales no están correlacionadas [17].

2.6. Modelos Ocultos de Markov

El ciclo cardíaco es acústicamente secuencial, normalmente consta de dos eventos principales conocidos como el primer sonido cardíaco, "S1" y el segundo sonido cardíaco "S2". La aparición de los sonidos pulmonares (LS) es también un proceso cíclico formado por dos eventos principales, la inhalación y la exhalación. En ambos casos, las señales HS y LS, tienen silencios entre sus eventos principales, que son diferentes según el evento que tenga lugar antes y después.

Por lo tanto, ambos casos son secuencias de eventos que pueden variar según las circunstancias, las condiciones de salud e incluso el estado de ánimo de la persona. En el análisis de HS, S1 y S2, así como los silencios, son las etapas de la señal a modelar, mientras que en LS, la inhalación, la exhalación y los silencios son las etapas de la señal a modelar. Luego, estas se modelan a través de los Modelos Ocultos de Markov (HMM), que es un autómata de estado finito.

Un HMM puede considerarse como un modelo basado en estados, en el que cada estado no se observa directamente, de hecho, cada estado se caracteriza por un GMM. Formalmente, se explica exhaustivamente en [18]. Como en el caso de los modelos GMM, los HMM se expresan convencionalmente por una tripleta $\lambda = (A, B, \pi)$, como se observa en la Fig. 1.

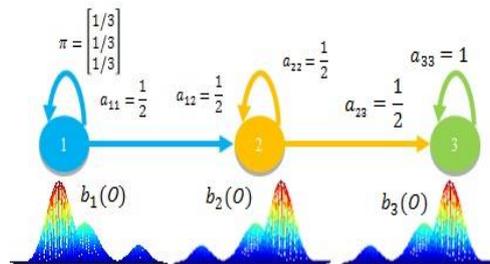


Fig. 1. Modelo HMM con $\lambda = (A, B, \pi)$.

Donde A es una matriz que contiene la probabilidad de transitar de un estado a otro, B (en nuestro caso GMM) es la probabilidad de que un vector acústico (MFCC o vector cuartil) se haya generado a partir de un estado, y π es la probabilidad de iniciar en algún estado. El entrenamiento de los parámetros HMM, dado un conjunto de secuencia de observación (vectores acústicos) $\{O_i\}$, normalmente se hace aplicando el algoritmo Baum-Welch [18], que obtiene los parámetros maximizando la probabilidad $P(O_i | \lambda)$. Para la etapa de evaluación, se requiere calcular $P(O | \lambda)$, dado el modelo λ y una secuencia O de observaciones; aquí se aplicó el algoritmo de avance-retroceso (normalmente conocido en inglés como *forward-backward*) [18].

2.7. Transformada de Hilbert

La transformada de Hilbert de una señal de tiempo continuo con valor real $y(t)$, se define como:

$$H\{y(t)\} = \hat{y}H(s) = \frac{1}{\pi} \oint_{-\infty}^{\infty} \frac{y(t)}{t-s} dt, \quad (6)$$

donde s es real y $H\{\cdot\}$ representa al operador de Hilbert. Aquí, la integración debe llevarse a cabo de acuerdo con el principio de Cauchy, es decir:

$$\oint_{-\infty}^{\infty} \frac{y(t)}{t-s} dt = \lim_{\epsilon \rightarrow 0} \left(\int_{-\infty}^{s-\epsilon} + \int_{s+\epsilon}^{\infty} \right) \frac{y(t)}{t-s} dt, \quad \epsilon > 0. \quad (7)$$

Sin embargo, las señales en el mundo real son discontinuas y deben ser expresadas como señales de tiempo discreto [6].

