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

Lung Water Detection using Acoustic Techniques

Feng Yang, *Member, IEEE*, Wee Ser, *Senior Member, IEEE*, Jufeng Yu, David Chee-Guan Foo, Daniel Poh Shuan Yeo, Pow-Li Chia, and Jennifer Wong

Abstract **– The presence of an excessive amount of water in lung is a sign of pulmonary edema which can be caused by heart failure. The current solutions for lung water detection involve the use of X-ray, CT scan or serum biomarkers, which require bulky and expensive instruments as well as long measurement duration. This paper reports on a study conducted on the use of a different sensing modality to detect the presence of water in lung. The main contributions of the paper are twofold: 1) we propose to employ acoustic (or sound) based techniques for lung water detection. The design is simple and can be implemented on a portable or wearable system; 2) we establish the feasibility of sound-based techniques for lung water detection, by carrying out experimental studies using four feature extraction methods combined with two classification methods. The findings of this study will be beneficial to the design of portable devices for rapid and accurate lung water detection.**

I. INTRODUCTION

Pulmonary edema can be caused by many critical conditions, such as acute lung injury and heart failure which are either cardiogenic or non-cardiogenic. A common sign of pulmonary edema is the increased extravascular lung water that is usually accompanied by a high mortality rate. Consequently, reliable tools for monitoring and detecting lung water are increasingly needed in modern intensive care and clinic therapy. Existing lung water detection techniques can be classified into two types: Biochemical techniques and Engineering techniques. Examples of the former techniques include double indicator dilution method, single-indicator transpulmonary thermodilution method, blood gases method, re-breathing method, serum biomarkers, etc. [1]; while some examples of the latter techniques are X-ray, CT scan, microwave or ultrasonography, magnetic resonance imaging, Impulse Radio Ultra Wide band (IRUWB) radar, etc. [2,3]. The common drawbacks of the abovementioned techniques include the use of bulky equipments (i.e. not portable) and a relatively long measurement duration (i.e. cannot be measured readily). This is undesirable for some time-critical operations. Some of the equipments are expensive and their operations require trained professionals. These make them not practical for use outside the hospital (e.g. ambulatory applications). Another more practical and commonly used approach is auscultation where a stethoscope is used (by the clinician) to listen to the lung sound. The drawback of this approach is that, the outcome is not quantifiable and the accuracy depends highly on the experience of the clinician using it.

This work was supported by the Nanyang Institute of Technology in Health and Medicine (NITHM) and the VALENS Centre of Excellence at the School of EEE at Nanyang Technological University in Singapore.

Feng Yang, Wee Ser (correspondence author) and Jufeng Yu are with the School of EEE, Nanyang Technological university, Singapore 639798 (phone:+6567905951; fax +65-67912383; e-mail: {yangfeng, ewser, jfyu}@ntu.edu.sg).

David Chee-Guan Foo, Daniel Poh Shuan Yeo, Chia Pow-Li, and Jennifer Wong are with the department of Cardiology, Tan Tock Seng Hospital, Singapore 308433, email: ({david_foo, daniel_yeo, pow_li_chia, jennifer_wong}@ttsh.com.sg).

Lung or breath sounds are produced during the process of breathing due to the flow of air through abruptly branching respiratory passage. Analysis of lung sounds has been used to detect the presence of some respiratory disorders symptoms such as snoring, wheezing, etc. [4,5]. However, there are no (or very few) works reported on the use of lung sound to detect water in lung. The closest work is the paper published by Mulligan et al. in [6]. In [6], a system consisting of a speaker (that injects white Gaussian noise into the mouth) and four electronic stethoscopes (that record sound signals from the chest wall) was used to measure changes in the distribution of lung fluid in the respiratory system.

This study proposes to detect lung water based on respiratory sounds collected via a stethoscope and tries to establish the feasibility of using lung sounds for detecting water in lungs. The sound-based approach proposed by this study has the advantage of being able to overcome the drawbacks of existing techniques as described above. Such an advantage is particularly important for time-critical situation when the clinicians need to have the information quickly during episodes of acute breathlessness.

II. MATERIALS AND METHODOLOGIES

A. Overview of the Lung Water Detection System

The main purpose of this study is the establishment of the feasibility of detecting water in lung using a sound sensor. The process involved in our proposed method is shown in Figure 1.

