A multi-feature approach for noise detection in lung sounds

A. Leal, C. Teixeira, I. Chouvarda, N. Maglaveras, J. Henriques, R. Paiva and P. Carvalho

Abstract— During the acquisition of lung sounds, several sources of noise can interfere with the recordings. Therefore, the detection of noise present in lung sounds plays an important role in the correct diagnosis of several pulmonary disorders such as in chronic obstructive pulmonary diseases.

Denoising tools reported so far focus mainly in the detection of abnormal lung sounds from the background noise (usually vesicular background) or even just in the discrimination of normal from abnormal lung sounds. Algorithms for heart sound cancellation have also been proposed. However, it can be noticed that there is a lack of signal processing methods to efficiently detected and/or remove artifacts introduced in the acquisition environment or produced by the subject (e.g., speech).

The present study focuses in the analysis of lungs sounds recorded in two different populations containing events of cough, speech and other artifacts from the surrounding environment. Feature extraction and binary classification were performed achieving, on average, values of a sensitivity and specificity ranging from 76% to 97% for the classification of cough, speech and other artifacts and from 83% to 90% for the specific detection of cough events. The detection of artifacts achieved sensitivity and specificity values of 84% and 61%, respectively for one population and 88% and 52% for another population.

I. INTRODUCTION

The acquisition of lung sounds (LSs) using the stethoscope is considered an important, fast and noninvasive method in the detection of pulmonary disorders, namely chronic obstructive pulmonary diseases (COPD) [1]. The presence of abnormal LSs in the normal respiratory cycle is an indicator that leads to a positive diagnosis.

LS is a biological signal characterized by a stochastic non-stationary behavior, with sudden variations in short periods of time over the signal [1], [2]. The signal's bandwidth typically ranges from $50\,\text{Hz}$ to $2500\,\text{Hz}$ (when signals are acquired on the chest) [3], [4]. In fact, the morphology of the LS can vary from subject to subject, depending on the chest size and on the body mass, and it is also influenced by the variations in the air flow rate and on the position where the stethoscope is attached in the chest [5], [6].

A. Leal, C. Teixeira, J. Henriques, R. Paiva and P. Carvalho are with the Center for Informatics and Systems, Polo II, University of Coimbra, 3030-290 Coimbra, Portugal adriana.leal@student.fisica.uc.pt and {cteixei, jh, ruipedro, carvalho}@dei.uc.pt

I. Chouvarda is with the Laboratory of Medical Informatics, Medical School, Aristotle University of Thessaloniki, Thessaloniki 54124, Greece ioanna@med.auth.gr

N. Maglaveras is with the Laboratory of Medical Informatics, Medical School, Aristotle University of Thessaloniki, Thessaloniki 54124, Greece and the Institute of Applied Biosciences, Center for Research and Technology Hellas nic-mag@med.auth.gr

However, LSs can be contaminated with noise coming from the acquisition environment and also with physiological interferences, such respiratory muscle and heart sounds (HSs) [3], [4]. Specifically, concerning HSs interference, it is known that the predominant frequency components typically detected in those signals may overlap the frequency range of the LSs, making it hard to distinguish both types of sound [5], [7].

Then, in order to truthfully conclude about the clinical state of a subject, concerning pulmonary disorders, the LSs should ideally be free of any type of noise. Following that purpose, several approaches in literature have been proposed to detect abnormal LSs from the vesicular background (considered the background noise) present in the original LS. Those methods include higher-order statistics (HOS), fuzzy logic, wavelet transform (WT), empirical mode decomposition (EMD) and fractal dimension (FD) [1], [4]. In most of the studies, one infers that the acquisitions take place in controlled clinical environments, where external noise is intentionally reduced or just nonexistent.

There is also a wide number of authors who presented algorithms developed to remove HSs from the original LSs, using mainly adaptive HOS, WT, recursive least squares and time-frequency spectrum analysis [4].

However, only few articles were found to report signal processing methods to detect and remove artifacts from LSs, in m-Health environments. Two papers reported algorithms based on spectral subtraction of a noise estimate obtained from: a period when the patient sustained breathing [8] or a reference sensor that solely and simultaneously recorded the ambient noise [6]. In another study a denoising filter based on a combination of EMD and FD was described, returning 91.8% of sensitivity and 97.7% of specificity [1].

