34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

Modified Classification of Normal Lung Sounds applying Quantile Vectors

Mayorga P.[1], Druzgalski C.[2], González O. H.[1], López H.S.[1]
[1]Instituto Tecnológico de Mexicali, Mexicali, B.C., México
[2]Electrical Engineering, California State University, Long Beach, CA, USA

Abstract—**In this paper a novel Lung Sound Automatic Verification (LSAV) system and front-end Quantile based acoustic models to classify Lung Sounds (LS) are proposed. The utilization of Quantiles allowed an easier and objective assessment with smaller computational demand. Moreover, less-complex Gaussian Mixture Models (GMM) were computed than those previously reported. The LSAV system allowed us to reach practically negligible error in healthy (normal) LS verification. LASV system efficiency and the optimal GMM's were evaluated by using Equal Error Rate (EER) and Bayesian Information Criterion (BIC) techniques respectively. These approaches could provide a tool for broader medical evaluation which does not rely, as it is often the case, on a qualitative and subjective description of LS.**

Keywords: ***Quantile Vectors, Lung Sounds, Automatic Verification, Gaussian Mixture Models (GMM).***

I. INTRODUCTION

Utilizing the signals generated by the human body, different applications can be implemented, such as detection of diseases or identification of individuals at risk as a part of screening [1]. In both cases there are two crucial stages: The first is the feature extraction and data representation to emphasize attributes allowing identification [2-4]; the second stage is intended to generate models of the class the data belongs to.

The capacity to classify normal or adventitious lung sounds (*LS*) strengthens the objective aspects of medical diagnoses; in this domain, digital tools can be very accurate and reliable for the LS classification [1, 3]. Our studies were motivated by an increasing number of cases of asthma in children and limited diagnostic capabilities to screen the condition [3-5]. In previous works [3, 4], normal LS recognition was the main concern, while in this paper normal LS classification became the main focus, which led to the development of the LSAV system, and significantly better results were obtained.

In this paper, the term "verification" is used instead of "recognition", due the LSAV system goal is not the explicit identification of a particular pathology pattern, but instead it is designed to verify if a lung sound exhibits, or not, a normal healthy acoustic behavior.

An essential aspect of our studies was to first determine if LS are normal or adventitious to eventually identify what type of pathology the patient presents. Since certain diseases cause characteristic adventitious sounds, the idea is to acquire acoustic signals through a digital stethoscope and to preprocess them based on spectral density. The main motivation is to exploit the development in speech research and respiratory diseases medical care [6-8]. Studies have demonstrated that airflow measurements could reflect morphological changes in the airways of asthma patients by using Quartiles [9, 10]. Adapting these experiments to LS verification, physiology alterations would influence airflow and its spectral representations [3]; therefore, it will be an important criterion for classifying normal or adventitious sounds related pathologies.

II. ACOUSTIC VECTORS

The front-end signal processing in this case involves representation of the acoustic parameters of corresponding vectors. In this context, it is necessary to determine the stationary range in LS for inspiratory and expiratory portions of respiratory cycle durations. The Parameter extraction is the process of measuring particular attributes of the signal to distinguish between normal and pathological signals [3, 4]. The LS signal analysis (Fig. 1) and their spectrogram (Fig. 2) are performed on a succession of theoretically stationary segments called *analysis windows* or *frames*, which may be of 400 ms, overlapped each 100 ms.

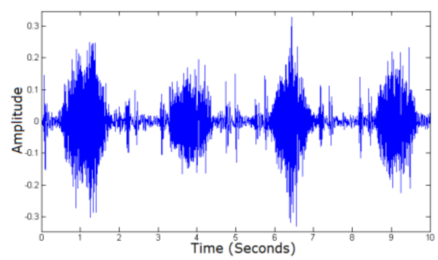

Figure 1. LS signal (Time vs. Amplitude).

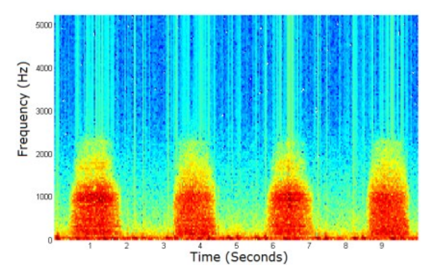

Figure 2. Normal (healthy) LS spectrogram.

Although there are different approaches for acoustic signal representation such as the MFCC (*Mel Frequency Cesptral Coefficients*) vectors, which are an extension of the cepstral methods, and its transformation to a nonlinear frequency space is related to the human hearing [3, 4, 6-8]. In this paper we propose Quantile vectors for LS verification, discussed in details in the next section.