In Figure 1, there are six functional blocks namely data collection, data pre-processing, feature extraction, feature selection, classification and performance evaluation. The designs of these blocks are elaborated below.

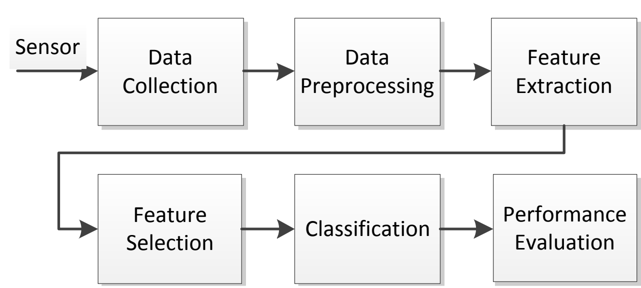

Fig. 1. The process involved in our proposed lung water detection method

B. Data Collection and Pre-processing

Real data of 12 volunteers (including 6 patients and 6 normal persons) were collected in practical environment (i.e. either in the ward of a local hospital or a research lab in the university). The data collection was conducted with approval from the Institutional Review Board (IRB) and the informed consent of the patients.

All the volunteers were selected by experienced clinicians working in this project.

Respiratory sounds were first recorded from the back of the test subject using an electronic stethoscope which is connected to a laptop computer. The A/D conversion is performed by the computer and the recorded data are time-tagged. There was no direct contact between the stethoscope and the volunteers' skin as the stethoscope was placed over the clothing worn by the volunteers. The sampling rate is 8000 Hz which is sufficient to capture the bulk of the energy of the lung sounds. The recording durations are around 40 seconds for the patients and about 2 minutes for healthy subjects. This is to reduce the participating time needed from the patients.

As the data recorded are of relatively good quality, minimum or no pre-processing was performed in our experiment.

C. Feature Extraction

It is well known that discrete time signals or samples cannot be used for classification directly as there are too many sampled values to consider. A common approach is to perform feature extraction from sampled signals so that the signals can be represented by a manageable set of feature values. In this study, four feature extraction techniques are investigated and they are as described below.

1) Mel-Frequency Cepstral Coefficients (MFCCs)

In acoustic signal analysis, Mel-Frequency Cepstrum (MFC) is commonly used to represent the short time power spectrum of the signal. This feature is determined by performing a linear cosine transformation of the log of the power spectrum using a nonlinear MEL scale of frequencies. Mel-Frequency Cepstral Coefficients (MFCCs) [7] are the coefficients of MFC.

A common way to derive MFCCs is by the following four steps [8]:

(i) Compute the Fourier transform and hence the power spectrum of the signal.

(ii) Map the power spectrum onto the MEL scale using a triangular overlapping window.

(iii) Compute the log transform of the powers at each of the MEL frequencies.

(iv) Compute the Discrete Cosine Transform of MEL log powers. The MFCCs are the amplitudes of the resulting spectrum.

2) Perceptual Linear Predictive Coefficients (PLPCs)

Perceptual linear prediction coefficient (PLPC) is similar to MFCC. The main differences between MFCC and PLPC are fourfold [9]:

(1) The shape of the filter-banks in PLPC is trapezoidal, instead of triangular;

(2) PLPC uses equal-loudness pre-emphasis to weight the filter-bank outputs;

(3) PLPC uses cube-root compression instead of logarithmic compression;

(4) PLPC uses a (parametric) linear-predictive model to determine cepstral coefficients, instead of (non-parametric) discrete cosine transform.

More details of PLPC are given in [10,11].

3) Linear Prediction Coefficients (LPCs)

Linear Prediction (LP) is a mathematical operation where the future values of a discrete-time signal are estimated as a linear function of the previous samples. The most common representation is:

$$\hat{x}(n) = \sum_{i=1}^{p} a_i x(n-i) \tag{1}$$

where $\hat{x}(n)$ is the predicted signal value, $x(n-i)$ are the previous observed values, a_i are the predictor coefficients and p is the order.

The error generated by this estimate is:

$$e(n) = x(n) - \hat{x}(n) \tag{2}$$

where $x(n)$ is the true signal value.

There are two widely used methods for estimating the linear prediction coefficients, one is autocorrelation and the other is covariance [12].

