

Detection of wheezes using their signature in the spectrogram space and musical features

L. Mendes, I. M. Vogiatzis, E. Perantoni, E. Kaimakamis
I. Chouvarda, N. Maglaveras, V. Tsara, C. Teixeira, P. Carvalho, J. Henriques, R. P. Paiva

Abstract— In this work thirty features were tested in order to identify the best feature set for the robust detection of wheezes. The features include the detection of the wheezes signature in the spectrogram space (WS-SS) and twenty-nine musical features usually used in the context of Music Information Retrieval. The method proposed to detect the signature of wheezes imposes a temporal Gaussian regularization and a reduction of the false positives based on the (geodesic) morphological opening by reconstruction operator. Our dataset contains wheezes, crackles and normal breath sounds.

Four selection algorithms were used to rank the features. The performance of the features was asserted having into account the Matthews correlation coefficient (MCC). All the selection algorithms ranked the WS-SS feature as the most important. A significant boost in performance was obtained by using around ten features. This improvement was independent of the selection algorithm. The use of more than ten features only allows for a small increase of the MCC value.

I. INTRODUCTION

The automatic detection of adventitious sounds (additional respiratory sounds superimposed on breath sounds) is a valuable non-invasive tool to detect and follow-up respiratory diseases such as chronic obstructive pulmonary disease (COPD). Adventitious sounds include wheezes (continuous sounds), stridor, squawks and crackles (discontinuous sounds).

Crackles are short explosive sounds that are associated with cardiopulmonary diseases and typically present a very characteristic waveform. These sounds seem to result from an abrupt opening or closing of the airways [1].

Wheezes are continuous sounds that are usually associated with obstructions in the air passages. These whistling sounds are characterized by periodic waveforms with duration equal or longer than 100 ms [2]. Due to their musical nature, these sounds have a distinct signature in the spectrogram space (see Fig. 2).

This work was supported by iCIS (CENTRO-07-ST24-FEDER-002003) and by EU Project WELCOME (FP7 - 611223).

L. Mendes, P. Carvalho, C. Teixeira, R. P. Paiva and J. Henriques are with CISUC, Department of Informatics Engineering, University of Coimbra, Portugal (e-mail: {lmgendes, carvalho, cteixeira, jh, ruipedro}@dei.uc.pt).

I. Vogiatzis is with General Hospital of Imathia- Health Unit of Naoussa, Greece (email: ioanvogiatzis@gmail.com).

V. Tsara, E. Perantoni and E. Kaimakamis are with the General Hospital of Thessaloniki ‘G. Papanikolaou’, Greece. (email: bpneumonologiki@yahoo.gr, elperantoni@gmail.com, vkaimak@med.auth.gr).

N. Maglaveras and I. Chouvarda are with the Laboratory of Medical Informatics, Medical School, Aristotle University of Thessaloniki, Thessaloniki, Greece. (email: {nicmag, ioanna}@med.auth.gr).

Different methods were proposed to automatically detect wheezes, such as 1) based on the detection of the wheezes signature in the spectrogram space [3][4][5][6], 2) using the Mel-frequency cepstral coefficients combined with Gaussian mixture model [7][8], 3) based on auditory modeling [9], 4) based on tonal index combined with the correlation function [10] and 5) based on sample entropy [11].

In this study we assert the capacity of thirty features to detect wheezes. Our features include the result of the detection of the wheeze signature and twenty-nine musical features. A new method for the detection of the signature of the wheezes in the spectrogram space is presented. Our dataset contains not only wheezes but also crackles and normal breath sounds.

The performance of the feature(s) to discriminate respiratory sounds with wheezes was studied taking into account the Matthews correlation coefficient (MCC) [12] measured after classifying the data using the logistic regression classifier (LR) [13] or the random forest classifier (RF) [13]. Four feature selection methods were tested.

