WHEEZING SOUNDS DETECTION USING MULTIVARIATE GENERALIZED GAUSSIAN DISTRIBUTIONS

S. Le Cam, A. Belghith, Ch. Collet

LSIIT UMR CNRS 7005 Université Strasbourg 1 (ULP) France

ABSTRACT

A wheeze is a continuous, coarse, whistling sound produced in the respiratory airways during breathing, commonly experienced by persons suffering from asthma. In this paper, we present a new method for the detection of wheezing sounds in the normal breathing sounds. In our study we perform an accurate statistical analysis of breathing signals. We suggest a modeling for wheezing and normal sounds in the wavelet packet domain using generalized gaussian distributions. Our detection method is based on a specific multimodal Markovian modeling proposed in a bayesian framework. We cope with the multidimensional aspect of the generalized gaussian distribution by using the theory of copulas. Experimental results are given in detail in this paper.

Index Terms— Adventitious Respiratory Sounds, Data Fusion, Hidden Markov Chain, Generalized Gaussian Distribution, Copulas Theory.

1. INTRODUCTION

It is assumed that asthma, with diabetes, will be the most common chronic illness in the XXIst century and is constantly growing. In many countries roughly five percent of the population suffer from asthma and other related respiratory illnesses. Its detection is still carried out by pulmonary auscultation using a stethoscope and implies limitations due to the subjectivity of this process. Indeed, it depends on the individual's own hearing, experience and its ability to differentiate patterns. Nowadays, there is a clear need for a normalization of the diagnosis methodology and for the development of a common framework for all the medical community [1]. In this context, much of the knowledge gained in recent years has resulted from the use of modern digital processing techniques, which leads to objective analysis and comparisons of respiratory sounds [2].

Adventitious lung sounds fall into two main categories: crackles and wheezes [1, 3]. Wheezes are musical adventitious respiratory sounds, also called continuous, and their presences keep relation to partial airway obstruction. According to the American Thoracic Society, their duration is usually greater than 100 ms and smaller than 250 ms, which is significantly higher relatively to other abnormal sounds, such as crackles, typically lasting less than 20 ms [1]. Its dominant spectral range is sharp but it can be highly variable between one patient and another, as well as one pathology and another. It is commonly observed between 100 and 1600 HZ [3]. Wheezing with unforced breathing is correlated with severity of airway obstruction, so its measurement is a useful tool for evaluating the severity of asthma. F. Salzenstein

Laboratoire INESS, UMR CNRS 7163 Université Strasbourg 1 (ULP) France

A survey of literature from the last decade shows that the main methodologies used for wheezes analysis is Fourier peaks detection and spectrogram image analysis, based on the definition of a threshold above which wheezing signals can be distinguished from the noise of a normal breathing sound [4, 3]. In recent years, several efforts for more sharpness and precision for the modeling and the detection of wheezing sounds have been achieved. The time/scale representation is often selected for its good time/frequency localization properties. Hadjileontiadis *et al.* [5] use the time/scale representation of breathing sound recordings, while Bahoura *et al.* [6] use the Subband based Cepstral parameters (SBC) to characterize wheezing sounds.

We propose a segmentation using Hidden Markov Chain (HMC) model in the wavelet packet domain. In a scaling and fusion prospect of multivariate data, we go further in respiratory sounds analysis, by detecting the respiratory phases (inspiration and expiration) [7] in order to perform multivariate treatments on observations recorded on several points of the chest. The data likelihood is a generalized gaussian probability density function chosen for its good ability to model a wide range of symmetric distributions. We will show up that the shape parameter α of this probability density function can be seen as a discriminant value for the respiratory sounds classification.

First part of the paper gives a review on Generalized Gaussian Distribution (GGD) and its extension to a multivariate expression by using the theory of copulas. Second part presents the segmentation method and provides details on its computation using Mode of Posterior Marginals (MPM) criterion and Iterative Conditional Estimation (ICE) of the parameters. In particular, we specify how the estimation of the variance and the shape parameter for GGD are computed. Results on real data are then exposed and discussed.

