Automatic cough episode detection using a vibroacoustic sensor

Marcel Młyńczak[1], Katarzyna Pariaszewska[1] and Gerard Cybulski[1,2]

***Abstract*— Cough monitoring is an important element of the diagnostics of respiratory diseases. The European Respiratory Society recommends objective assessment of cough episodes and the search for methods of automatic analysis to make obtaining the quantitative parameters possible. The cough "events" could be classified by a microphone and a sensor that measures the vibrations of the chest. Analysis of the recorded signals consists of calculating the features vectors for selected episodes and of performing automatic classification using them. The aim of the study was to assess the accuracy of classification based on an artificial neural networks using vibroacoustic signals collected from chest. Six healthy, young men and eight healthy, young women carried out an imitated cough, hand clapping, speech and shouting. Three methods of parametrization were used to prepare the vectors of episode features - time domain, time-frequency domain and spectral modeling. We obtained the accuracy of *95%* using artificial neural networks.**

I. INTRODUCTION

Chronic cough is a very common symptom of respiratory diseases like asthma, chronic obstructive pulmonary disease (*COPD*) or lung cancer. Usually, cough monitoring consists of the subjective assessments of patients (collected during the medical interviews) and the results of questionnaires based on a visual analogue scale. However, the European Respiratory Society (*ERS*) recommends searching for methods of automatic analysis of cough in order to make obtaining quantitative parameters possible [1], [2]. Analysis of cough frequency and intensity gives an objective method to diagnose and assess the progression of therapy [3], [4]. Currently, during cough monitoring, it is very important to evaluate cough episodes comprehensively with combinations of subjective and objective tools [5].

Cough can be treated as an example of vibroacoustic signal, what determines the choice of methods for signal recording. Most often, the cough sensors consist of microphones for recording the sound component of the signal and accelerometers for recording the chest vibrations [1], [2], [4], [6]. The microphones in cough monitoring systems are very small in order to allow attachment to patients and design of a portable system for daily activity registration. Mostly, omnidirectional electret condenser microphones with sensitivity at the level of $-40dB$ (Sony, Panasonic, Bruel & Kjaer) are used [7], [8], [9]. The chest vibrations are mainly recorded using accelerometers and piezoelectric sensors [1], [4], [10], [11]. Accelerometers are placed on the surface of the patient's body, usually between the cartilage of the thyroid and the highest position of the sternum [2], [4]. Furthermore, other additional sensors and modules have been already used, e.g. *ECG* and *EMG* channels, oxygen saturation measuring units (SpO_2) or esophageal pressure topography ones [3], [6], [12], [13].

Automatic analysis is based on extraction and classification of cough "events" from the recorded signals. Parametrization of the signals through calculation of vectors of episodes features seems to be necessary for this task. This could be accomplished by processing the signals or their envelopes in the time domain (with amplitude peak detection or zero-crossing rate calculation algorithms), in the time-frequency domain (with spectral analysis) or with spectral modeling [6], [9], [14]. The methods for calculating spectral parameters are often those widely implemented in speech recognition, e.g. mel-frequency cepstrum analysis (*MFCA*) or linear predictive coding (*LPC*) [14], [15]. In some cases, the parameters of discrete wavelet analysis are also determined [11], [16]. The most commonly used classification methods are statistical models, such as Hidden Markov Models (*HMM*), in which the aim of the learning procedure is to create unique models for each class of event signal, represented as feature vectors or artificial neural networks (*ANN*), where the results in the output layer of the network determine membership of the appropriate class of signals [14], [15], [16], [17], [18].

The main objective of this work was to assess the accuracy of cough episode detection based on artificial neural networks using vibroacoustic signals recorded from chest. We also wanted to evaluate the suitability and impact of chosen descriptive episode parameters which were served as input feature vectors.

II. METHODS

We registered vibroacoustic signals for the imitated "events" of cough, single hand clap, shouting and speech with a single Polish consonant or vowel. To record the data we used an omni-directional condenser microphone (BCM, Bestar) and a piezoelectric sensor placed on a belt (MLT1132 Piezo Respiratory Belt Transducer, ADInstruments) and positioned them on the patient's chest, as in Fig. 1. Both signals were measured by the multichannel recorder, ADInstruments PL3516 PowerLab 16/3, with sampling frequencies of *20kHz* and *200Hz* for sounds and vibrations, respectively. We performed signal acquisition using the dedicated software, LabChart8. The recorded signals were next pre-processed by filtering, silent-part removal and signal envelope calculation.

