Biomedical Signal Processing and Control 21 (2015) 126–136

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

Automatic cough segmentation from non-contact sound recordings in pediatric wards

Yusuf A. Amrulloh[a,c], Udantha R. Abeyratne[a,*], Vinayak Swarnkar[a], Rina Triasih[b], Amalia Setyati[b]

[a] School of Information Technology and Electrical Engineering, The University of Queensland, Brisbane, Australia
[b] Department of Child Health, Sardjito Hospital, Gadjah Mada University, Yogyakarta, Indonesia
[c] Department of Electrical Engineering, Faculty of Industrial Engineering, Islamic University of Indonesia, Yogyakarta, Indonesia

ARTICLE INFO

Article history:
Received 13 October 2014
Received in revised form 20 March 2015
Accepted 4 May 2015
Available online 23 June 2015

Keywords:
Pediatric respiratory diseases
Cough
Automatic cough segmentation

ABSTRACT

Cough is a common symptom of almost all childhood respiratory diseases. In a typical consultation session, physicians may seek for qualitative information (e.g., wetness) and quantitative information (e.g., cough frequency) either by listening to voluntary coughs or by interviewing the patients/carers. This information is useful in the differential diagnosis and in assessing the treatment outcome of the disease. The manual cough assessment is tedious, subjective, and not suitable for long-term recording. Researchers have attempted to develop automated systems for cough assessment but none of the existing systems have specifically targeted the pediatric population. In this paper we address these issues and develop a method to automatically identify cough segments from the pediatric sound recordings. Our method is based on extracting mathematical features such as non-Gaussianity, Shannon entropy, and cepstral coefficients to describe cough characteristics. These features were then used to train an artificial neural network to detect coughs segment in the sound recordings. Working on a prospective data set of 14 subjects (sound recording length 840 min), proposed method achieved sensitivity, specificity, and Cohen's Kappa of 93%, 98%, and 0.65, respectively. These results indicate that the proposed method has the potential to be developed as an automated pediatric cough counting device as well as the front-end of a cough analysis system.

1. Introduction

Cough is a defense mechanism of the body to clear the respiratory tract from foreign materials which are inhaled accidentally or produced internally by infections [1]. It is a common symptom appearing in early to mid stages of respiratory diseases such as pneumonia, the leading cause of death in children less than five years of age. It has been estimated that pneumonia causes over 1.6 million deaths in this group per year [2], with more than 97% of cases occurring in the developing countries [3]. The world health organization (WHO) also reported that in those countries, pertussis (whooping cough) has also become one of the major childhood morbidities with an estimated 50 million cases and 300,000 deaths every year [4].

Even though cough is common in respiratory diseases and considered as an important clinical symptom, there is no gold standard to assess it. In a typical consultation session, physicians may listen to several episodes of spontaneous or voluntary coughs, to obtain qualitative information such as the "wetness" of a cough. Such qualitative information is extremely useful in diagnosis as well as the treatment of respiratory diseases.

During consultation sessions physicians may also seek quantitative information on coughs, such as the frequency of occurrence of cough events over a given time interval. This information can be used to determine the nature (e.g., acute, chronic) and the severity of coughs as well as to monitor the efficacy of a treatment [5]. However, to obtain this information, physicians heavily rely on subjective reports of patients or their carers. There is a great need for an automated device capable of counting the number of coughs, especially in childhood diseases. More importantly, technology capable of automatically extracting cough events from pediatric recordings is urgently needed in order to facilitate the diagnosis of diseases such as pneumonia [6,7].

Several approaches have been taken to develop automated cough counting systems (e.g., Hull Automatic Cough Counter

* Corresponding author at: School of Information Technology and Electrical Engineering, The University of Queensland, GP South Building (Building 78), Staff House Road, Brisbane, Australia. Tel.: +61 7 33469063; fax: +61 7 33654999.
E-mail address: udantha@itee.uq.edu.au (U.R. Abeyratne).

http://dx.doi.org/10.1016/j.bspc.2015.05.001

(HACC), Leicester Cough Monitor (LCM), LifeShirt, VitaloJAK, and PulmoTrack). The performances of these devices are varied. The HACC claimed 80 percent sensitivity and 96 percent specificity [8]. The figures for LifeShirt, Pulmotrack, LCM, and Vitalojak are (78%, 99%), (94%, 96%), (85.7%, 99.9%), and (97.5%, 97.7%) respectively [9–12]. To extract the sound events, HACC computed the standard deviation of cough sound intensities [8], while VitaloJAK [13] and LCM [11,14] used sound intensities themselves. This makes these methods susceptible to variations in recording conditions, such as the distance from the microphone to the patient, sensitivity of the recording instruments and the sound level of the coughs being recorded. Both LifeShirt [15] and PulmoTrack [16] described cough as having a distinguishable humped structure in the waveform. Then they identified coughs by counting the number and/or measuring the slope of these humps. None of these commercial devices have been tested on pediatric populations or subjects with diseases such as pneumonia. In such subjects, the cough sound intensities and waveform-shapes may vary widely. Therefore intensity or simple waveform-shape based methods are unlikely to be optimal and are less likely to be robust in field use.