The aforementioned methods can be of great usefulness when LSs need to be acquired in unpredictable noisy environments. In other words, denoising tools for LSs are required in telemonitoring systems that could provide continuous monitoring of patients with pulmonary disorders.

Herein, we analyze LSs that were deliberately contaminated with periods of cough followed by speech performed by each patient. Furthermore, other sources of noise can be heard that include voices and cough from other subjects in the room and hair rub. The purpose of the study is the detection of all types of artifact, to allow the further interpretation of the LSs. We follow a multi-feature approach to face the problem of noise detection in LSs.

II. MATERIAL AND METHODS

A. Data Collection

Pulmonary signals were recorded in two different populations: at the (1) George Papanikolaou General Hospital of Thessaloniki and (2) General Hospital of Imathia-Health Unit of Naoussa, Greece (see Table I). All subjects comprising the Thessaloniki dataset were diagnosed with COPD. Wheezes or wheezes and crackles were identified in the LSs of six subjects of the Naoussa population. The ethical committee of both hospitals authorized the acquisition of the data.

Auscultation was performed with the participants in a sitting position, using six channels that were set in different positions: four in the back and two in the front of the chest. During the acquisition, the volunteers were asked to simulate cough and next to count from one to ten. The different events were annotated in the timeline by the physicians who supervised the acquisition and we assigned them to four distinct classes: (1) cough and speech, (2) artifact, (3) abnormal LSs, and (4) normal LSs (see Table I). Cough and speech periods last on average 5-6 s in the Thessaloniki population and are also the predominant events in the Naoussa dataset.

B. Feature Extraction

The features were computed for each individual channel, since each channel corresponds to different positions in the chest and also to different recording times.

Four different features were tested in the analysis of LSs and will be briefly described in this section. The LSs signals were windowed, with a given percentage of overlap. The size of the window and the overlap were defined for the highest specificity and sensitivity, depicted in a receiver operating characteristic (ROC) curve.

1) Teager-Kaiser energy operator (TKEO): As a nonlinear energy operator, this feature is useful to highlight sudden discontinuities (either in amplitude or frequency) and reduce smooth transitions occurring in the signal [9].

In (1) the discrete version of TKEO is computed for the signal x_n. Energy is obtained for each window in each instant and depends on the strength of the previous and next sample

TABLE I: Population Description

Population	Thessaloniki	Naoussa
Sampling Rate (Hz)	10000	4000
Stethoscope	3M Littman 3200	Meditron
Number of subjects	7	9
Mean signal duration (s)	70	19
Abnormal lung sounds	wheezes, fine crackles, coarse crackles	
Artifacts	external voice, cough with artifacts	speech background, cough background, hair rub

[9]. So far only another author used this feature for the detection of crackles [10].

$$E_n = x_n^2 + x_{n-1}x_{n+1} \qquad (1)$$

2) Katz fractal dimension (KFD): This feature is a measure of the complexity of a signal in the time domain. KFD is estimated in (2), where L is the sum of the Euclidean distance between successive points, d is the distance between the first point of the window and the point of the window at which the distance is maximal, n is the number of steps in the window ($n = N - 1$), and N is the length of the input data [4], [11], [12].

$$KFD = \frac{\log_{10}(n)}{\log_{10}\left(\frac{d}{L}\right) + \log_{10}(n)} \qquad (2)$$

3) Mel-frequency cepstral coefficients (MFCC): The cepstral coefficients are obtained from the inverse cosine transform of the logarithm of the short-time Fourier transform (STFT) of each window and are given by (3):

$$c_n = \sum_{k=1}^{M} \log(E_k)\left(n(k - 0.5)\frac{\pi}{M}\right) \qquad (3)$$

where $n = 1, 2, ..., L$, L is the desired number of cepstral coefficients, M is the number of frequency subbands and E_k is the respective filter bank energy. This approach uses a different weighting frequency scale that closely resembles how human ear perceives sounds. The cepstrum, in addition to the spectrum which informs about the predominant frequencies present in the signal, allows to know how those frequencies change across time [6], [13]. The extracted feature corresponds to the coefficient with the highest module of specificity and sensitivity in the ROC curve.