III. QUANTILE PRINCIPLES FOR LS: OCTILES

The Quantiles are points taken at regular intervals of the cumulative distribution function (CDF) of a random variable.

The data is sorted and divided into segments of equal width, to form the empirical distribution function [2].

Complying with a basic principle for a probability distribution, the spectral distribution is normalized (Eq. 1).

$$F_N\,(f) = \int_{-\infty}^{\infty} \frac{f(t)e^{-j2\pi f t}dt}{area(\,F(f)\,)} \qquad (1)$$

This guarantees that the distribution of the frequency components obtained from the FFT will define an area that equals 1. Thereafter, the values for which the Octiles divide the distribution at $f_{.125},\ldots,f_{.875}$ are calculated using Eq. 2.

$$A_{.125} = \int_{-\infty}^{f_{.125}} F_N(f)\,df \ \ ,\ldots, A_{.875} = \int_{-\infty}^{f_{.875}} F_N(f)\,df \qquad (2)$$

It is important to note that calculating the last Quantile is not relevant, because in a normalized process it is always 1 (Fig. 3, Table I).

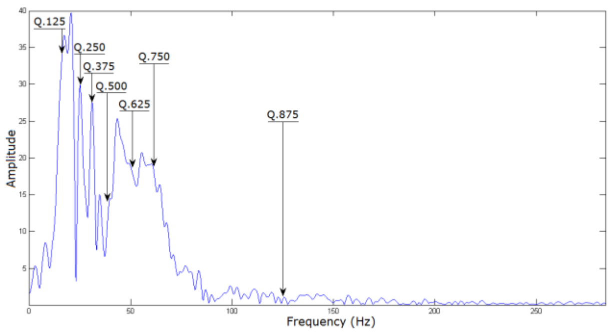

Figure 3. First 400 ms of the Normal LS (Frequency vs. Amplitude).

TABLE I –OCTILE COEFFICIENTS AND RESPECTIVE FREQUENCIES FROM FIG. 3

Frame No.	Octile Coefficient (Q.#)						
	125	250	375	500	625	750	875
Freq. (Hz)	16	21	30	43	52	63	125

IV. RALE AND ITM DATABASES

RALE is a signal repository developed at the University of Manitoba, Winnipeg, Canada [3, 4]. It contains a set of Lung Sounds that were recorded in groups of individuals who had normal breathing, and others with crackles, wheezes, and other peculiar sounds. RALE contains over 50 labeled **.wav* signals (recordings), and 24 additional unlabeled signals, which allow system tests and learning evaluation. The signals are filtered with a 7.5 Hz high-pass Butterworth filter to suppress any DC offset. Moreover, a low-pass 8th-order Butterworth filter set at 2.5 kHz was used to avoid aliasing. RALE signals sampling rate is 11025 Hz.

In order to implement a Normal LS model, numerous LS recordings were taken from ITM subjects (students from Instituto Tecnológico de Mexicali). The group's age ranged from 18 to 25 years-old. They were assessed by applying Stethographics' Interactive Software (STG) to distinguish those with normal breathing from those with normal breathing from those with adventitious sounds. The LS were considered Normal if the subject's breathing didn't exceed a STG's criteria threshold. The ITM recordings were made following a developed protocol with a sample rate of 11025 Hz.

It is known that the heart sounds overlap with LS at low frequencies. Regarding this issue, several pre-processing methods (and their combinations) were made: Sound Activity Detection (SAD), high-pass Butterworth filter for heart, pre-emphasis, etc. Despite all combinations, the database did not reflected significant recognition improvements. Actually, the results were worse because the LS information which is overlapped at heart frequencies was also altered in the process. LSAV system executions were done with unaltered ITM recordings (except for DC mean subtraction), while RALE signals were used as provided.

The adventitious LS signals were taken from RALE database and presented the following cases: 5 asthma, 4 crackles, 4 stridor, 7 wheezes, and 8 normal. Aside from RALE, the database was extended by adding 28 cases from ITM recordings of Normal LS, which corresponded to 7 individuals (4 recordings each), thus expanding the corpus to 36 cases of normal LS.