II. MATERIAL AND METHODS

A. Data

Data of twelve volunteers (nine patients and three healthy subjects) were acquired at the General Hospital of Thessaloniki ‘G. Papanikolaou’ and at the General Hospital of Imathia (Health Unit of Naoussa), Greece. The respiratory sounds of six patients contain wheezes or wheezes and crackles. The healthy subjects exhibit normal respiratory sounds, while another set of three patients had only crackle manifestations. Auscultations positions were selected among the six possible positions presented in Fig. 1. For each volunteer we selected the data acquired from the two positions where the adventitious sounds/normal sounds were better heard, i.e., twenty-four acquisitions of approximately 30 seconds were used in this study.

The data were acquired at 4000 Hz using a 3M Littman electronic stethoscope (model 3200), which complies with the EMC requirements of the IEC 60601-1-2. The acquisitions were done with the volunteers in the sitting position.

One hundred and thirteen wheezes were annotated in the temporal space. Using this information, the partitions on the spectrogram space (see section II-B) were annotated as containing or not containing wheezes. The ethical committee of the General Hospital of Thessaloniki ‘G. Papanikolaou’ authorized the acquisition of the data.

B. Detection of the wheezes signature in the spectrogram space (WS-SS)

Fig. 2 presents the diagram of the proposed method to detect the signature of wheezes in the spectrogram space.

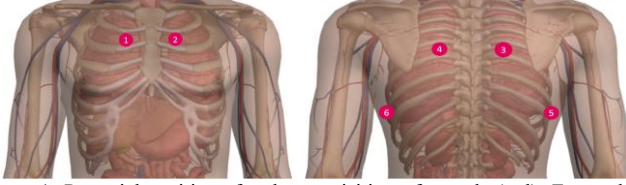


Figure 1. Potential positions for the acquisition of sounds (red). For each volunteer we selected the data acquired from the two positions where the adventitious sounds/normal sounds were better heard.

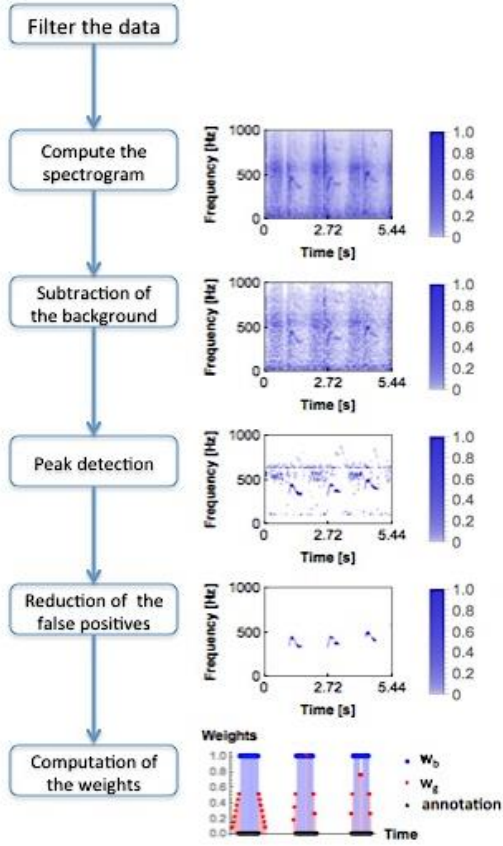


Figure 2. Diagram of the proposed method to detect the signature of wheezes in the spectrogram space (WS-SS).

The method has the following steps:

1 – Filter the signal

The first derivative of the discrete Gaussian kernel [14] was used to filter the signal. The kernel size was equal to 5 bins (on the time domain).

2 - Compute the spectrogram

The spectrogram [14] of the filtered signal, $S[t, f]$ (with f the frequency and t the time), is computed using a flat top window, partitions with the length equal to 512 bins and an offset equal to 128 bins (on the time domain).

3 - Subtraction of the background

The subtraction of the background (tread) was done using the method proposed in [3], i.e.,

$$S^B[t] = S[t] - B[t] \quad (1)$$

where $S[t]$ is an array with the frequencies values computed at t and $B[t]$ is the background estimated using a moving average filter applied to $S[t]$ and S^B the spectrogram without the background.