2.8. Detección de S1 y S2

Inicialmente, las señales de HS se escalaron en el rango ± 1 , luego se aplicó un filtro pasa bajo Butterworth con una frecuencia de corte de 150 Hz y de orden 10. Después de eso, las zonas S1 y S2 se segmentaron manualmente, y estas secciones se guardaron en formato wav. Del mismo modo, las zonas de silencio (que tenían ruido) se segmentaron, y estas secciones se guardaron en formato wav. El siguiente paso fue calcular un modelo GMM para señales S1 y S2, al mismo tiempo que se calculaba un modelo GMM para zonas de silencio, pero previamente todas las señales se transformaron en vectores de características. Se aplicó PCA sobre todos los vectores obtenidos, y se determinó visualmente que en el caso de MFCC solo los primeros 4 componentes eran necesarios. Los dos modelos constituían el sistema VAD, que se aplicaba a todas las señales para detectar y segmentar automáticamente S1, S2 y las zonas de silencio. Después de la detección automática de S1 y S2, se calculó la envolvente de Hilbert de la señal, donde se comparó la envolvente con la señal original normalizada. A continuación, la envolvente de Hilbert se suavizó con un filtro de Butterworth de quinto orden, con frecuencia de corte de 7-25 Hz (después de varios experimentos se seleccionó 8 Hz como la mejor opción) [6]. VAD detecta S1 y S2, pero sin diferenciarlos entre sí. Para diferenciar S1 de S2 es necesario calcular umbrales. Los umbrales se calculan para discriminar amplitudes y establecer que picos corresponden a S1 o S2. Esto se logra aplicando mínimos y máximos, considerando un mínimo como el inicio de un pico y otro mínimo como el final del mismo pico. Es importante distinguir entre los segmentos sistólico y diastólico,

por eso S1 y S2 podrían identificarse, y poder determinar cuál fue el primero en el registro de la señal. El propósito es separar la sístole y la diástole en pares o señales impares, cuando la señal comienza con S1, la señal corresponde a la sístole, mientras que en las señales pares corresponden a la diástole; si la señal comienza con S2, las señales pares corresponden a la sístole y las impares corresponden a la diástole. A partir de estas señales, es posible obtener las duraciones de tiempo (ancho) de S1 y S2.

2.9. Base de datos

RALE consiste en un conjunto de sonidos LS normales y adventicios como grabaciones en formato *.wav, que se desarrolló en Winnipeg, Canadá [19]. Las señales de RALE se filtraron con paso de banda de 7.5 Hz para suprimir cualquier componente de DC mediante un filtro Butterworth de primer orden. Además, se aplicó un filtro pasa baja Butterworth de octavo orden a 2.5 kHz para evitar traslape. Las señales de la base de datos se muestrearon a 11025 Hz. Las señales de LS Normal de RALE fueron segmentadas manualmente (por los autores) para obtener solo los segmentos de inhalación-exhalación de cada señal, haciendo un total de 13 inhalaciones y 13 exhalaciones.

Las señales HS utilizadas provienen de bases de datos que están disponibles para fines académicos y científicos [9, 20]. Por lo tanto, se usaron 13 señales normales, que se segmentaron manualmente, obteniendo 13 sonidos de S1 y 13 de S2. La frecuencia de muestreo fue de 11025 Hz, con formato *.wav y mono. Las señales utilizadas se dividieron en conjunto de entrenamiento y conjunto de evaluación. La evaluación se realizó con validación cruzada, aquí se deja una señal para evaluar, y se usan las señales restantes para calcular el modelo; este proceso es repetitivo cambiando las configuraciones hasta agotar todas las posibilidades.

Utilizando datos experimentales, se encontró que los intervalos de longitud de tiempo de las fases de señal del repositorio de datos RALE eran aproximadamente de 1.5 segundos para la inhalación y de 2.5 segundos para la exhalación. Con respecto al HS, la fase S1 dura aproximadamente 0.1 a 0.12 segundos; La fase S2 está entre 0.8 y 0.14 segundos [21].

2.10. Clasificación

La evaluación se llevó a cabo con validación cruzada, donde cada señal de inhalación y exhalación es leída y luego procesada con PCA. Debido a que se aplica validación cruzada, los conjuntos de señales de inhalación y exhalación se usan para calcular modelos por clase y su evaluación. Sin embargo, debe tenerse en cuenta que, en la validación cruzada, durante cada prueba se utiliza una señal del conjunto n ($n = 13$) para la evaluación, mientras que las señales $n-1$ restantes se utilizan para calcular el modelo. Se realizan n evaluaciones, pero en cada evaluación, se cambia la señal de prueba y las $n-1$ señales restantes se usan para calcular el modelo.

3. Resultados

Una vez aplicadas las metodologías descritas anteriormente, es importante verificar si son realmente efectivas, por lo cual se llevaron a cabo experimentos de clasificación

aplicando Modelos Ocultos de Markov (HMM). La detección automática con VAD-GMM arroja un gráfico que se muestra en la Fig. 2, al mismo tiempo esto nos da el índice inicial y final de los eventos extraídos de estos datos. Para el primer experimento de clasificación se utilizó una base de datos con 13 señales de inhalación y 13 señales de exhalación.