2. MODELING OF PULMONARY SOUNDS IN WAVELET PACKET DOMAIN

2.1. Data Representation Space and Data Scaling

Some previous studies on wheeze detection [4, 3] have been carried out in the time-frequency domain, through a local power spectrum analysis using a Fourier transform. It is well known that this transform suffers from a lack of precision and flexibility due to the Gabor-Heisenberg uncertainty principle. Thus, in order to improve the detection accuracy, we use a wavelet packet representation. The decomposition based on the so-called Daubechies wavelet - using 20 filter coefficients - provides a well sparsed representation for our respiratory signals and is well adapted to our segmentation method.

Wheezing sounds are observed in a frequency range between 100 and 1600 HZ, and their main frequency characteristics are well

Thanks to ANR ASAP and to Region Alsace for funding.

localized, usually in a 50 Hz wide frequency band. Analyzing the whole [100, 1600] Hz window can lead to an overwhelming of the relevant information, and we then need to chose wavelet packet corresponding to sharper frequency bands. The sounds we worked on are sampled at 8000 Hz. The wavelet packets of level 5 were selected, they correspond to a 125 Hz resolution per packet which is acceptable for the detection of wheezes.

The auscultation is often performed on several points on the chest of the patient, the physicians exploit the whole information to construct their diagnoses. This can be seen like a kind of *mental information fusion*, and we wish to imitate this methodology in our model by developing fusing algorithms of these multivariate data. It is obvious that the information consistency is a main issue for a fusion method, thus a selection and a scaling of the relevant data is necessary. Furthermore, wheezing signals usually appear at the same time in each respiratory phase, and our fusion method will be helpful to localize the wheezing waveform in several contaminated respiratory phases.

In previous works, we proposed a non-invasive method to perform the detection of both inspiration and expiration phases in recorded respiratory signal [7]. This detection is helpful to scale homogeneous data (*e.g* an inspiration phase with another) and open the possibility of a consistent data fusion. The scaling is done in the time domain through spline cubic interpolation, which guarantees to preserve the frequency properties of the signal [8]. The effective fusion of the information is then performed using markovian model (see section 3) with a multivariate GGD as the data likelihood (see section 2.2).

2.2. Wavelet Coefficients Modeling

In order to propose appropriate priors for the wavelet coefficients, the histograms related to the normal and pathological sounds have been studied (see fig. 1). For crackling to normal respiratory noises, it confirms the sparse property commonly formulated for the wavelet distribution, stating that "the wavelet transform of a large class of signals results in a *large* number of *small* coefficients and in a *small* number of *large* coefficients", and that many signals can be well approximated by a *small* number of their wavelet coefficients [9]. On the other hand, the waveform of a wheezing signal owns the shape of a sinusodal signal, and thus its distribution is more like a uniform distribution.

Then we need a distribution which could model both peaky and heavy tailed distributions, and uniform-like distributions. Not many of these types of distributions can be found. Gaussian mixture have already been exploited to tackle this kind of issue [6], but are difficult to estimate in this context where the data are centered for each class. The generalized Gaussian distribution (GGD) showed better abilities for the modeling of our data, and has already been approved by Mallat [10] for the modeling of wavelet coefficients. The GGD is given by:

$$f_{GG}(x,\alpha,\sigma,\mu) = \frac{\eta(\alpha)\alpha}{2\Gamma(1/\alpha)} \exp\left(-(\eta(\alpha)|x-\mu|)^{\alpha}\right) \quad (1)$$

With μ , σ and α respectively the mean, variance and shape parameters, $\Gamma(.)$ the gamma function and $\eta(\alpha) = \left[\frac{\Gamma(3/\alpha)}{\sigma^2\Gamma(1/\alpha)}\right]^{\frac{1}{2}}$.