V. GMM MODELING

A Gaussian Mixture Model (GMM) is characterized by its means, covariance(s) and weights; each class was represented by a GMM model λ. A Gaussian mixture density is a weighted sum of M component densities, as depicted by the Eq. 3:

$$p(\vec{x}\mid\lambda) = \sum_{i=1}^{M} m_i b_i(\vec{x}) \qquad (3)$$

Where $\vec{x}$ is a D-dimensional random vector ($\vec{x}$ is a Octile vector), $b_i(\vec{x})$ i=1,..,M are the component densities, and m_i , i=1,...,M, are the mixture weights. Each component density is a D-variate Gaussian function as shown in Eq. 4.

$$b_i(\vec{x}) = \frac{1}{(2\pi)^{D/2}|\Sigma_i|^{1/2}} \exp\left\{-\frac{1}{2}(\vec{x}-\vec{\mu}_i)'\Sigma_i^{-1}(\vec{x}-\vec{\mu}_i)\right\} \qquad (4)$$

Here, $\vec{\mu_i}$ is the mean vector and Σ_i is the covariance matrix. During the training stage the models for each class (normal and adventitious sounds) were computed by using the correspondent recordings, thus creating a *codebook*. A signal of the same class must be recorded from as many patients as possible to be representative. After the LS signals were recorded, each class was represented as a set of Quantile Vectors. The GMM method uses the EM (Expectation Maximization) algorithm to build the models $\lambda_i = \{\,m_i, \vec{\mu_i}, \Sigma_i\,\}$; this computation was made for each class, *i.e.* creating their classes. In each model, the mean $\vec{\mu_i}$ represents the average of all vectors, while the covariance matrix Σ_i describes the variability of the acoustic class [3, 4].

Gaussian Mixture Models are obtained from multidimensional density histograms (which have been normalized), built from the acoustic vectors. So, instead of examining the distribution of the data, Gaussian mixture curves are adjusted to the multidimensional histograms.

VI. LSAV SYSTEM

The efficiency of the LS automatic verification (LSAV) system depends on the available quantity of signals and training data. In this kind of system there are two error types: *False Acceptance* (FA), when the system classifies an adventitious signal as Normal LS; and *False Rejection* (FR) when a Normal signal is misclassified as Adventitious. System efficiency is measured by using both error types as rates (Eq. 5 & 6).

$$FAR = \frac{FA\ Quantity.}{Qty.\ of\ Adventious\ Signals\ Trials} \quad (5)$$

$$FRR = \frac{FR\ Quantity}{Qty.of\ \ Normal\ Signals\ Trials} \quad (6)$$

System efficiency is expressed in terms of the Equal Error rate (ERR), which is an interpolated point value in which FAR equals FRR. Determining this value is critical for optimal system performance. If the threshold is too high, FRR will rise (Normal LS may be considered Adventitious and rejected); similarly, if the threshold is too low, FAR will increase (Adventitious LS are wrongly accepted). With the purpose of selecting a proper Decision Threshold (Δ_i), the system is adjusted over a validation data set. Efficiency could also be expressed with ROC curves [2].

In order to partition the database, we started with an Automatic Speaker Verification defined procedure, named Lausanne protocol [11], where the Normal LS serves as the *Client*, while the Adventitious serves as the *Imposter*. The Normal and Adventitious models are computed from a Training Data set, and the decision threshold is obtained by using an Evaluation set [12].

The *score* is the $\log_{10}$ of the rate between the Normal LS probability and the Adventitious probability for each evaluated LS input signal; when the Normal LS probability is higher than the Adventitious probability, the result then becomes positive, on the contrary, if the result is lower than the Adventitious probability, the result is negative (Fig. 4, Fig. 5).

In order to run the decision process (Fig. 4), the recognition was computed by using a measurement of model similarity. The RALE+ITM database was divided in two sets: Normal and Adventitious. 18 recordings from the Normal partition were used to build the Normal LS model: 9 for evaluation and 9 for testing. From the Adventitious partition, 7 signals were used for Evaluation and 13 for Testing. The Lausanne protocol was the criterion to select Evaluation and Test populations (partitions).

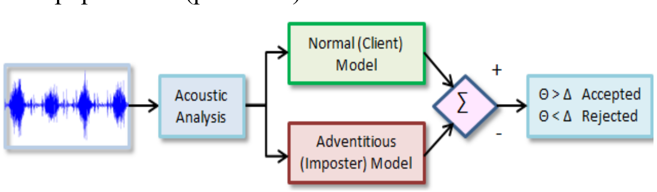

Figure 4. General depiction of the LSAV system workflow.

VII. Results and Discussion

Quantile Vectors (in short-time) were computed from the FFT (Fig. 3). The short-time computation was performed on frames of 400 ms for each signal, as shown in Table I and Fig. 3. As mentioned, the FFT is calculated for each frame, thus obtaining its Quantile Vector. This process is repeated frame by frame over the entire signal, resulting in an *N*x7 matrix, where *N* is the number of obtained vectors along the signal.