4 - Peak detection

In this step we identify the elements of $S^B[t, f]$ that should be part of a wheeze following the method used in [3]. As in [3] we restrict our search to the interval of frequencies between 100 and 1000 Hz. All the wheezes in our dataset are within this interval. We use two frequency bands, $B_1 = [100, 600[$ Hz and $B_2 = [600, 1000]$ Hz. For each frequency band B_k , with $k = \{1, 2\}$, we calculate a binary matrix, P , with the same size as the spectrogram S using

$$P_k[t, f] = \begin{cases} 1 & \text{if } P_k[t, f] \geq \overline{S_k^B[t]} + C_k \sigma(S_k^B[t]) \\ 0 & \text{else} \end{cases}, \quad (2)$$

where the value of 1 in the matrix P correspond to a possible element of a wheeze, the $\overline{S_k^B[t]}$ and $\sigma(S_k^B[t])$ correspond to the mean value and the standard deviation, respectively, of $S_k^B[t]$. The $C_k = \{1.5, 2.5\}$ are the thresholds used.

5 - Reduction of false positives

For the reduction of the false positives we propose to use the (geodesic) morphological opening by reconstruction operator. We apply the Mathematica function GeodesicOpening [14] to P using as structuring element a box matrix.

6 - Computation of the array of weights (w)

After the reduction of the false positives we compute a binary array of weights, w_b , (see Fig. 3) using

$$w_b[t] = \begin{cases} 1 & \text{if } \sum_f P[t, f] \geq 1 \\ 0 & \text{else} \end{cases} \quad (3)$$

In order to improve the accuracy of the classification we propose to add a temporal Gaussian regularization to the binary weights (see Fig. 3). Fig. 4 describes the method used to compute the additional Gaussian weights, w_g . The final array of weights, w , is the sum of w_b with w_g .

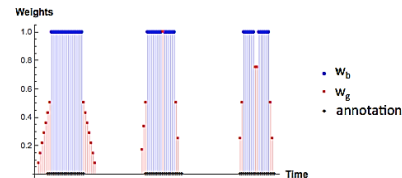


Figure 3. Example of the values of the weights, computed in the step 6 of the proposed method to detect the WS-SS, as a function of the time. The binary weights, w_b , are represented in blue and the gaussian weights, w_g , in red. The black points correspond to the annotation.

Computation of the gaussian weights, w_g , and the (final) weights w

Input: w_b array with the binary weights
Output: w_g array with the gaussian weights
 w array with the final weights

1. Create a constant array, $w_{g\text{tmp}}$, of zeros with the same size as w_b .
2. Identify the groups of elements of ones in w_b .
3. For each group of ones

- 3.1 Create a constant array, $w_{g\text{tmp}}$, of zeros with the same size as w_g .
- 3.2 Compute the length, L , the start index, $s\text{Ind}$, and the end index, $e\text{Ind}$.
- 3.3 Compute $\sigma = L / (2 \cdot \sqrt{2 \log(2)})$.
- 3.4 Compute

$$fwhm_{1/2} = \left\lceil \frac{2 \cdot \sqrt{2 \log(2)} \cdot \sigma}{2} \right\rceil;$$

$$fwtm_{1/2} = \left\lceil \frac{2 \cdot \sqrt{2 \log(10)} \cdot \sigma}{2} \right\rceil;$$

$$\Delta = fwtm_{1/2} - fwhm_{1/2} \text{ and } \phi = 0.5 / \Delta.$$

- 3.5 Set $w_g[e\text{Ind} + inc] = 0.5 - (inc - 1) \cdot \phi$ and

$$w_g[s\text{Ind} - inc] = 0.5 - (inc - 1) \cdot \phi \text{ with } inc = 1, 2, \dots, \Delta.$$

- 3.6 Compute $w_g = w_g + w_{g\text{tmp}}$.