The shape parameter $\alpha > 0$ detemines the rate of exponential decay of the PDF. Note that for $\alpha = 2$ the density reduces to the Gaussian density, whereas for $\alpha = 1$ it becomes the Laplacian density. Furthermore, the uniform distribution ($\alpha \rightarrow \infty$) and δ (dirac)

function ($\alpha \rightarrow 0$) are also special cases of GGD. This distribution can cover a wide range of distribution and explains its success for data modeling, and especially for wavelet distribution fitting.



Fig. 1. Wavelet histograms and GGD fitting for: (a) a normal sound ($\alpha = 1.8$); (b) a wheezing sound ($\alpha = 19.6$).

2.3. Multivariate expression of the generalized gaussian distribution using copulas

Each multivariate observation $\mathbf{y} = (y^1, ..., y^d)$ takes its values in \mathbb{R}^d We wish to use multivariate distribution to model the observation likelihood $f(y^1, ..., y^d)$. In the general non-Gaussian case, the computation of multivariate distribution is not trivial and one assumes the likelihood on each mode to be independent. Since such assumption is often not verified, this problem can be solved with the copulas theory. The basis of this theory is the Sklar Theorem [11] which asserts the existence of a function C, called copula and defined on $[0, 1]^d$, binding the joint distribution function $f(y^1, ..., y^d)$ to the marginal distribution functions $f^1(y^1), ..., f^d(y^d)$, as follows [11]:

$$f(y^1, ..., y^d) = f^1(y^1) \times ... \times f^d(y^d) C(F^1(y^1), ..., F^d(y^d))$$
(2)

provided that the cumulative marginals $F^1(y^1), ..., F^d(y^d)$ are continuous, and that C is differentiable.

For multivariate Gaussian copula C_g , the later is given by:

where $\tilde{\mathbf{u}} = (\Phi^{-1}(u^1), ..., \Phi^{-1}(u^d))$, with Φ the standard gaussian cumulative distribution, R the inter-band correlation matrix and I the $d \times d$ identity matrix.

To model non-gaussian multivariate densities, we use the equation (2) with a gaussian copula density given by the equation (3) and and generalized Gaussian marginal densities defined by the equation (1).

3. MULTIMODAL MARKOV CHAIN MODELING FOR THE DETECTION OF WHEEZING

3.1. Hidden Markov Model

٢

Let $X = (x_n)_{1 \le n \le N}$ and $\mathbf{Y} = (\mathbf{y_n})_{1 \le n \le N}$ be two stochastic processes. X is hidden and takes its value in a finite set $\Omega = \{\omega_1, ..., \omega_K\}$. **Y** model the multivariate observation $(\mathbf{y_n})_{1 \le n \le N}$. The problem consists in estimating X from **Y**. The Hidden Markov Chain (HMC) model has been widely used in this context [12, 13], for its adaptability and time performance. We will use it in this paper with the classical assumption of independence of **Y** conditionnally on X, and the assumption of stationnarity for the chain X (HMC-IN). The *priors* on X are then given by the initial probability $\pi(.)$, and the $K \times K$ square transition matrix $a_{ij} = p(x_n = \omega_j | x_{n-1} = \omega_i)$. The distribution of the couple (X, \mathbf{Y}) is then given by:

$$P(X = \{\omega_1 \dots \omega_N\}, \mathbf{Y}) = \pi(x_1 = \omega_1) f_{\omega_1}(\mathbf{y_1})$$
$$\cdot \prod_{n=2}^N a_{(n-1) n} f_{\omega_n}(\mathbf{y_n}) \quad (4)$$

This factorization allows one to compute the *posterior* probabilities $P(x_n|\mathbf{y})_{1 \le n \le N}$ using the well-known forward-backward probabilities. The final decision on X is then obtained through a bayesian approach, using the Maximal Posterior Marginals criterion (MPM) [14]. The parameters estimation is performed using the ICE algorithm. Next section provides details on the computation of the parameters estimation.

