MobiCough: Real-Time Cough Detection and Monitoring Using Low-Cost Mobile Devices

Cuong Pham(✉)

Computer Science Department and Machine Learning and Applications Lab,
Posts and Telecommunications Institute of Technology, Km 10, Hadong, Hanoi, Vietnam
cuongpv@ptit.edu.vn

Abstract. In this paper we present MobiCough, a method and system for cough detection and monitoring on low-cost mobile devices in real-time. MobiCough utilizes the acoustic data stream captured from a wirelessly low-cost microphone worn on user's collar and connected to the mobile device via Bluetooth. MobiCough detects the cough in four steps: sound pre-processing, segmentation, feature & event extraction, and cough prediction. In addition, we propose the use of a simple yet effective robust to noise predictive model that combines Gaussian Mixture model and Universal Background model (GMM-UBM) for predicting cough sounds. The proposed method is rigorously evaluated through a dataset consisting of more than 1000 cough events and a significant number of noises. The results demonstrate that cough can be detected with the precision and recall of more than 91 % with individually trained models and over 81 % for subject independent training. These results are really potential for health-care applications acquiring cough detection and monitoring using low-cost mobile devices.

Keywords: Cough detection · Monitoring · Machine learning · Healthcare · Ubiquitous computing · Mobile devices · Acoustic sensors · Universal background model · Gaussian mixture model · Health monitoring

1 Introduction

Cough is a prevalent symptom of many related respiratory diseases from minor ailment to severely lung tuberculosis or chronic cough. According to a report [1] the prevalence of cough-related diseases is around 29 % of hospital consultant episodes. One of the most serious cough-related diseases is pneumonia killing nearly 6.6 million children under five every year [2], and it is the leading disease causing of death in children. Especially, the situation is even worse in Vietnam due to climate characteristics of hot and humid leading to the high rate of pharyngitis and pneumonia infection. Every year, up to 4,500 children died and 2.9 million under five years old infectious pneumonia, adding about 40 children are hospitalized to be treated for respiratory disease everyday [2]. This high rate indicates the lack of local healthcare and an effective monitoring pneumonia method for patients and people who are highly potentially infectious and living in remote areas.

N.T. Nguyen et al. (Eds.): ACIIDS 2016, Part I, LNAI 9621, pp. 300–309, 2016.
DOI: 10.1007/978-3-662-49381-6_29

Monitoring cough is a task that is able to provide cough frequency information to users and doctors. Such information is useful for diagnose and treatment of the cough-related diseases. Traditionally, cough is monitored by manually diarizing from nurses and patients themselves. This would possibly report imprecise cough and whooping frequency as manually counting cough for every time and everywhere is infeasible. Therefore, an alternative is to automatically monitor cough using electronic devices. Recent technologies such as wearable computing can easily allow users to capture cough sounds which are analysed for detecting and counting the cough. Recently, although research on cough detection has made significantly progress such as detection of cough using mobile phones [5, 14] while preserving privacy [5], or analysis sounds from an audio recorder [6, 9, 13]. These works achieved relatively high accuracies of cough detection rate. However, real-time cough detection and monitoring is omitted from these works. As cough relates to the diversity of respiratory diseases, it is understandable that the detection of cough in real-time is crucial for opportune interventions, diagnosis, cure, treatment and even emergency aids made by the doctors. Therefore, in this work we propose a method and system that can detect and monitor cough by analysing sounds from a low-cost, wireless microphone in real-time. In addition, to improve the detection accuracy, we propose the Gaussian Mixture model (GMM) combined with the Universal Background models (UBM), an adaptive version of Gaussian Mixture models with maximum a posteriori scheme, for discriminating cough and noise sounds (i.e. from human's speaking, environment etc.). The proposed method is evaluated on a cough dataset consisting of more than 1,000 cough events, and a significant number of background noise events. The results of this research can be a complementary tool for real-time monitoring pneumonia as well as other respiratory related diseases. In addition, we have implemented the pre-trained GMM-UBM models (the GMM-UBM models after trained with offline data) on the smart phones for real-time cough detection and the monitoring module on the smart phones.

2 Related Work

Automatic cough detection has attracted by researchers, medical experts, and doctors for long years as cough is the most frequent symptom appearing on the people when asking medical advice [5]. Approaches to cough detection can be the use of array of audio sensors [6, 8, 10, 13], a single microphone worn on user's body [4] or mobile phones [3, 14]. Previously, arrays of sensors installed in the environment surroundings are proven effectively for context recognition and situated services [11], while wearable sensors commonly utilized for human activity recognition and fall detection [12]. While the use of either multiple sensors installed in the environment surroundings can possibly be limited by the range of sensing signals (i.e. a room or a house), the mobile phones or wearable sensors, in contrast, can allow users to detect and to monitor coughs at everywhere and at every time.