4. Compute $w = w_b + w_g$.
-

Figure 4. Pseudocode to compute the array of gaussian weights, w_g , from the array of binary weights, w_b . The final array of weights, w , is the sum of w_b with w_g .

C. Musical features

In this study we tested twenty-nine musical features computed using the MIRtoolbox [15]. The timbre and tonal features were obtained taking into consideration only the frequencies between 100 Hz to 1000 Hz (as we have done in II-B). The frame duration used was 128 ms and the hop factor (ratio) was 0.25. Table 1 presents the musical features used in this study.

TABLE 1. MUSICAL FEATURES AND THE CORRESPONDENT LABELS USED IN THIS STUDY.

Dynamic	Label
Root-mean square energy	2
Timbre	-
Brightness	3
Centroid	4
Flatness	5
Irregularity	6
Keyclarity	7
Kurtosis	8
13 Mel-frequency cepstral coefficients	[9,21]
Rolloff 85	22
Rolloff 95	23
Roughness	24
Spread	25
Skewness	26
Zerocross	27
Tonal	-
Chromagram centroid	28
Chromagram peak	29
Mode	30

D. Performance criteria

The MCC, measured after classifying the data using the LR classifier or the RF classifier, was used to assert the capacity of the features to detect wheezes. The MCC is a balanced performance measure, specially suitable when the

dataset are unbalanced [12]. The sensitivity (Se), specificity (Sp) and the accuracy (Acc) were also measured. Each partition/frame was classified as containing or not containing wheezes.

Four state-of-the-art feature selection algorithms were used to rank the importance of the features: the least absolute shrinkage and selection operator (LASSO) [13], the variable importance estimated using a RF (RF-VI) [13], the sequential feature selection [13] in the forward direction (SFS_Forward) and in the backward direction (SFS_Backward). The objective function used in the sequential feature selection was the maximization of the MCC measured after classifying the data using the LR classifier.

For each classification a stratified 10-fold cross-validation approach with ten Monte Carlo repetitions was used.

III. RESULTS

Table 2 presents the first fifteen features, ranked by importance, selected by the four different feature selection algorithms. Fig. 5 and Fig. 6 show the MCC measured after classifying the data using the LR classifier and the RF, respectively, as a function of the number of features ranked by the different selection algorithms. Table 3 and Table 4 present the MCC, Se, Sp, and Acc measured after classifying the data using the LR and the RF, respectively, using one, ten and thirty (all) features.

TABLE 2. RANK OF THE FIRST 15 FEATURES SELECTED (#) USING FOUR DIFFERENT FEATURE SELECTION ALGORITHMS: LASSO, RF-VI, SFS_FORWARD AND SFS_BACKWARD. SEE TABLE 1 TO IDENTIFY THE CORRESPONDENCE BETWEEN LABELS AND THE MUSICAL FEATURES. THE LABEL 1 CORRESPONDS TO THE WS-SS FEATURE.

#	LASSO	RF-VI	SFS_Forward	SFS_Backward
1	1	1	1	1
2	6	6	13	13
3	23	24	18	17
4	20	7	23	12
5	3	2	19	5
6	14	19	6	21
7	5	20	28	28
8	24	16	22	6
9	27	17	9	22
10	19	29	5	15
11	29	28	29	19
12	28	14	21	18
13	21	11	20	10
14	10	21	27	23
15	22	15	2	4

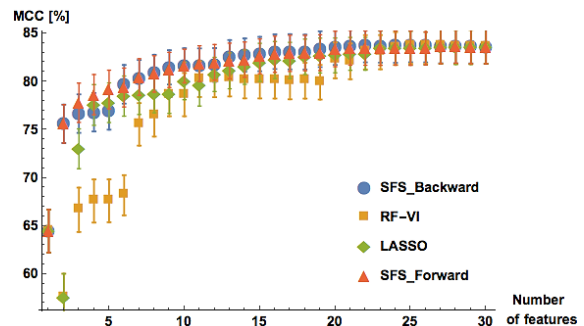


Figure 5. MCC measured after classifying the data using the LR classifier as a function of the number of features ranked by LASSO (green), RF-VI (yellow), SFS_Forward (red) and SFS_Backward (blue).