3.2. Model Parameters Estimation

In the case of unsupervised classification, all the HMC-IN parameters are unknown and must be estimated from the observed data **Y**. The first step consists in estimating the *prior* parameters and the data driven parameters. We use here the ICE procedure [15, 16]. The distribution of the stationary Markov chain X is driven by the initial probabilities $\pi(.)$ and the transition matrix a_{ij} . It is possible to determine exact computation for these *prior* parameters. On the contrary, the computation of exact estimation expressions for the data driven parameters is intractable. Their update is then compute in a stochastic way using samples from $P(X|\mathbf{Y})$. The data driven parameters to estimate are the means $\mathbf{M} = {\mu_k^j}$, the form parameters $\mathbf{A} = {\alpha_k^j}$ and the covariance matrix $\mathbf{R} = {R_k}$ for each class ω_k and each mode j.

Unsupervised classification first implies an initialization of the parameters:

• Initialization: $\Theta^{[0]} = \{\pi^{[0]}, a_{ij}^{[0]}, \mathbf{R}^{[0]}, \mathbf{M}^{[0]}, \mathbf{A}^{[0]}\}.$

Prior parameters on X are initialized in a deterministic way:

$$\pi^{[0]}(i) = \frac{1}{K}, \quad a_{ij}^{[0]} = \begin{cases} \frac{3}{4} \text{ if } i = j\\ \frac{1}{4(K-1)} \text{ if } i \neq j \end{cases}$$
(5)

Data driven parameters $\{\mathbf{R}, \mathbf{M}, \mathbf{A}\}\$ are initialized using the segmental K-means algorithm. The initialization of the covariance matrix R_i associated with the class $\omega_i \in \Omega$ is a diagonal matrix containing the variance for each mode previously estimated with the K-means algorithm.

• For each q in \mathbb{N}^* , $q \leq Q$ (Q is set by the user), we compute the following ICE estimation for the parameters update $\Theta^{[q]} = \{\pi^{[q]}, a_{ij}^{[q]}, \mathbf{R}^{[q]}, \mathbf{M}^{[q]}, \mathbf{A}^{[q]}\}$:

- Computation of the *posterior* probabilities $\xi_n^{[q]}(i) = P(x_n = \omega_i | Y)$ and $\Psi_n^{[q]}(i, j) = P(x_n = \omega_i, x_{n+1} = \omega_j | \mathbf{Y})$ using the forward-backward probabilities. A realization $\hat{X}^{[q]} = \{\hat{x}_n^{[q]}\}_{1 \le n \le N}$ is then computed using $\xi_n^{[q]}(i)$.

- Equation for the maximization step [16]:

Deterministic estimation of the prior parameters:

$$\pi^{[q]}(i) = \xi_1^{[q]}(i), \qquad a_{ij} = \frac{\sum_{n=1}^{N-1} \Psi_n^{[q]}(i,j)}{\sum_{n=1}^{N-1} \xi_n^{[q]}(i)} \tag{6}$$

Stochastic estimation of the data driven parameters:

$$\boldsymbol{\mu}_{k}^{[q]} = \frac{\sum_{n \in I_{k}^{[q]}} \mathbf{y}_{n}}{N_{k}} \tag{7}$$

$$R_{k}^{[q]} = \frac{\sum_{n \in I_{k}^{[q]}} (\mathbf{y}_{n} - \boldsymbol{\mu}_{k}^{[q]}) (\mathbf{y}_{n} - \boldsymbol{\mu}_{k}^{[q]})^{t}}{N_{k}}$$
(8)

With $I_k^{[q]} = \{n | \hat{x}_n^{[q]} \in \hat{X}^{[q]}, \hat{x}_n^{[q]} = \omega_k\}$ and $N_k = \operatorname{card}(I_k^{[q]})$ refers to the cardinal of this set.