In Table I, the Octile Vector 1 (No actv.) corresponds to the first analysis frame, which contains neither expiration nor inspiration activity. The 10th frame Vector (actv.) corresponds to a segment of breathing activity. Since the frequency amplitudes are higher, the Quantile coefficients will present higher values. This key aspect reflects the potential of Quantile Vectors, because their frequency attributes evidence acoustic activity by relating energy and power spectral density.

Regarding Normal LS, the experiments performed with a single Gaussian density model achieved excellent results. It can be noted that a single Gaussian density curve could suffice for Normal LS modeling, given that its Quantile distribution adjusts well to a normal PDF, contrary to the Adventitious LS. It is important to achieve a good classification in normal LS signals, because it could allow determining the possible existence of pathologies. In fact, in previous works [3, 4], this was the principal classification drawback. From the experiments, it was observed that a good selection of database partitions (or *cohort*) for the Adventitious model improves the results. Given that there was a limited quantity of adventitious signals, it was necessary to build the Adventitious model by using both Adventitious partitions (Test and Evaluation). The results were excellent: the LSAV system achieved a 0% EER along an interval obtained in the Evaluation stage (Fig. 5). The results confirm that the system could greatly improve objective auscultatory diagnosis.

On the other hand, Quantile Vectors show their capacity to capture the Normal LS characteristics. If the system classifies an individual's breathing as Adventitious, it could be possible to proceed with a system oriented to pathology detection, such as the one discussed in the literature [3, 4].

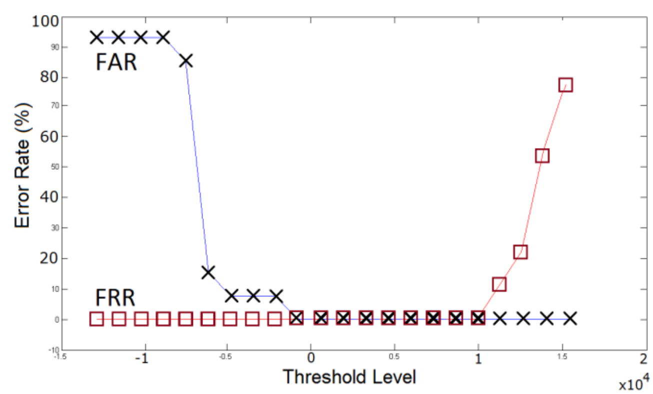

Figure 5. FAR and FRR Curves (*i.e.* EER).

One way to evaluate the model size is by applying either the Bayesian or the Akaike Information Criterions (BIC, AIC). The model with the optimal BIC value is chosen as the 'best' model [13]. The BIC allows estimating how a model adjusts to the data, regarding the number of components, parameter estimates, and a form of the covariance matrices [13]. In order to just visualize this criterion for our Validation system, 50 separate models were built for both Normal and Adventitious LS, and their AIC and BIC values were computed (Fig. 6) with the aim of selecting the best model.

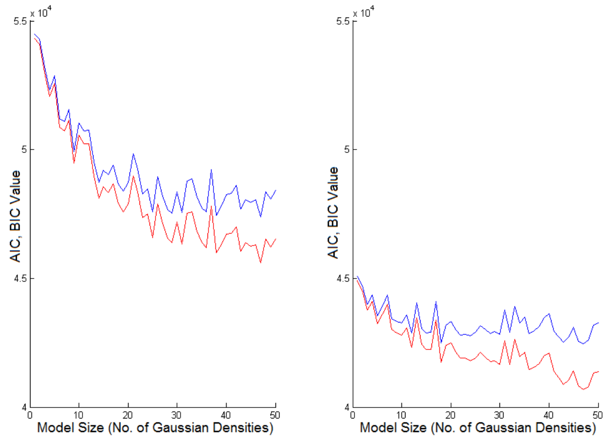

Figure 6. BIC Criterion results for each Model size, for both Normal model (*left*) and Adventitious model (*right*).

In Fig. 6 the horizontal axes represent the number of Gaussian Densities used in the Models, while the vertical axes show their AIC and BIC values. Even though AIC and BIC values decreased as the number of densities increased, neither FAR nor FRR curves exhibited evident behavior improvements; ERR remained 0%. Also, lower AIC and BIC values are observed in normal LS than those in adventitious signals for the same densities number by using Quantiles. At the moment, the experiments oriented to model selection based on the AIC and BIC methodologies are not conclusive.

VIII. Conclusion

In this paper we showed a novel approach in Lung Sounds signal representation, based on the Quantile Vectors, which were obtained through FFT analysis, from signal frames of 400 ms each, and overlapping by 100 ms.