4) Normalized mutual information (NMI): Mutual information quantifies the statistical dependence between random variables or, in other words, the amount of information shared by those variables. The I between two variables X and Y, denoted $I(X;Y)$, is obtained using the corresponding entropies, $H(X)$ and $H(Y)$ and the joint entropy $H(X,Y)$ [14]:

$$I(X;Y) = H(X) + H(Y) - H(X,Y) \qquad (4)$$

$H(X)$ and $H(Y)$ are given by:

$$H(X) = -\sum_{x \in X} p(x)\log p(x) \qquad (5)$$

With x taking any of the nonrepeated values in the variable X and $p(x)$ corresponding to the frequency counts of the discrete variable X. Joint entropy is given by:

$$H(X,Y) = -\sum_{x,y} p(x,y) \log p(x,y) \qquad (6)$$

In order to scale the values of $I(X;Y)$ between 0 (statistically independent variables) and 1 (variables with the same information) normalization was performed:

$$NMI = \frac{I(X;Y)}{H(X)} \qquad (7)$$

NMI was computed for a reference window, free of noise, and for each of the windows into which the LS was segmented (named test windows). The final feature corresponds to the comparison between the NMI of the reference window and each of the test windows. This method of finding a reference window was developed in another study which purpose was the detection of artifact in phonocardiogram signals [15]. However, this algorithm to denoise HSs has suffered slight changes aiming to adapt it to the study of LSs. Shortly, the algorithm comprises two steps: (1) search for a reference window considered free from noise and (2) comparison of the reference window with all the windows into which the signal was segmented (named test windows), in order to distinguish contaminated from clean windows.

In the first phase, differently from the HSs' algorithm, the LSs signal is windowed and the TKEO is computed for each window. The reference window will correspond to the window with the minimum value of TKEO. Other two features, spectral similarity (R) and PSD ratio, described in [15], were also computed for the LSs.

C. Classification

Detection of artifacts is performed in three stages. Firstly, events other than LSs (cough, speech and other artifacts) were classified using different features (TKEO, KFD, MFCC, R, and PSD ratio). In the second part, the events of cough were classified using features NMI and MFCC, depending on dataset. In the final phase, the detection of artifact, including the speech segments, took place using the output of the two previous classifications (see Fig. 1).

The binary classification outputs a vector of 0's and 1's (called "detected"), where one corresponds to the detection of the artifact and zero to a nondetected event, respectively. Consequently, classification only takes place if a fixed threshold is set for each of the features. If the value of the feature in the test window is above this threshold, that window is classified as contaminated. In order to obtain the best values of specificity and sensitivity for each feature several thresholds were tested and further analyzed using the ROC curve. The 100 tested thresholds ranged from the minimum to the maximum of the feature. In other words, the tested thresholds were a percentage of the signal, making them adaptive thresholds. The discrimination capacity of the features was assessed in terms of specificity and sensitivity, by comparing the output of the classification, the "detected", with the target (annotated by the physicians).
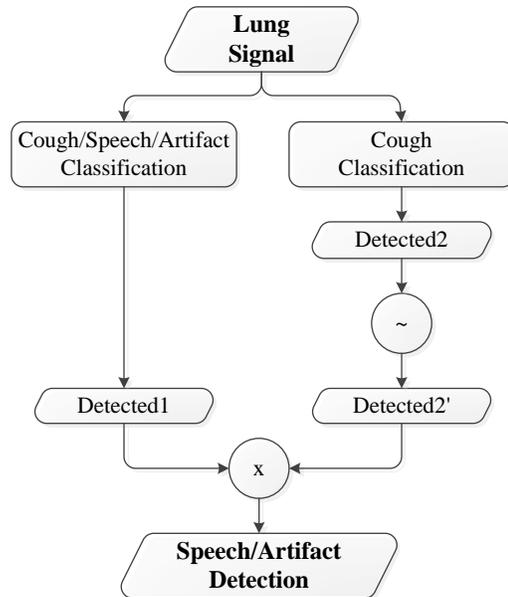
Fig. 1: Multi-stage algorithm for the classification of artifact.

The algorithm was implemented in Matlab R2013b on Windows 8.1 using an Intel® Core™i7-4790K CPU at 4 GHz.

III. RESULTS AND DISCUSSION

Tables II and III contain the results of the classification in terms of sensitivity and specificity, averaged across the subjects for each one of the parameter combination (window, overlap and threshold).