The estimation of the shape parameter is solution of the following equation:

$$\alpha_{k}^{j \, [q]} + \phi(1/\alpha_{k}^{j \, [q]}) + \log(\alpha_{k}^{j \, [q]}/N_{k}) +$$

$$\log(G_{\alpha_{k}^{j \, [q]}}) - \alpha_{k}^{j \, [q]} \frac{G_{\alpha_{k}^{j \, [q]}}'}{G_{\alpha_{k}^{j \, [q]}}} = 0$$
(9)

With $G_{\alpha_k^j[q]} = \sum_{n \in I^k} |y_n^j - \mu_k^j[q]|^{\alpha_k^j[q]}, G'_{\alpha} = \frac{\partial G_{\alpha}}{\partial \alpha}$ and $\phi = \frac{\Gamma'(x)}{\Gamma(x)}.$

This equation is solved numerically in our algorithm.

4. RESULTS AND DISCUSSION

Our fusion methodology has been applyed on real respiratory data with a frequency sampling of 8000Hz. The wavelet packet for the level 5 have been exploited in reason of their good compromise between time and frequency resolutions regarding our problem.

The signals shown in fig. 2 (mode 1 and mode 2) are two real inspiration signals with overlaying wheezing, extracted from the auscultation of a same patient. These signals are represented in the wavelet packet domain, the frequency band picked is [250, 375] Hz. Once the detection of the inspiratory phases is done for both signals, the scaling of the relevant phases is computed. Thus, representations of the both inspiratory signals in the wavelet packet selected contain the same number of wavelet coefficients. This scaling offers the possibility to apply our fusion method. Figure 2 shows the data fusion result of the proposed method and compares it with independant segmentation of the two signals, in order to show up the relevance of our method facing the high variability of respiratory sounds.

For the first mode, the segmentation problem is difficult because of high vesicular coefficients, and its processing does not allow a sharp detection of the wheezing waveform. Wavelet coefficients from vesicular noises have been misclassified as wheezing coefficients (see fig. 2(a)). The segmentation of the second mode is relatively easy since the wheezing waveform is well above the vesicular coefficients. The processing on this signal is successfully performed (see fig. 2(b)). We will then gain from the information contained in this second mode to rectify the segmentation error on the first mode. As can be seen on the last figure 2(c), the fusionned segmentation we obtain represents well the location of the wheezing waveform in both signals.



Fig. 2. Two wavelet packets (level 5, [250, 375] Hz) of two real inspiration phases with overlaying wheezing (mode 1 and 2), (a) gives the segmentation result on the mode 1, (b) on the mode 2 and (c) gives the result of the fusion method. ω_1 is the 'normal class' and ω_2 is the 'wheezing class'.

The form parameter α seems to be a good descriptor for the quantization of the wheezing strength in a contaminated signal, since it increases with the amplitude of the wheezing waveform. For the wheezing class (ω_2) in the mode 1, the estimation of our algorithm gives $\alpha_2^1 = 4.5$, while it estimates $\alpha_2^2 = 9.8$ for the mode 2. This parameter can then be regarded as a descriptor of the state of a pathology, and will be useful to measure the effectiveness of a treatment or the degeneration of a pathology.

These results show that the proposed method uses both signals to increase the effectiveness of the segmentation. However the segmentation map does not return a sharp resolution, and in more rough case where the wheezing is not defined as well as in the exemple above, our fusion method fails to detect the wheeze. This is in part due to the regularizing aspect of the HMC model, which tends to drown some significant details of the signal. Also, only a few of the information available is exploited (only one packet for each signal). Further work will be based on hierarchical treatment using Hidden Markov Tree (HMT) models, which include the scale-to-scale dependency and are more reliable to cope with the sharp details.