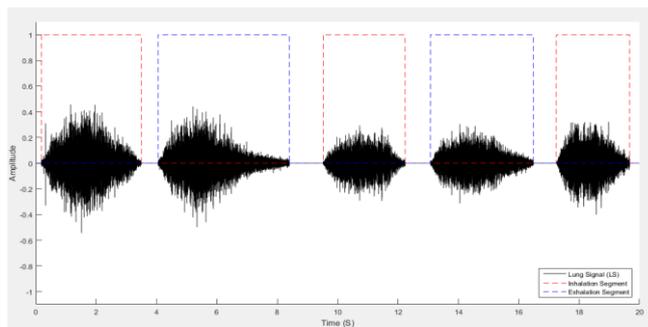


Fig. 2. Segmentación automática de LS con algoritmo VAD-GMM.

Para evaluar la eficiencia de este proceso, se experimentó con diferentes configuraciones de arquitectura de modelos HMM, así como dos tipos de vectores acústicos (cuartiles y MFCC). Los resultados de la eficiencia de clasificación se muestran en la Tabla 1, donde solo se muestran los mejores resultados.

Tabla 1. Eficiencia de clasificación aplicando VAD [4] y detección automática en LS.

Núm. de Estados	Núm. de Gaussianas	Vectores Acústicos	Eficiencia de Clasificación
3	3	Cuartiles	77.5%
3	3	MFCC	76.25%
2	3	Cuartiles	75%
2	3	MFCC	70%

Para el segundo experimento, se utilizaron los mismos conjuntos de señales LS que el experimento anterior, pero en este caso se aplicó VAD-GMM propuesto, como se muestra en la Tabla 2.

Tabla 2. Eficiencia de clasificación con detección automática en LS aplicando VAD-GMM.

Núm. de Estados	Núm. de Gaussianas	Vectores Acústicos	Eficiencia de Clasificación
3	3	Cuartiles	85.625%
3	3	MFCC	95%
2	3	Cuartiles	88.125%
2	3	MFCC	91.25%

En el tercer experimento, se utilizó un conjunto de señales S1 y S2 obtenidas por detección automática de HS (Fig. 3). Así como en la Tabla 2 se aplicó VAD-GMM; el mejor resultado de clasificación fue 92.6223% como se muestra en la Tabla 3. Como se puede ver en la Tabla 3, ambos vectores acústicos se usaron para calcular modelos,

con dos configuraciones de arquitecturas HMM-GMM y los mejores resultados se obtuvieron con 3 estados y 3 gaussianos por estado.

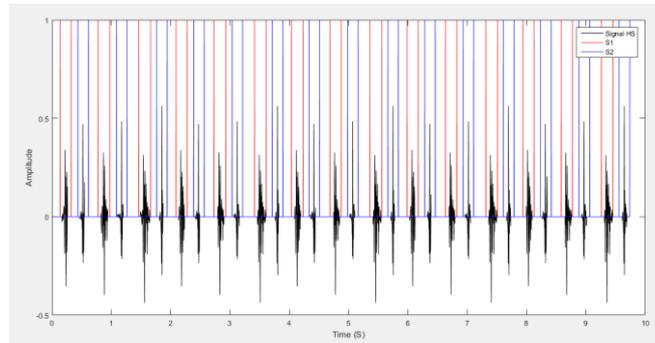


Fig. 3. Segmentación automática de señales HS con el algoritmo VAD-GMM.

Tabla 3. Eficiencia de clasificación con detección automática en HS y con VAD-GMM.

Núm. de Estados	Núm. de Gaussianas	Vectores Acústicos	Eficiencia de Clasificación
3	3	Cuartiles	92.6923
3	3	MFCC	90.3846%
2	3	Cuartiles	87.3077%
2	3	MFCC	91.25%

Como se observa en la Tabla 1, el mejor resultado de clasificación se obtuvo con cuartiles, es 77.5%, mientras que con VAD-GMM se logró 95% como se muestra en la Tabla 2, y esto demuestra la superioridad de VAD-GMM. Incluso en los eventos de HS (S1 y S2) la detección y extracción con VAD-GMM tuvieron un buen rendimiento, lo que significa que VAD-GMM puede determinar a qué evento pertenece un vector, sin importar el tipo de vector (cuartil o MFCC). La utilización de GMM mejora la capacidad de VAD para asociar un vector acústico con su clase. Esto es razonable ya que en los esquemas de reconocimiento de locutor, la identificación de la clase a la cual pertenece un vector acústico, utilizan sistemas VAD. Además de la detección y extracción automáticas, VAD-GMM podría proporcionar un registro documentado para el monitoreo a largo plazo y el análisis comparativo.