A. Detection of cough, speech and artifact events

The detection results obtained for the five features previously described in section II-B are presented in Table II. The best results were obtained for the features TKEO and KFD in the Thessaloniki dataset and MFCC in the Naoussa population. Naoussa signals were of lower size comparing to the files from Thessaloniki, explaining the variation in the window size across the two populations.

B. Detection of cough events

Cough is an important manifestation of several pulmonary diseases (e.g. pulmonary fibrosis, COPD, lung cancer, etc.). Consequently, counting and classification of cough is considered a useful diagnostic tool [16]. The cough events were classified using the NMI and MFCC features in the Thessaloniki and Naoussa populations, respectively. Unlike other tested features, NMI and MFCC (the latter using a different window size) allowed the identification of the cough events rather than the detection of both cough and speech segments. Results are presented in Table III.

C. Detection of speech and artifact events

For this classification the "detected" vectors obtained in sections III-A and III-B were used to compute the "detected"

TABLE II: Classification results for cough, speech and artifact detection.

Population	Feature	Window (s)	Overlap (%)	Threshold*	SS(%)	SP(%)
Thessaloniki	TKEO	3.9	90	9	92.04	97.24
	KFD	4	90	10	90.91	99.24
	MFCC	4	80	58	86.52	93.71
	PSD ratio	4	80	9	85.98	92.81
	R	4	90	5	86.14	96.37
Naoussa	TKEO	2.3	60	57	63.09	99.50
	KFD	2.8	70	53	69.18	98.79
	MFCC	1	90	64	84.66	96.40
	PSD ratio	1	90	20	78.47	94.18
	R	0.5	90	3	84.38	96.29

*Percentage of the features' amplitude.

TABLE III: Classification results for cough detection.

Population	Feature	Window (s)	Overlap (%)	Threshold*	SS(%)	SP(%)
Thessaloniki	NMI	4	90	81	84.49	83.05
Naoussa	MFCC	0.5	90	83	90.36	90.09

*Percentage of the features' amplitude.

for the third phase of the algorithm (see Fig. 1). The results of the classification of the class speech and artifact versus the class cough and LSs are presented in Table IV.

The results are strongly dependent on the annotations (which will influence the target definition). Moreover, the detection of artifact, the main purpose of this study, was affected by the size of the windows selected through the ROC curve. In the specific case of the Thessaloniki dataset, using a $4\,\text{s}$ window lead to good results in the classification of cough and speech since those events' duration is superior to the window used to segment the signal. On the contrary, other artifact events such external voices or hair rub last less than $4\,\text{s}$ making these type of events hard to detect when cough and speech are the predominant segments in the signal.

IV. CONCLUSIONS

The detection of artifact in lung sounds becomes an important step when those signals need to be acquired in unpredictable conditions, e.g., when there are different sources of noise in the acquisition environment. Therefore, in the search for a suitable methodology to identify interferences in LSs, several features were tested and optimal parameters were selected. The features analyzed reached a high detection capacity in the classification of cough, speech and other artifacts occurring in uncontrolled environments. This performance was reflected in the values of sensitivity and specificity, which average value for all the features was 88% and 96%, respectively for Thessaloniki population and 76% and 97%, respectively for the Naoussa dataset. The results for the classification of cough were equally high, with sensitivity and specificity ranging, respectively from 83% to 90% in both populations. The classification of artifact, including speech against the other events, lead to lower sensitivity (ranging from 84% to 88%) and even lower specificity (ranging from 61% to 56%) in both populations.

As future work, the set of features tested will be considered as inputs of more complex classifiers able to define more flexible, nonlinear decision boundaries. Data from each channel, i.e., each position of the stethoscope can also be analyzed separately in order to conclude about the more suitable locations to detect noise. Finally, it must be noticed that a careful inspection of the annotations of the events will have a significant impact in the specificity and sensitivity.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

TABLE IV: Classification results for speech and artifact detection.