IV. DISCUSSION AND CONCLUSION

The objective of this study was to test the performance of thirty features to detect wheezes (adventitious lung sounds that have a musical character). One of the features attends to detect the signature of the wheezes in the spectrogram space (see Fig 2). Our algorithm to detect this signature imposes a temporal Gaussian regularization (see Fig. 3) and a reduction of the false positives based on the (geodesic) morphological opening by reconstruction operator. The other twenty-nine features are usually used in the context of Music Information Retrieval (see Table 1). Our dataset contains wheezes, crackles and normal breath sounds.

Two classification algorithms, the logistic regression and the random forest, were used to evaluate the performance of the different features. As expected, the performance criteria (MCC, Sp, Se and Acc) were higher with the random forest classifier (see Table 3 and Table 4). Using thirty features combined with RF it is observed a MCC, Se, and Sp equal to $92.7 \pm 1 \%$, $90.9 \pm 2 \%$ and $99.4 \pm 1\%$, respectively.

In this study we evaluated four different selection algorithms to rank the features. All of these algorithms ranked the WS-SS feature as the most important (see Table 2). Using only this feature and the LR to classify the data (see Table 3), we measured a MCC, Se, and Sp equal to $64.4 \pm 2 \%$, $84.4 \pm 2\%$ and $88.7 \pm 1\%$, respectively. As can be observed in Fig. 5 and Fig. 6, the addition of more features allows for the improvement of the MCC value. When the RF was used to classify the data, similar values of MCC were obtained using the features ranked by the different selection algorithms (see Fig. 6).

Independently of the selection algorithm applied, a significant improvement was obtained using around ten features (see Fig. 5 and Fig. 6). The addition of more features only allows for a smaller improvement.

In the future, after validating the results of this study with more data (acquired under the EU WELCOME [16] project), the identified features will be utilized to monitor the presence of wheezes in patients with COPD.

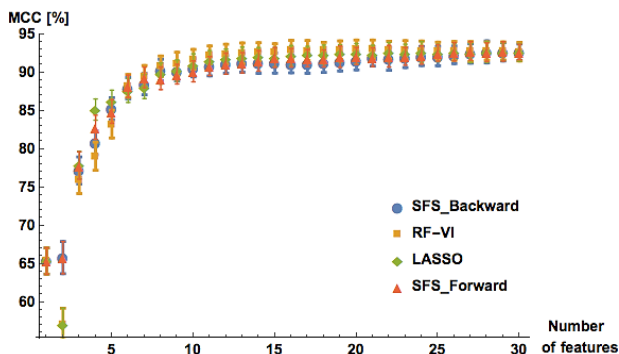


Figure 6. MCC measured after classifying the data using the RF classifier as a function of the number of features ranked by LASSO (green), RF-VI (yellow), SFS_Forward (red) and SFS_Backward (blue).

TABLE 3. MCC, SE, SP AND ACC MEASURED AFTER CLASSIFYING THE DATA USING THE LOGISTIC REGRESSION CLASSIFIER USING 1, 10 AND ALL FEATURES. THE FEATURES WERE SELECTED USING THE SFS_FORWARD.

Nº of Features	MCC [%]	Se [%]	Sp [%]	Acc [%]
1	64.4 ± 2	84.4 ± 2	88.7 ± 1	87.9 ± 1
10	81.6 ± 2	82.0 ± 2	97.6 ± 1	94.9 ± 1
30	83.6 ± 2	82.7 ± 2	98.1 ± 1	95.5 ± 1

TABLE 4. MCC, SE, SP AND ACC MEASURED AFTER CLASSIFYING THE DATA USING THE RANDOM FOREST CLASSIFIER USING 1, 10 AND ALL FEATURES. THE FEATURES WERE SELECTED USING THE RF-VI.