Cough recording on children, especially the younger ones, pose several additional challenges. Younger children are unable to produce voluntary coughs upon request. Any method targeting pediatric populations should be capable of using spontaneous coughs recorded over a period of interest. In pediatric recordings, crying, vocalization, and grunting are found abundantly, intermixed with cough sounds. Consequently, technology developed for adults are unlikely to be optimal for use on children.

Existing commercial cough counting devices such as LifeShirt, Vitalojak, and Pulmotrack employ contact sensors. While the use of contact sensors may have some advantages, they also carry several drawbacks. Contact sensors, compared to non-contact (free-air) microphones are robust against background sound propagated through air. However, they are more vulnerable to sound conducted through tissue and bones; spurious rubbing sounds due to sensor movement can also be an issue. In infectious diseases, elaborate efforts are needed to avoid cross contamination of patients through contact instrumentation. Furthermore, in pediatric subjects, contact sensors can also be difficult to attach because of patient discomfort. In this paper we address these issues and propose a novel technology for the automated segmentation of cough events from recordings obtained using non-contact microphones in a pediatric ward. In particular:

- We design our algorithms to target the pediatric population (age <6 years), addressing a fundamental gap in current technology.
- We develop a method for the segmentation of cough events, with algorithms capable of discounting background sounds such as crying, vocalization, and grunting.
- We develop new techniques that are robust against the variation of cough sound intensity levels and waveform-shapes. This is an unprecedented approach in this field, to the best of our knowledge.

The method has the potential to be developed as an automated cough counting device as well as the front-end of a cough based diagnostic system.

2. Materials and methods

2.1. Cough recording protocol

The cough recording system consisted of a low-noise microphone having a cardioid beam pattern (Model NT3, RODE®, Sydney, Australia), followed by a pre-amplifier and an A/D converter (Model Mobile Pre-USB, M-Audio®, CA, USA). The output of the Mobile Pre-USB was connected to the USB port of a laptop computer. The nominal distance from the microphones to the mouth of subjects was 50 cm. The actual distance could vary from 40 cm to 100 cm due to the subject movement. We kept the sampling rate at $F_s = 44.1\ k$ samples/s and 16-bit resolution to obtain the best sound quality.

Fig. 1. Cough recording system set up.

The data for this work were recorded at Sardjito Hospital, Yogyakarta, Indonesia, from pediatric patients admitted on respiratory complaints. The inclusion criteria used in the recruitment was patients with at least two of the following symptoms: cough, sputum, breathlessness, and temperature higher than 37.5 °C. We excluded patients having advanced disease where recovery is not expected, diseases with droplet precautions and patients undergoing ventilation treatment. The recordings were started after physicians had examined the subjects, begun the initial treatment, and informed consent had been completed. We acquired data in the natural hospital environment, without modifying it anyway, other than placing our sound recording system by the bed (see Fig. 1). The duration of recording for each subject was between four and 6 h. The research protocol had received ethics clearances from Sardjito Hospital and The University of Queensland, Australia.

2.2. Database and classification of cough segments

In this paper, we used recordings from 24 pediatric subjects. We divided these subjects into two datasets namely:

(i) *Model design dataset (MDD)*: Dataset MDD denotes the training dataset that was used to train the neural network (NN) classifier. The main criterion in designing MDD is that it should contain the whole range of cough types and their variations as well as non-cough sounds expected in the practical setting. That way, the NN can learn the characteristics of the variety of coughs and learn to differentiate them from non-cough sounds. Data set MDD was designed by manually picking representative cough and non-cough sounds from each subject in the training dataset. All such sound events were then concatenated as a single data stream to form the set MDD. Thus, the dataset MDD is not a natural sequence of sound events, but the combination of a large number of handpicked cough and non-cough sounds. To develop MDD, we took cough and non-cough sound samples from $D_1 = 10$ subjects. The overall length of MDD was 15 min.

(ii) *Prospective study dataset (PSD)*: Dataset PSD denotes the testing dataset and contains the actual sound stream recorded in the hospital. This way, our results on PSD can be taken as a true indication of performance in the clinical environment. Testing set PSD included the first 60 min of the sound streams recorded