Population	SS(%)	SP(%)
Thessaloniki	84.39	61.54
Naoussa	87.72	56.38

REFERENCES

[1] L. Hadjileontiadis, "Empirical mode decomposition and fractal dimension filter: a novel technique for denoising explosive lung sounds," *IEEE Engineering in Medicine and Biology Magazine*, vol. 26, no. 1, pp. 30–39, Jan. 2007. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4069353

[2] A. Bohadana, G. Izbicki, and S. S. Kraman, "Fundamentals of lung auscultation." *The New England journal of medicine*, vol. 370, no. 8, pp. 744–51, Feb. 2014. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/24552321

[3] A. Gurung, C. G. Scrafford, J. M. Tielsch, O. S. Levine, and W. Checkley, "Computerized lung sound analysis as diagnostic aid for the detection of abnormal lung sounds: a systematic review and meta-analysis." *Respiratory medicine*, vol. 105, no. 9, pp. 1396–403, Sep. 2011. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3227538&tool=pmcentrez&rendertype=abstract

[4] L. J. Hadjileontiadis, "Lung Sounds: An Advanced Signal Processing Perspective," *Synthesis Lectures on Biomedical Engineering*, vol. 3, no. 1, pp. 1–100, Jan. 2008. [Online]. Available: http://www.morganclaypool.com/doi/abs/10.2200/S00127ED1V01Y200811BME009

[5] Z. Moussavi, "Fundamentals of Respiratory Sounds and Analysis," *Synthesis Lectures on Biomedical Engineering*, vol. 1, no. 1, pp. 1–68, Jan. 2006. [Online]. Available: http://www.morganclaypool.com/doi/abs/10.2200/S00054ED1V01Y200609BME008

[6] G.-C. Chang and Y.-F. Lai, "Performance evaluation and enhancement of lung sound recognition system in two real noisy environments." *Computer methods and programs in biomedicine*, vol. 97, no. 2, pp. 141–50, Feb. 2010. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0169260709001734

[7] T. H. Falk and W.-Y. Chan, "Modulation filtering for heart and lung sound separation from breath sound recordings." *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference*, vol. 2008, pp. 1859–62, Jan. 2008. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/19163050

[8] A. Yadollahi and Z. Moussavi, "On arithmetic misconceptions of spectral analysis of biological signals, in particular respiratory sounds." *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference*, vol. 2009, pp. 388–91, Jan. 2009. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/19964737

[9] E. Kvedalen, "Signal processing using the Teager Energy Operator and other nonlinear operators," Jul. 2003. [Online]. Available: https://www.duo.uio.no/handle/10852/8823

[10] L. Mendes, P. Carvalho, C. A. Teixeira, R. P. Paiva, and J. Henriques, "Robust features for detection of crackles: An exploratory study." *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference*, vol. 2014, pp. 1473–6, Aug. 2014. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/25570247

[11] L. Hadjileontiadis and I. Rekanos, "Detection of explosive lung and bowel sounds by means of fractal dimension," *IEEE Signal Processing Letters*, vol. 10, no. 10, pp. 311–314, Oct. 2003. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1232728

[12] S. Charleston-Villalobos, G. Dorantes-Méndez, R. González-Camarena, G. Chi-Lem, J. G. Carrillo, and T. Aljama-Corrales, "Acoustic thoracic image of crackle sounds using linear and nonlinear processing techniques." *Medical & biological engineering & computing*, vol. 49, no. 1, pp. 15–24, Jan. 2011. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/20652429

[13] M. Bahoura, "Pattern recognition methods applied to respiratory sounds classification into normal and wheeze classes," *Computers in Biology and Medicine*, vol. 39, no. 9, pp. 824–843, Sep. 2009. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0010482509001267

[14] G. Brown, A. Pocock, M.-J. Zhao, and M. Luján, "Conditional likelihood maximisation: a unifying framework for information theoretic feature selection," *The Journal of Machine Learning Research*, vol. 13, no. 1, pp. 27–27–66–66, Jan. 2012. [Online]. Available: http://dl.acm.org/citation.cfm?id=2503308.2188387

[15] D. Nunes, A. Leal, R. Couceiro, J. Henriques, L. Mendes, P. Carvalho, and C. Teixeira, "A Low-Complex Multi-Channel Methodology for Noise Detection in Phonocardiogram Signals," in *37TH ANNUAL INTERNATIONAL CONFERENCE OF THE IEEE ENGINEERING IN MEDICINE AND BIOLOGY SOCIETY*, 2015, p. 4. [Online]. Available: http://emb.citengine.com/event/embc-2015/paper-details?pdID=5652

[16] C. Lucio, C. Teixeira, J. Henriques, P. de Carvalho, and R. P.

Paiva, "Voluntary cough detection by internal sound analysis," in *2014 7th International Conference on Biomedical Engineering and Informatics*. IEEE, Oct. 2014, pp. 405–409. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7002808