[1]M. Młyńczak, Student Member, IEEE, K.Pariaszewska and G. Cybulski are with Institute of Metrology and Biomedical Engineering, Faculty of Mechatronics, Warsaw University of Technology, Boboli 8, Warsaw, Poland `mlynczak at mchtr.pw.edu.pl`

[2]G. Cybulski is also with the Department of Applied Physiology, Mossakowski Medical Research Centre, Polish Academy of Sciences, Pawinskiego 5, Warsaw, Poland

978-1-4244-9270-1/15/$31.00 ©2015 IEEE

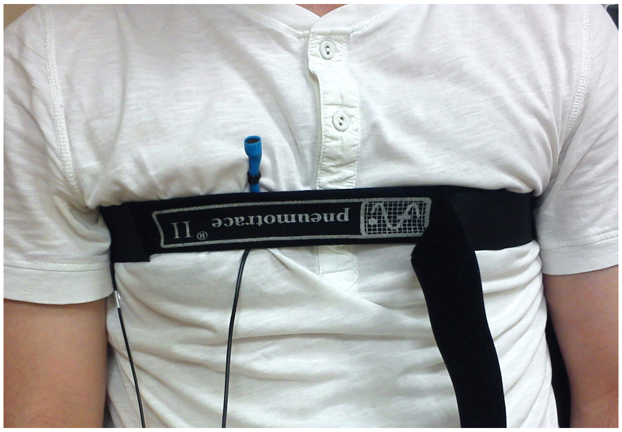

Fig. 1. Positioning of the microphone and MLT1132 Piezo Respiratory Belt Transducer

The measurements were carried out on 14 healthy students: 6 males aged 22-24 (M: *23.7*; SD: *0.8*) and 8 females aged 23-26 (M: *24.0*; SD: *0.9*) - without any reported respiratory or asthmatic diseases.

The study procedure consisted of 3 series. In each one imitated cough, reading the Polish alphabet (24 phonemes), hand-clapping and shouting (basically loudly pronouncing an extended "a" vowel) were performed, in silent condition. The measurements were conducted in sitting body posture. The experimental procedures involving human subjects described in this paper were approved by the Institutional Review Board.

While the signals were collected, the offline data analysis was performed. First, we carried out episode segmentation using the single-threshold technique. We applied a visually determined threshold to the changes in standard deviation of the signal over time. Next, for each specified episode, the set of parameters was provided.

We calculated five parameters for the piezoelectric signal and the same five for the envelope of the sound signal. They were based on the shapes of the determined signals and are listed below:

- the mean value of the episode signal (*1st*)
- the standard deviation of the episode signal (*2nd*)
- the area under the mean value (*3rd*)
- the percentage of the samples' values greater than the mean value (*4th*)
- the ratio of the mean values of the episode's first and second halves (*5th*).

In addition, for the microphone signal, we used spectral parameters derived from mel-frequency cepstrum analysis and the linear predictive coding. *MFCA* is a procedure based on the determination of the inverse transform of signal spectrum logarithm. The results are presented in a *mel* scale, which relates them to subjective sound perception - this is achieved using a dedicated filter bank. *MFCA* provided the set of 13 parameters created from the coefficient matrix and calculated in subsequent frames of the episode's duration as ratios of the sum of the specific frame values to the sum of all frames values. Nine *LPC* parameters were simply the 8th-order model coefficients. Therefore, total number of parameters calculated for each episode was 32.

Differentiation between cough episodes and other recorded signals was achieved using the artificial neural network, a multilayer perceptron. We tested different numbers of input parameters (16, 24 and all 32), reduced using the *mRMR* algorithm [19] after parametrization procedure. In the diminished set of 16 parameters all time-related parameters of sound envelope, mean and standard deviation of the vibration signal, first 3 parameters of *LPC* model and first 6 coefficients estimated from *MFCA* analysis (for lower frequencies) were used. We added third and fourth vibration-related parameters and remaining *LPC* ones to them to obtain the set of 24 parameters.

In the learning phase, we classified the input vectors into 4 output states: cough, single hand clap, speech and shout.

In order to test classifiers we used cross-validation scheme with 10 repetitions, with random data division into training and testing parts with relative proportions of 0.7/0.3. As the speech episodes were significantly more numerous than the others (24 phonemes), we chose for training only those phonemes whose waveforms showed the greatest mutual differences in Polish pronunciation ("h", "j", "k", "s", "y"). The mean accuracy for considered classifier were calculated, and for the best one we added the information about its standard error, sensitivity and specificity.

Various network topologies were considered. We performed the analysis for two-layer and three-layer perceptrons with different numbers of neurons in the hidden layers. The numbers were chosen arbitrarily. The training technique was Levenberg-Marquardt back-propagation.