4. Conclusiones

La detección automática de eventos acústicos en LS aplicando VAD sería útil en diagnóstico respiratorio, en particular para documentar eventos y monitoreo a largo plazo. Debido al aumento en los casos de EPOC y que no se conoce una cura, un mejor manejo en datos a largo plazo y con medicamentos selectivos aumentaría la calidad general de atención médica.

Los resultados obtenidos son prometedores, ya que la eficiencia de clasificación aumentó aplicando detección automática con VAD-GMM. Con el primer algoritmo de

VAD obtenido de la literatura, se logró el 77.5% de la eficiencia, mientras que con VAD-GMM propuesto se alcanzó un 95% de eficiencia.

Por lo tanto, nuestra propuesta podría mejorar la auscultación tradicional cualitativa apoyada en la evaluación cuantitativa y al diagnóstico automatizado asistido mediante la identificación de indicadores acústicos anormales. Además, este método podría usarse en medicina general.

Referencias

1. Fatma Ayaria, M. K., Alouanib, A.T.: Lung sound extraction from mixed lung and heart sounds FASTICA algorithm. Electrical Engineering Department, National School of Engineers Tunis (ENIT), Electrical Engineering Department, Tennessee Technological University (2012)
2. Ghaderi, F., Mohseni, H. R., Sanei, S.: Localizing Heart Sounds in Respiratory Signals Using Singular Spectrum Analysis. Biomedical Engineering, IEEE Transactions on, vol. 58, pp. 3360–3367 (2011)
3. Sohn, J., Sung, W.: A Statistical Model-Based Voice Activity Detection. IEEE signal processing letters, vol. 6 (1999)
4. Faezipour, A. A. a. M.: Acoustic Signal Classification of Breathing Movements to Virtually Aid Breath Regulation. IEEE journal of biomedical and health informatics, vol. 17 (2013)
5. Wang, Y. X. H. , Li, M.: Study on the MFCC similarity-based voice activity detection algorithm. In: Proc. 2nd IEEE Int. Conf. Artif. Intell. Manage. Sci. Electron. Commerce, pp. 4391–4394 (2011)
6. Ashok Monda, P. B. a. G. 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, India (2013)
7. 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, Grenoble, France (2003)
8. Mayorga, P., Druzgalski, C., Vidales, J.: Quantitative Models for Assessment of Respiratory Diseases. Pan-American Health Care Exchange (PAHCE), pp. 25–30 (2010)
9. Boston Children's Hospital [Online]. Available: <http://www.childrenshospital.org/>
10. Pearce, D.: Developing the ETSI Aurora advanced distributed speech recognition front-end and what next? In: Automatic Speech Recognition and Understanding, ASRU '01. IEEE Workshop on, pp. 131–134 (2001)
11. Mayorga, P., Besacier, L., Lamy, R., Serignat, J. F.: Audio packet loss over IP and speech recognition. In: Automatic Speech Recognition and Understanding, ASRU '03, IEEE Workshop on, pp. 607–612 (2003)
12. Proakis, J. G., Salehi, M.: Digital Communications. 5th ed. (2008)
13. Mayorga, P., Druzgalski, C., González, O. H., Lopez, H. S.: Modified classification of normal Lung Sounds applying Quantile Vectors. In: Engineering in Medicine and Biology Society (EMBC), Annual International Conference of the IEEE, pp. 4262–4265 (2012)
14. 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, vol. 2004, pp. 1–22 (2004)
15. Martínez W. L., Martínez A. R.: Computational Statistics Handbook with Matlab. Second ed., (2008)
16. Martínez W. L., Martínez A. R., Solka, J. L.: Exploratory Data Analysis with MATLAB. Second Edition (2011)

17. 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*, vol. 34, pp. 131–144 (2013)
18. Rabiner, L. R., Juang, B. H.: *Fundamentals of speech recognition*. Englewood Cliffs, N.J.: PTR Prentice Hall (1993)
19. RALE, Lung Sounds Repository. [Online]. Available: <http://www.rale.ca/Repository.htm>
20. Texas Heart Institute. [Online]. Available: <http://www.texasheartinstitute.org/AboutUs/index.cfm>
21. Echeverry, J. D. , López, A. F., López, J. F.: Reconocimiento de valvulopatías cardíacas en señales de fonocardiografía empleando la transformada Gabor. *Scientia Et Technica*, vol. XIII, pp. 139–144 (2007)