4) Wavelet Transform-Based Features (WTFs)

With its first introduction by Grossmann and Morlet [13] in the mid-1980s, wavelet transform has been increasingly applied in many areas such as pattern recognition, processing and synthesizing signals (e.g., speech), image analysis, and so on.

One advantage of Wavelet Transform is that the signal could be simultaneously analyzed in physical (time, coordinate) and frequency spaces. Wavelet transform decomposes the signal into approximations and detailed space represented by wavelet coefficients in a series of sub-bands [14]. The coefficients can then be used for feature extraction. In this study, the following features are extracted [15]:

(1) Maximum value of the wavelet coefficients in each sub-band.

(2) Minimum value of the wavelet coefficients in each sub-band.

(3) The standard deviation of the wavelet coefficients in each sub-band.

(4) The mean of the absolute values of the coefficients in each sub-band.

(5) The average power of the wavelet coefficients in each sub-band.

(6) The mean of the coefficients in each sub-band.

(7) The ratio of the mean of the absolute values of the coefficients of a sub-band and that of the adjacent sub-bands.

D. Feature Selection

Feature selection, also known as variable selection or attribute selection, is commonly used as a dimension reduction technique [16]. The purpose of feature selection is to select the most informative features for a specific purpose, e.g. detecting the presence of water in lungs.

In this study, we employ a simple but effective feature selection method -- "Fisher's Ratio" which measures the linear discriminating power of a feature X_j as the ratio of squared inter-class divergence to intra-class spread [16]:

$$\mathrm{FR}(\mathrm{X}_j) = \frac{(m_{j(1)}-m_{j(2)})^2}{\sigma_{j(1)}^2+\sigma_{j(2)}^2} \quad (3)$$

where $m_{j(c)}$ and $\sigma_{j(c)}^2$ are the sample mean and variance of feature X_j respectively in a class c, for $c = 1, 2$. The larger the FR value, the more discriminative the feature is.

E. Classification

In pattern classification, a classifier can be viewed as a mapping from a pattern space to a class-label space, or more specifically, a decision boundary in the pattern space which segments the pattern space into meaningful regions. In this study, we employ two popular classifiers: Support Vector Machines (SVMs) and k-Nearest-Neighbor (kNN). The former is a linear classifier while the latter is a nonlinear classifier.

F. Performance Evaluation

The performance of a pattern classification method is usually assessed based on the classification accuracy (ACC). For medical studies, the sensitivity or true positive rate (TPR) and the specificity or true negative rate (TNR) are often used instead.

In order to determine the classification performance of the methods investigated, the data are divided into training data and testing data. A classifier is first trained on the training data and then tested on the testing data. There are several validation techniques for the estimation, such as the "leave-one-out" method, the "repeated k-fold cross validation" method and the "bootstrap" method. Although the cross validation is more widely used, it may exhibit large variability when the sample size is small [17]. In this study, the .632 bootstrap method [18] was used. For the purpose of reliable estimations, the number of repeats was set to about 100 times in our study.

III. Experimental Study

A. Experimental Setup

There are 12 volunteers in our study. Two sets of data, one from the left lung and the other from the right lung, were recorded for each volunteer, with durations ranging from 40 seconds to 2 minutes. As the duration of a human's respiratory period (cycle) is approximately 3 to 4 seconds, we segment every record into segments of 4 seconds each (so that each segment contains one breathing cycle). Our study shows that, the results are not affected by using different parts of the breathing cycle and there is therefore no need to segment the signals from the start of the breathing cycle. The segmentation produces a total of 136 segments of signals for patients and 165 segments of signals for healthy subjects.

The segments of the same set of records are placed randomly as either the training sets or the testing sets of samples.

The parameters of the classifiers were set to C=0.01 for SVM and k=3 for k-nearest-neighbor classifier. The order of LPC is set to $p = 10$.

B. Results and Discussion

Fig. 2 shows the classification performance which includes the classification accuracy (ACC), sensitivity (TPR) and specificity (TNR), of the four feature extraction methods and the two classifiers. Table I gives the best classification accuracy (in percentage) where the numbers shown in the bracket are the number of features used.

From the classification results shown in Fig. 2 and Table I, we observe that good classification performance can be achieved, especially by MFCC, PLPC and WTF (Fig. 2 (a), (b) and (d)). On TPR and TNR, we observe that TPR are better than TNR when MFCC and PLPC are used (Fig. 2 (a) and (b)), and the reverse is true when LPC and WTF are used (Fig. 2 (c) and (d)).