We determined overall accuracies as statistics of the results for the 10 neural network classifiers for every considered topology and every number of input parameters.

Signal processing, parameter calculations, neural network training and validation were performed using the MATLAB software with corresponding toolboxes.

III. RESULTS

For each subject, the signals of imitated cough, single hand clap, speech and shout were recorded. For further analysis, we used the data obtained after pre-processing and features extraction. Fig. 2. presents the sample waveforms of the unprocessed sound signal of cough, and Fig. 3. single hand clap.

We carried out mel-frequency cepstrum analysis for the microphone signals, the sample results of which are presented in Fig. 4. The top subfigure shows the course of an episode of cough in time, and the results of the analysis are presented below in the form of the logarithm of the signal spectrum and the calculated mel-frequency cepstrum coefficients.

The results of the assessment of classification accuracy are shown in Table I, in the form of mean value and standard deviation for the 10 iterations of cross-validated data for

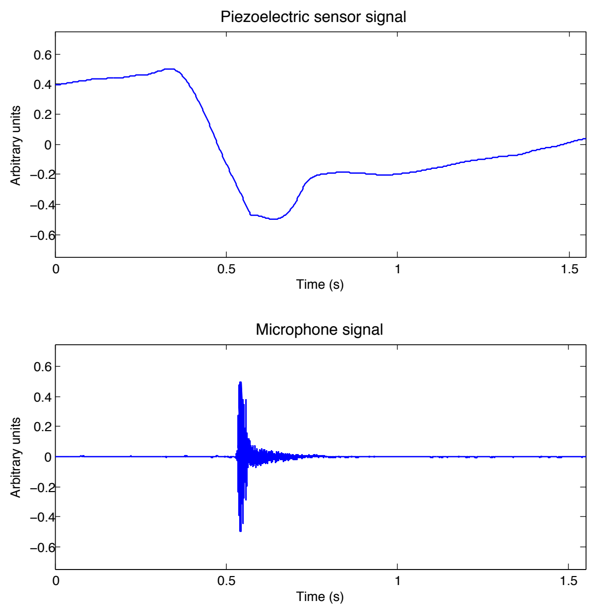

Fig. 2. An example of the recorded signals of imitated cough

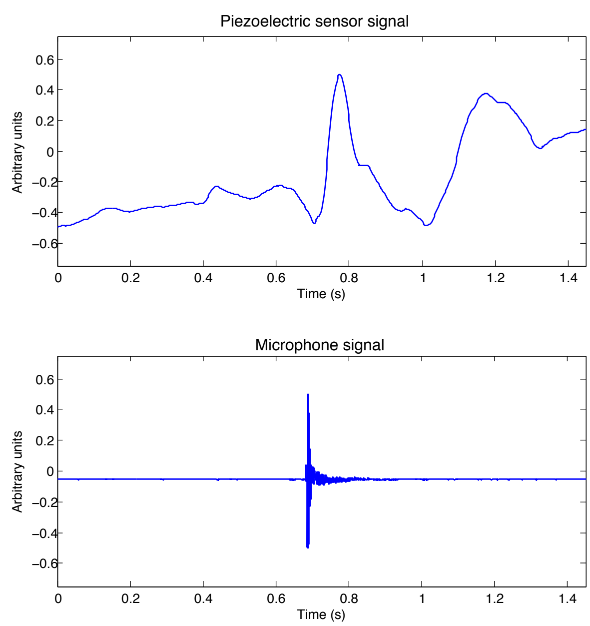

Fig. 3. An example of the recorded signals of single hand clap

every neural network topology and every number of input parameters.

For the best classifier, dealing with 32 input parameters and consisting of two hidden layers, of which each had 32 neurons, we obtained *77%* sensitivity and *97%* specificity for cough/other differentiation. Standard error equaled *0.04*.

IV. DISCUSSION

In order to record cough ”events” two types of sensors (microphone and piezoelectric one) were used. This choice is justified by the nature of the received vibroacoustic signals (e.g. behavior of the chest movements during cough). It