Multiple acoustic sensors are widely used for cough detection as the cough detector can achieve performance accuracies as high as over 95 %. Work by Drugman, T. et al. [6, 8] for example, proposed a cough detection method based on Artificial Neural

Networks which were trained feature vectors comprising of 222 feature elements, While [6] investigated how the acoustic sensors worn on different positions of user's body can impact on the performance accuracy of the cough detector, work by Vizel E. et al. [7] analyzed acoustic data captured from both microphones worn on user's chest and ambient sensors (installed in the environment surroundings). Similarly, Zheng, S. et al. [12] proposed CoughLoc which analyzes acoustic signals from a non-intrusively wireless sensor network. CoughLoc exploits the location of cough occurrence to enhance the detection accuracy. [8] uses various sensors including Siemens EMT 25 C accelerometer (Siemens); PPG 201 accelerometer (PPG); Sony ECM-T150 electret condenser microphone with air coupler etc. connected to an electret condenser microphone for comparisons of the effectiveness of lung sound transducers. In contrast, cough detection using a single microphone analyses only one audio stream from a microphone worn on chest of the user. For instance, [4] proposed Leicester Cough Monitor (LCM) using an audio recorder on patient's chest. LCM was rigorously evaluated and achieved sensitivity and specificity over 91 % on their (offline) dataset of 15 patients with chronic cough and 8 healthy subjects. Another study [11] proposed hidden Markov models trained with more than 800 min of ambulatory recordings and achieved the detection accuracy rate of 82 % with the false alarm rate is as low as 7 events per hour.

With the incorporation of multi-processors, cache memory, and many sensors such as accelerometer, GPS, gyroscope, digital camera, microphone etc., mobile phones become powerful platforms for health-care applications. A cough detector on mobile phones utilizes audio stream captured by the microphone embedded inside the mobile phone. For example, [3] proposed a cough detection method based on mean decibel energy, component weight features extracted from Fast Fourier transform coefficients from acoustic raw data captured from the mobile phone carried in the participant's shirt pocket or using a neck strap (the phone's microphone facing up in the direction of the mouth). [3] achieved the true positive rate is as high as 92 % while preserving privacy for the users, using a neck strap might be inconvenient for users. Although multiple acoustic sensors might enhance the cough detection accuracy, it might lead to several limitations such as high cost and being limited within the range of the sensing signals, and being uncomfortable and invasive for the users. Mobile device based cough detection, in contrast, is low-cost while cough (and many other diseases) might be monitored every time and everywhere. Therefore, in this investigation we propose the cough detection and monitoring method using mobile devices that is convenient for the users while being cheap. Our proposed method is distinct from multiple sensor based cough detection as we use only one single microphone for capturing the sound from the user while cough detection system runs on a mobile platform. Our proposed method is more convenient for the users than [3, 13, 16] as the users are not be required to wear the phone on the user's chest area while the detection performance is not affected by phone's positions.

3 Real-Time Cough Detection & Monitoring

As depicted in the Fig. 1, cough is detected in 4 steps: *sound pre-processing*, *audio segmentation*, *feature & event extraction,* and *cough prediction*.

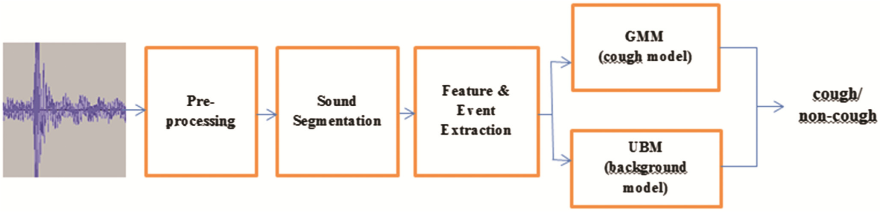

Fig. 1. Cough detection system

3.1 Hardware

The hardware used in this study includes a tiny wifi microphone and a mobile phone platform. The microphone is a low-cost, mini wireless Bluetooth earphone speaker stereo with microphone provided by the OEM (cost at $5.00). The microphone can communicate to the mobile phone via Bluetooth wireless technology. The battery is completely embedded inside the case, and it has a long talking time of 6 h without any external power. With the size of 33 × 12 × 7 mm, the microphone is very easy to be worn on the user's collar for capturing sounds from the user while the user's mobile phone can be flexible to be positioned at hand or pocket.