Another interesting observation is that kNN is observed to perform better than SVM.

TABLE I. THE BEST CLASSIFICATION ACCURACY

MFCC	SVM	90.7 (1)	PLPC	SVM	92.0 (1)
	kNN	95.7 (5)		kNN	94.7(11)
LPC	SVM	66.6 (5)	WTF	SVM	85.0 (4)
	kNN	85.6 (9)		kNN	92.5 (4)

IV. Conclusion

In this paper, we propose to use the acoustic (or sound) approach to detect the presence of water in lung (the sound sensor used for the experiment is a stethoscope). Four feature extraction methods and two classification methods have been investigated. The results obtained using data collected from patients and healthy subjects are very promising, with the best classification accuracy exceeding 95%, the best sensitivity exceeding 96%, and the best specificity exceeding 97%. These results show clearly the feasibility of detecting lung water using the lung sounds analysis approach. The sound-based approach proposed in this study will be beneficial to the design of portable devices for rapid and accurate lung water detection.

References

[1] J. Lemson, P. Merkus and J. G. van der Hoeven ,"Extravascular lung water index and global end-diastolic volume index should be corrected in children," *Journal of Critical Care*, vol. 26 , pp. 432.e7- 432.e12, 2011.

[2] R. Gagarin, H. Youn, N. Celik and M. Iskander, "Noninvasive microwave technique for hemodynamic assessments," *APSURSI 2010*, Toronto, pp. 1-4.

[3] X. Li, E. Pancera, L. Zwirello, H. Wu and T. Zwick, "Ultra wideband radar for water detection in the human body," *German Microwave Conference*, 2010, Berlin, pp. 150 – 153.

[4] J. Zhang, W. Ser and Y. Goh, "A novel respiratory rate estimation method for sound-based wearable monitoring systems," *EMBS2011*, Boston, USA, pp. 3213 – 3216.

[5] J. Abhishek and V. Jithendra, "Lung sound analysis for wheeze episode detection," EMBS2008, Vancouver, Canada, pp. 2582 - 2585.

[6] K. Mulligan, A. Adler and R. Goubran, "Detecting regional lung properties using audio transfer functions of the respiratory system," *EMBC2009*, Minneapolis, USA, pp. 5697 – 5700.

[7] R. Hasan, M. Jamil, G. Rabbani and S. Rahman, "Speaker identification using mel frequency cepstral coefficients," *ICECE 2004*, 28-30 December 2004, Dhaka, Bangladesh, pp. 565-568.

[8] M. Sahidullah, S. Goutam, "Design, analysis and experimental evaluation of block based transformation in MFCC computation for speaker recognition". Speech Communication, 54 (4): 543–565.

[9] G. Zweig and M. Picheny, "Advances in large vocabulary continuous speech recognition," *Signal Processing,*, Vol. 60, pp. 249–291, 2004.

[10] H. Hermansk, "Perceptual linear predictive (PLP) analysis of speech," *J. Acoust Soc. Am.,* vol. 87, no. 4, pp.1738-1752, 1990.

[11] Florian Hönig et al., "Revising perceptual linear prediction (PLP)," *Interspeech*, pp. 2997-3000, 2005.

[12] Tamanna Islam, *Interpolation of Linear Prediction Coefficients for Speech Coding*, PhD thesis, McGill University, 2000.

[13] A. Grossmann and J. Morlet, “Decomposition of hardy functions into square integrable wavelets of constant shape,” SIAM J. Math. Anal, vol. 1, pp. 723–736, 1984.
[14] S. Mallat, “A theory for multiresolution signal decomposition: the wavelet representation,” *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, no. 7, pp. 674–693, 1989.
[15] J. Hadjileontiadis, *Lung Sounds: An Advanced Signal Processing Perspective*, Morgan & Claypool, 2008.
[16] I. Guyon and A. Elisseeff. “An introduction to variable and feature selection,” *J. Mach. Learn. Res.*, vol. 3, pp. 1157–1182, March 2003.
[17] U. M. Braga-Neto and E. R. Dougherty, “Is cross-validation valid for small-sample microarray classification?” *Bioinformatics,* vol. 20, no. 3, pp. 374–380, Feb 2004.
[18] M. R. Chernick, *Bootstrap Methods: A Guide for Practitioners and Researchers*, 2nd ed. New York: John Wiley & Sons, October 2007.