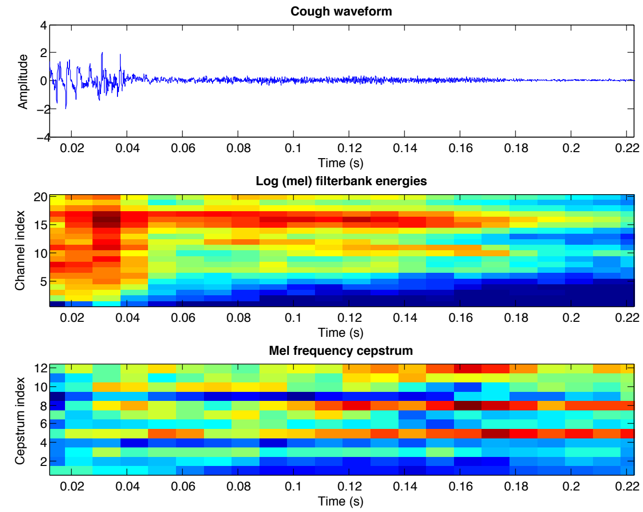

Fig. 4. The results of the mel-frequency cepstrum analysis for imitated cough

TABLE I

SUMMARY OF THE ACCURACY STATISTICS FOR A MULTI-LAYER ARTIFICIAL NEURAL NETWORK FOR EVERY CLASSIFIER TOPOLOGY AND EVERY INPUT PARAMETER VECTOR; A SINGLE NUMBER IN THE FIRST COLUMN DENOTES THE NUMBER OF NEURONS IN A SINGLE HIDDEN LAYER; TWO NUMBERS CORRESPOND TO THE NUMBERS OF NEURONS IN DOUBLE HIDDEN LAYERS

Network topology	Input parameters	Accuracies statistics
16	16	75.0 ± 18.2 %
32	16	87.2 ± 11.3 %
16,16	16	84.2 ± 13.6 %
32,32	16	84.4 ± 12.0 %
24	24	84.6 ± 14.0 %
32	24	87.2 ± 11.9 %
24,24	24	90.4 ± 9.8 %
32,32	24	92.2 ± 9.5 %
32	32	85.7 ± 13.5 %
32,32	**32**	94.5 ± 6.4 %

seems that the number of sensors is sufficient, taking into account the simplicity and accuracy of measurements.

The use of time parameters based on the waveforms is justified because of the differences in signals’ shapes observed during visual, exploratory data analysis. Both the mel-frequency cepstrum and *LPC* parameters have already been used to describe cough signals [8], [12]. However, in our analysis, we used a different combination of input feature vectors. Moreover, calculating the generalized 13 parameters for the episode allowed reduction of the number of *MFCA* coefficients.

Signals of cough, as well as other classified states, were obtained in laboratory conditions, with relatively similar duration of different episodes. A limitation of the study is that the identification of the cough ”events” was carried out on signals which did not differ each other much and did not contain the environmental noise component, which is present in ambulatory conditions. Therefore, it seems necessary to perform testing of the classifiers with imitated or natural noise outside the laboratory in the future research.

In addition, due to the preliminary nature of the measurements, only healthy people participated in the study, thus the cough episodes were imitated. The number of subjects taken into account should also be greater to make investigation stronger in terms of accuracy, sensitivity and specificity calculations.

The relatively low sensitivity of the study drew attention to the need to evaluate other parameters of the signals (especially estimated from piezoelectric sensor).

The accuracies for larger network topologies were definitely smaller than for the best possible one included in the Table I. Furthermore, the computational times of training procedure were disproportionately longer. Therefore, we did not consider them.

We did not perform the analysis, what is the impact of variance of events duration, however in our opinion considered parameters were not very sensitive to episode length.

Although the method seems to provide promising results, further measurements should take into account measurements performed outside the laboratory, for a longer time, and for a greater number of considered output states.

V. CONCLUSIONS

The main objective of this work was to assess the accuracy of a system for cough episode detection using a vibroacoustic sensor for recording and artificial neural networks as classifiers. Despite the relatively small number of surveyed people and the limitations mentioned above, we obtained classification accuracy of *95%* for signals undisturbed by background noise. Moreover, analysis of the accuracies of classification suggests that the selection of measurement sensors was sufficient to assess the incidence of cough.

The analysis of accuracies for selected network topologies and the reduced number of parameters showed that the methods used for parameterization could achieve accuracy in range of *80-95%* and that accuracy increased with the number of the considered parameters. However, further studies with greater numbers of subjects and with recording of noise could improve the assessment.

ACKNOWLEDGMENT

This study was supported by the research programs of institutions the authors are affiliated with.

REFERENCES

[1] A. H. Morice, G. A. Fontana, M. G. Belvisi, S. S. Birring, K. F. Chung, P. V. Dicpinigaitis, J. A. Kastelik, L. P. McGarvey, J. A. Smith, M. Tatar and J. Widdicombe, ERS guidelines on the assessment of cough, European Respiratory Journal, no. 29, pp. 1256–1276, 2007.