In previous works [3, 4], good results were obtained in Adventitious LS detection, but not the same with Normal LS. Therefore, we carried out the presented verification experiments. As the results show, the Octile Vectors were successful for LS verification, and achieved 0% EER. Our analysis of Acoustic LS was done by applying successful technologies from voice processing that highlight their potential. The application of Quantile Vectors was a follow up of overall studies of breathing mechanism including airflow, volume, and their relationship [9]. The Quantile Vectors showed to be more efficient than MFCC vectors in the case of Normal LS extraction features [3, 4]. Moreover, they led to simpler models, compared to those built from MFCC in the case of Normal LS. In fact, the models computed in our experiments were created with a single Gaussian density and 7-dimensional vectors.

Although the results are highly encouraging, it would be useful to validate the experiments with a higher quantity of LS signals, and also to focus on patient groups organized by age, gender, and body weight. Additionally, it would be useful to include studies of infants and overall asthma evaluation using these techniques. Similarly, it would be useful to implement a verification system for specific types of adventitious sounds. This system could include channel normalization techniques, such as CMN (Cepstral Mean Normalization), CMS (Cepstral Mean Substraction), Sound Activity Detection (similar to Voice Activity Detection, or *VAD*), and experiments regarding model selection. In addition, we will adopt another *ad-hoc* results representation such as ROC curves, or Sensitivity/Specificity measurement.

References

[1] Charleston-Villalobos S., Martinez-Hernandez G., Gonzalez-Camarena R., Chi-Lem G., Carrillo J., Aljama-Corrales T.; Assessment of multichannel lung sounds parameterization for two-class classification in interstitial lung disease patients; Computers in Biology and Medicine; Vol. 41(7), 2011, 473-482; DOI:10.1016/j.compbiomed.2011.04.009

[2] Martinez W.L. and Martinez A.R., "Computational Statistics Handbook with MatLab", Second Edition, Chapman & Hall/CRC, 2008, ISBN 1-58488-566-1.

[3] Mayorga P., Druzgalski., González O.H., Zazueta A., Criollo M.A.;Expanded Quantitative Models for Assessment of Respiratory Diseases and Monitoring, PAHCE 2011, Rio de Janeiro. March 2011; ISBN 978-1-4244-6291-9; DOI: 10.1109/PAHCE.2011.5871938

[4] Mayorga P., Druzgalski C., Vidales J.; Quantitative Models for Assessment of Respiratory Diseases. PAHCE 2010, Lima, Perú, March 2010. IEEE ISBN: 978-1-4244-6

[5] Lev S., Glickman Y., Kagan I., Shapiro M., Moreh-Rahav O., Dahan D., Cohen J., Grinev M., Singer P.; Computerized Lung Acoustic Monitoring Can Help to Differentiate between Various Chest Radiographic Densities in Critically Ill Patients; Respiration Vol. 80, No. 6, 2010: 509-516 (DOI: 10.1159/000274382)

[6] Milner B. P., Semnani S., Robust Speech Recognition over IP Networks, ICASSP2000, IEEE International Conference on Acoustics, Speech and Signal Processing, Istambul, Turkey, 9 June 2000.

[7] Pearce D., An Overview of ETSI Standards Activities for Distributed Speech Recognition Front-Ends, AVIOS 2000: The Speech Applications Conference, San Jose, CA, USA, May 22-24, 2000.

[8] Istrate Dan 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, le 16 decembre 2003, 183 p.

[9] Kiyokawa H., and Pasterkamp H.; Volume-dependent variations of regional lung sound, amplitude, and phase; J.of Applied Physiology Sept.2002 vol. 93 (3). 1030-1038

[10] Gnitecki J., Moussavi Z., Pasterkamp H.; Diagnostic potential in state space parameters of lung sounds; Medical and Biological Engineering and Computing, 2008; Vol. 46, N. 1, 93-99, DOI: 10.1007/s11517-007-0270-2

[11] Poh N., Bengio S., Database, Protocol and Tools for Evaluating Score-Level Fusion Algorithms in Biometric Authentication, IDIAP Research Report, AVBPA2005 New York 2005.

[12] Fredouille C., Mariethoz J., Jaboulet C., Hennebert J., Mokbel C., and Bimbot F. Behavior of a bayesian adaptation method for incremental enrollment in speaker verification. In ICASSP2000 (IEEE International Conference on Acoustics, Speech and Signal Processing), Istanbul, Turkey, June 5.9 2000.

[13] Martinez W., Martinez A., Solka J., Exploratory Data Analysis with MATLAB, 2nd edition. Chapman & Hall/CRC ISBN 978-1-4398-1220-4 p. 207-208, 227-228