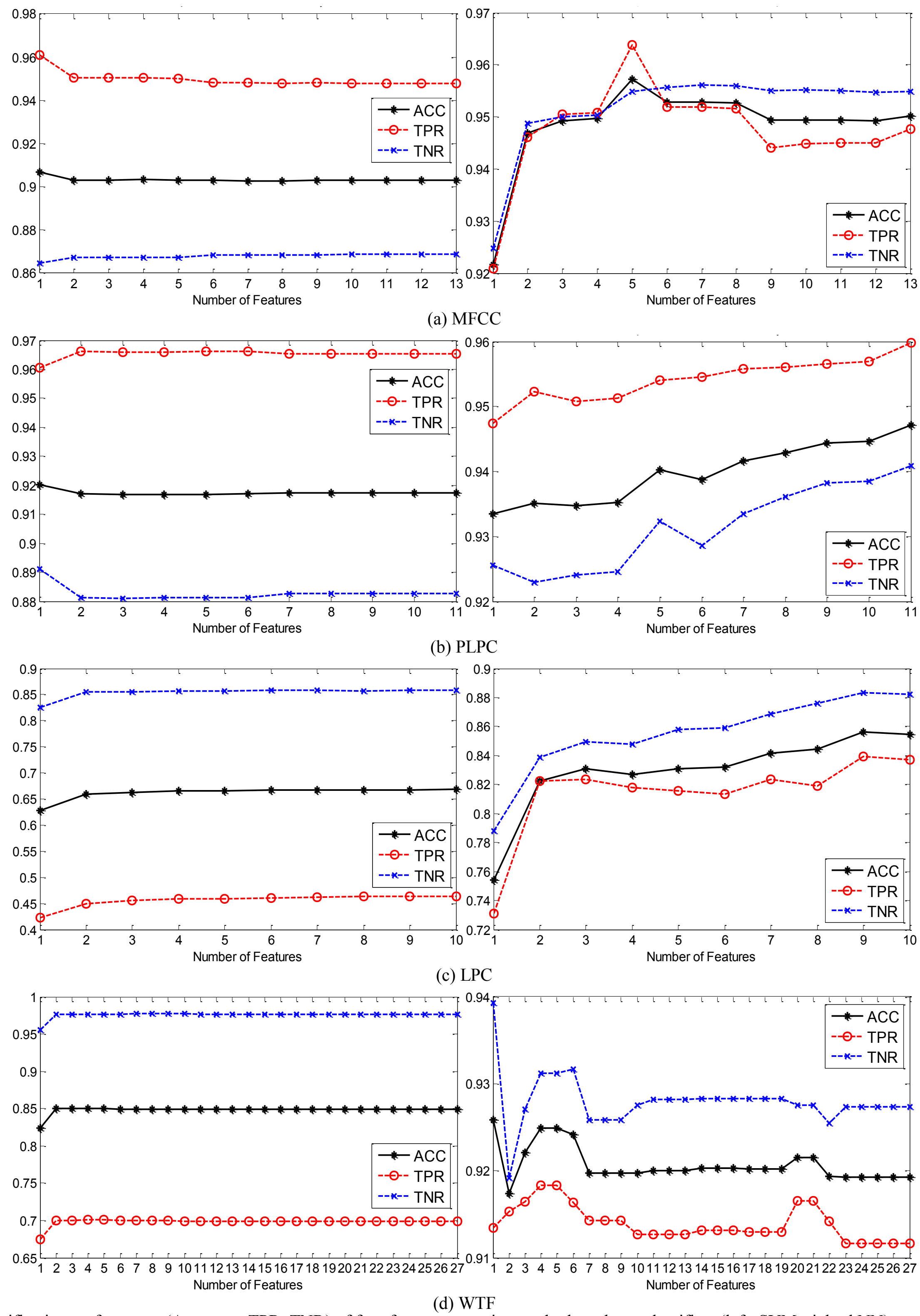

Fig. 2. Classification performance (Accuracy, TPR, TNR) of four feature extraction methods and two classifiers (left: SVM, right: kNN). y-axis gives the performance (in %) and x-axis represents the number of features used in the classification.