[2] T. Drugman, J. Urbain, N. Bauwens, R. Chessini, C. Valderrama, P. Lebecque and T. Dutoit, Objective Study of Sensor Relevance for Automatic Cough Detection, IEEE Journal of Biomedical and Health Informatics, vol. 17, no. 3, pp. 699-707, 2013.

[3] S. Ranjani, V. Santhiya and A. Jayapreetha, A Real Time Cough Monitor for Classification of Various Pulmonary Diseases, in Proc. Third International Conference on Emerging Applications of Information Technology (EAIT), 2012, pp. 102-105.

[4] M. Krajnik, I. Damps-Konstańska, M. Grabowski, P. Pozański and E. Jassem, An evaluation of cough-generated sound and body surface vibration by the MEPIM Vibroacoustic System of Cough Registration (MEPIM VSCR), Advances in Palliative Medicine, vol. 7, pp. 3–8, 2009.

[5] A. Spinou and S. S. Birring, An update on measurement and monitoring of cough: what are the important study endpoints?, J Thorac Dis, vol. 6, Suppl 7., pp. 728–734, Oct 2014.

[6] M. A. Coyle, D. B. Keenan, L. S. Henderson, M. L. Watkins, B. K. Haumann, D. W. Mayleben and M. G. Wilson, Evaluation of an ambulatory system for the quantification of cough frequency in patients with chronic obstructive pulmonary disease, Cough Journal, vol. 1, no. 3, 2005.

[7] J. Smith, J. Earis and A. Woodcock, Establishing a gold standard for manual cough counting: video versus digital audio recordings, Cough Journal, vol. 2, no. 6, 2006.

[8] S. Barry, A. Dane, A. Morice and A. Walmsley, The automatic recognition and counting of cough, Cough Journal, vol. 2, no. 8, 2006.

[9] M. Grabowski, P. Pozański, I. Damps-Konstańska, M. Krajnik, E. Jassem, Analiza sygnałów wibroakustycznych generowanych podczas kaszlu w przewlekłych chorobach układu oddechowego, Acta Bio-Optica et Informatica Medica, vol. 16, pp. 87-90, 2010.

[10] M. Barlow, V. Varadan, J. Di and K. Phan, A Compact, Self-Contained Cough Monitoring System, in Proc. IEEE Region 5 Technical Conference, 2007, pp. 167-170.

[11] I. M. Paul, K. Wai, S. J. Jewell, M. L. Shaffer and V. V. Varadan, Evaluation of a new self-contained, ambulatory, objective cough monitor, Cough Journal, vol. 2, no. 7, 2006.

[12] S. Matos, S. Birring, I. Pavord and D. Evans, An Automated System for 24-h Monitoring of Cough Frequency: The Leicester Cough Monitor, IEEE Transactions on Biomedical Engineering, vol. 54, no. 8, pp. 1472-1479, 2007.

[13] Y. Xiao, D. Carson, L. Boris, J. Mabary, Z. Lin, F. Nicodeme, M. Cuttica, P. J. Kahrilas, J. E. Pandolfino, The acoustic cough monitoring and manometric profile of cough and throat clearing, Dis Esophagus, vol. 27, no. 1, pp. 5-12, 2014.

[14] S. Birring, T. Fleming and S. Matos, The Leicester Cough Monitor: preliminary validation of an automated cough detection system in chronic cough, European Respiratory Journal, no. 31, pp. 1013–1018, 2008.

[15] P. Moradshahi, H. Chatrzarrin and R. Goubran, Improving the Performance of Cough Sound Discriminator in Reverberant Environments using Microphone Array, in IEEE Instrumentation and Measurement Technology International Conference, 2012, pp. 20-23.

[16] S. M. Stec, E. M. Grabczak, P. Bielicki, B. Zaborska, R. Krenke, T. Kryński, M. Dabrowska, J. Domagała-Kulawik, A. Domeracka-Kolodziej, A. Sikorska, P. Kulakowski, R. Chazan, Diagnosis and Management of Premature Ventricular Complexes-Associated Chronic Cough, Chest Journal, no. 135, pp. 1535-1541, 2009.

[17] W. Thorpe, M. Kurver, G. King, C. Salome, Acoustic analysis of cough, in Proc. Seventh Australian and New Zealand Intelligent Information Systems Conference, 2001, pp. 391-394.

[18] S.H. Shin, T. Hashimoto and S. Hatano, Automatic detection system for cough sounds as a symptom of abnormal health condition, IEEE Transactions of Information Technology in Biomedicine, vol. 13, no. 4, pp. 486-493, 2009.

[19] H. Peng, F. Long and C. Ding, Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27, no. 8, pp. 1226-1238, 2005.