3.2 Sound Pre-Processing

The sampling rate is down-sampled to 8 kHz. It is noticed that the sampling rate of 8 kHz is good enough for practical applications. After that sound signals are filtered using a low-pass filtering procedure to eliminate background noise and silent. Also, a high-pass filtering with a transfer function $H(x) = 1-\alpha x^{-1}$ where $\alpha = 0.95$ for emphasizing on higher frequency bands of the cough signals.

3.3 Audio Segmentation

Continuously audio stream is segmented into 2-second sliding windows with 50 % overlap between two consecutive sliding windows. A sliding window can contain a silent, a cough, or an unknown event. Silent sliding windows are discarded on the ground using a simple energy feature based threshold. Each sliding window is segmented into 25 ms hamming windows. Therefore, a sliding window comprises of 80 hamming windows, each contains 200 samples. The reason for using small segments such as hamming windows is the acoustic signals are constantly changing, but a short time scale of 25 ms is assumed statistically stationary.

3.4 Feature & Event Extraction

For each hamming window, the features *Mel-frequency Cepstral Coefficient* (MFCC), *Zero-crossing Rate* (ZCR) and *Entropy* are extracted. These features also contain rich sound information that we expected to effectively discriminate cough sounds and others. We describe the feature extraction procedure as followings.

Mel-frequency Cepstral Coefficient (MFCC): known as a feature type widely used in speech recognition, an MFCC is an accurate representation of the shape of the vocal tract (sounds coming out). Features extracted from each hamming window are the first 12 MFCC coefficients.

Zero-crossing rate (ZCR) feature is the rate of sign changes along with a signal, ZCR is computed from a hamming window as:

$$\mathrm{ZCR} = \frac{1}{199}\sum_{i=0}^{199} I\{x_i * x_{i-1} < 0\} \tag{1}$$

x_i is the sample i^{th} of the hamming window, I{A} is an indicator function.

$$\mathrm{I}\{\mathrm{A}\} = \begin{cases} 1 & \text{if A is true} \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

The entropy feature over a hamming window, the measure of the amount of information that is missing reception, can be calculated as:

$$\text{Entropy (x)} = -\sum_{i=0}^{199} p(x_i)\log(p(x_i)) \tag{3}$$

Where x_i is a sample value; $p(x_i)$, a probability distribution of x_i within the hamming window, can be estimated as the number of x_i in the hamming window divided by 200; and the probability 0*log(0) is assumed to be 0.

12 MFCC, ZCR, and Entropy features are combined into one feature vector sized of 14. Features are normalized to ensure all feature values in the range of [0, 1]. With feature size of 14, we can avoid unnecessary delay (good for real-time implementation) while we can achieve reasonable detection accuracy.

As an effective event extraction and processing can significantly reduce the computation resources, we develop a simple threshold-based event extraction algorithm for pruning the audio stream, and searching (highly potential) cough event candidates over the audio stream. The thresholds are estimated by using 4-fold cross validation procedure on the subset of the dataset. To select appropriate features for searching cough events, we test feature by feature over the dataset under 4-fold cross validation protocol, and then manually select 4 out of 14 features that perform best with high true positive rate and lowest false positive rate. The candidate cough events are afterward used for predicting cough.

3.5 Cough Prediction

Two models Gaussian Mixture model (GMM) and Universal Background model (UBM) are proposed for the prediction of cough. In this work, cough sound is modeled using

GMM while UBM is used for modeling background sounds which are any sounds out of cough, including noise, speech, etc.

In brief, the Gaussian Mixture model for modeling the cough G is a triple of 3 parameters: $\vec{\mu}_i$, C_i, $\vec{w}_i$ extracted from the training data:

In which $\vec{\mu}_i$ is the mean vector; C denotes the covariance matrix; and $\vec{w}_i$ represents the prior probabilities (w_i is the prior probability of i^{th} mixture component). Given the feature vector $\vec{f}$ computed from a hamming window, the likelihood for the cough is computed as follows.

$$\mathrm{p}(\vec{f}|\mathrm{G}) = \sum_{i=1}^{M} w_i N(\vec{f}\vec{\mu}_i, C_i) \tag{4}$$

Where N denotes the Normal probability distribution with the mean vector $\vec{\mu}_i$ and covariance matrix C:

$$N(\vec{f}|\vec{\mu}_i, C_i) = \frac{1}{|\sqrt{2\pi C}|} e^{-\frac{1}{2}(\vec{f}-\vec{\mu}_i)^T C^{-1}(\vec{f}-\vec{\mu}_i)} \tag{5}$$

A sliding window w, comprising of 80 hamming windows, the likelihood is approximated based on the independent assumption over hamming windows:

$$\mathrm{p}\,(\mathrm{w}|\mathrm{G}) = \prod_{i=1}^{80} \mathrm{p}\left(\vec{\mathrm{f}}_{\mathrm{i}}|\,\mathrm{G}\right) \tag{6}$$

$\vec{f}_i$ is the feature vector computed from the hamming window i^{th}

Mixture models for (known) cough are trained using feature vectors computed from audio data labeled with cough. The training process is straight forward by using k-Means clustering and Maximum Likelihood (ML) optimization. Model parameters $\vec{\mu}_i$, C_i, $\vec{w}_i$ are estimated on class-specific training data. The number of Gaussian mixtures M = 7 is estimated by a 4-fold cross validation procedure on the training dataset.