Nº of Features	MCC [%]	Se [%]	Sp [%]	Acc [%]
1	65.3 ± 2	86.5 ± 2	88.3 ± 1	88.0 ± 1
10	91.7 ± 1	89.6 ± 2	99.3 ± 2	97.7 ± 1
30	92.7 ± 1	90.9 ± 2	99.4 ± 1	97.9 ± 1

REFERENCES

- [1] P. Piirilä and A. R. A. Sovijärvi, "Crackles: recording, analysis and clinical significance," *Eur. Respir. J.*, vol. 8, no. 12, Dec. 1995.
- [2] R. A. Sovijärvi, F. Dalmasso, J. Vanderschoot, L. P. Malmberg, and G. Righini, "Definition of terms for applications of respiratory sounds," *Eur. Respir. Rev.*, vol. 10, issue 7, pp. 597–610, 2000.
- [3] S. A. Taplidou and L. J. Hadjileontiadis, "Wheeze detection based on time-frequency analysis of breath sounds," *Comput. Biol. Med.*, vol. 37, issue 8, pp.1073-1083, 2007.
- [4] R. J. Riella, P. Nohama and J. M. Maia, "Method for automatic detection of wheezing in lung sounds", *Braz. J. Med. Biol. Res.*, vol. 42, issue 7, pp. 674-684, 2009.
- [5] D. Emmanouilidou, K. Patil, J. West, and M. Elhilali "A multiresolution analysis for detection of abnormal lung sounds", in *Conf Proc Eng Med Biol Soc.*, pp. 3139-3242, 2012.
- [6] Y. Shabtai-Musih, J. B. Grotberg, and N. Gavriely, "Spectral content of forced expiratory wheezes during air, He, and SF6 breathing in normal humans," *J Appl. Physiol.*, vol. 72, pp. 629–635, 1992.
- [7] M. Bahoura, "Pattern recognition methods applied to respiratory sounds classification into normal and wheeze classes," *Comput Biol Med.*, vol 39, issue 9, pp. 824-843, 2009.
- [8] P. Mayorga, C. Druzgalski, R. L. Morelos, O. H. Gonzalez and J. Vidales, "Acoustics based assessment of respiratory diseases using GMM classification," in *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, pp. 6312-6316, 2010.
- [9] Y. Qiu, A. Whittaker, M. Lucas and K. Anderson, "Automatic wheeze detection based on auditory modeling," in *Proc. Inst. Mech. Eng. H.*, vol. 219, issue 3, pp. 219-227, 2005.
- [10] M. Wisniewski and T. P. Zielinski, "Fast and Robust Method for Wheezes Recognition in Remote Asthma Monitoring," in *Proc. Int. Conf. on Inf. Tech. in Bio.*, pp. 568- 572, 2012.
- [11] J. Feng, S. Farook and G. Daniel," Automatic wheeze detection using histograms of sample entropy," in *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, pp: 1890 – 1893, 2008.
- [12] G. Jurman, S. Riccadonna, and C. Furlanello, "A comparison of MCC and CEN error measures in multi-class prediction," *PLoS One*, vol. 7, 2012.
- [13] T. Hastie, R. Tibshirani and J. Friedman, *The elements of statistical learning: data mining, inference and prediction*, Springer, 2008.
- [14] Wolfram Research, Inc., *Mathematica*, Version 10.0, Champaign, IL, 2014.
- [15] O. Lartillot and P. Toiviainen, "MIR in Matlab (II): A Toolbox for Musical Feature Extraction From Audio," in *Int. Conf. Mus. Inf. Ret.*, Vienna, 2007.
- [16] "WELCOME Project" [Online]. Available: <http://www.welcome-project.eu/about-the-project.aspx>. [Accessed: 10-Mar-2015].