5. ACKNOWLEDGEMENTS

This work is a part of the ASAP (Analysis of Auscultatory and Pathological Sounds) project overseen by the French National Research Agency. This project is done in collaboration with the "Hôpitaux Universitaire de Strasbourg". We wish especially to thank the professors Emmanuel Andrès and Anne Charloux for their advices and for the data collection.

6. REFERENCES

- H. Pasterkamp, S.S. Kraman, and G.R. Wodicka, "Respiratory sounds: advances beyond the stethoscope," *Am. J. Respir. Crit. Care Med.*, vol. 156, pp. 974–987, 1997.
- [2] A. R. A. Sovijarvi, J. Vanderschoot, and J. E. Earis, "Standardization of computerised respiratory sound analysis," *Eur. Respir. Rev.*, vol. 10, pp. 586, 2000.
- [3] ARA Sovijarvi, LP Malmberg, G. Charbonneau, J. Vanderschoot, F. Dalmasso, C. Sacco, M. Rossi, and JE Earis, "Characteristics of breath sounds and adventitious respiratory sounds," *Eur Respir Rev*, vol. 10, no. 77, pp. 591–596, 2000.
- [4] A. Homs-Corbera, R. Jane, JA Fiz, and J. Morera, "Algorithm for timefrequency detection and analysis of wheezes," *Engineering in Medicine* and Biology Society, 2000. Proceedings of the 22nd Annual International Conference of the IEEE, vol. 4, 2000.
- [5] SA Taplidou, LJ Hadjileontiadis, IK Kitsas, KI Panoulas, T. Penzel, V. Gross, and SM Panas, "On Applying Continuous Wavelet Transform in Wheeze Analysis," *Engineering in Medicine and Biology Society*, 2004. EMBC 2004. Conference Proceedings. 26th Annual International Conference of the, vol. 2, 2004.
- [6] M. Bahoura and C. Pelletier, "Respiratory sounds classification using Gaussian mixture models," *Electrical and Computer Engineering*, 2004. Canadian Conference on, vol. 3, 2004.
- [7] S. Le Cam, Ch. Collet, and F. Salzenstein, "Acoustical respiratory signal analysis and phase detection," Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on, pp. 3629–3632, 2008.
- [8] M. Unser, "Splines: a perfect fit for signal and image processing," Signal Processing Magazine, IEEE, vol. 16, no. 6, pp. 22–38, 1999.
- [9] MS Crouse, RD Nowak, and RG Baraniuk, "Wavelet-based statistical signal processing using hidden Markov models," *Signal Processing*, *IEEE Transactions on*, vol. 46, no. 4, pp. 886–902, 1998.
- [10] S. Mallat, "A theory for multiresolution signal decomposition: the wavelet represention," *IEEE Transactions on PAMI*, vol. 11, no. 7, pp. 674–693, 1989.
- [11] M. Sklar, "Fonctions de répartition à n dimensions et leur marges," Publ. Inst. Stat. Univ. Paris, vol. 8, pp. 229–231, 1959.
- [12] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–286, 1989.
- [13] K. Aas, L. Eikvil, and R.B. Huseby, "Applications of hidden Markov chains in image analysis," *Pattern Recognition*, vol. 32, no. 4, pp. 703– 713, 1999.
- [14] J. Marroquin, S. Mitter, and T. Poggio, "Probabilistic solution of illposed problems in computer vision," *Journal of the American Statistical Association*, vol. 82, no. 397, pp. 76–89, 1987.
- [15] W. Pieczynski, "Champs de Markov cachés et estimation conditionnelle itérative," *Traitement du Signal*, vol. 11, no. 2, pp. 141–153, 1994.
- [16] J.N. Provost, Ch. Collet, P. Rostaing, P. Pérez, and P. Bouthemy, "Hierarchical Markovian segmentation of multispectral images for the reconstruction of water depth maps," *Computer Vision and Image Understanding*, vol. 93, no. 2, pp. 155–174, 2004.