Training a Universal Background model (UBM) with background noise data is different from does the cough model (GMM). Similar to a GMM, a UBM is also a triple of $\mathrm{U} = \left\{\vec{\mu}_i, C_i, \vec{w}_i\right\}_{\mathrm{i}=1}^{\mathrm{M}}$ where M is the number Gaussian mixtures. As UBM is trained to capture general characteristics of all background noises, M is significantly larger than that is used in GMM for modeling cough. In our study, to estimate M, we vary M to 10, 20, 30, 50, 100 on a 4-fold cross validation procedure on a subset of the training data. M = 50 is selected as it performs best (highest true positive rate). This leads to the covariance matrices are large, to facilitate fast computation, only diagonal form of the covariance matrices are considered.

The UBM U is trained using Expectation-Maximization (EM) algorithm. The model parameters are initialized as.

$\mathrm{w_i} = \frac{1}{50}$; $\mathrm{C_i} = \mathrm{I}$ (unit matrix); and $\vec{\mu}_i$ is randomly selected from training data.

The EM algorithm performs iterations until parameters are stable. After training process, we have a pre-trained GMM model and a pre-trained UBM model. These models are combined into the prediction stage that we can deploy on the mobile devices for cough detection.

A cough in a sliding window is detected if $p(w|G) \geq p(w|U)$; that means:

$$\prod_{i=1}^{80} p(\vec{f}_i|G) \geq \prod_{i=1}^{80} p(\vec{f}_i|U) \tag{7}$$

Otherwise, a non-cough is detected.

3.6 Cough Monitoring

Once a cough is detected in real-time, its information including the time (precise in millisecond) and place (GPS data) will be written into a log file. In addition, we also remark the fits of coughing or whoops when there are more than 3 coughs are detected within 3 s. The Fig. 2 (right) is an example of a fit of coughing.

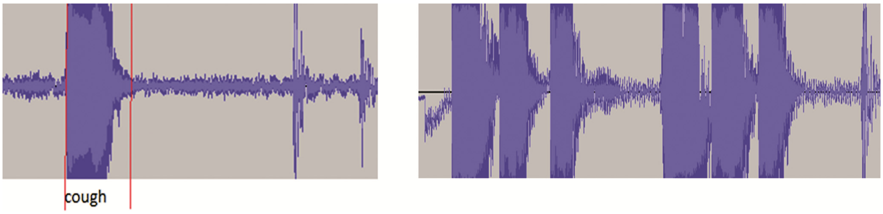

Fig. 2. A cough sound (*left*) and a fit of coughing (whoops) (*right*).

As cough information would need to be available for accessing by the doctor anytime, we have implemented some functions (as shown in the Fig. 3) that can easily be accessed and retrieved by both users and doctors.

Fig. 3. MobiCough GUI and Cough monitoring.

4 Experimental Evaluation

This section presents an empirical experiment for verifying our proposed method. The section is divided into two parts: data collection & annotation which describe the procedure of data collection from 10 pharyngitis or pneumonia patients; and followed by the evaluation.

4.1 Data Collection & Annotation

As so far no publicity of the cough dataset is available, so we ourselves need to collect cough data. We develop a simple audio logging program on Android phones for data collection. The logging program is deployed on 10 Android-based cell phones including Samsung, LG, FPT phones etc. 10 participants having with pharyngitis or pneumonia infection are willing to involve the study. The subjects were asked to use the cell phones installed audio logging program and to wear the mini wireless Bluetooth earphone speaker stereo on his/her collar for 3–6 h at different time of the day. The collected data is annotated by experimenters using Praat tool [15] with two labels: *cough* and *non-cough*. After annotation step, collected audio data is labeled with 1,117 cough events and more than 12,677 background events (any events out of the cough and silent in the audio logging files).

4.2 Evaluation

The cough event extraction can detect correctly 1,091 cough events out of 1,117 cough events of the ground truth. This results that our cough event extraction algorithm can achieve as high as 97.6 % of true positive rate. However, it misses out about approximately 3 % cough events while it incorrectly extracts 2,661 non-cough events. Details are shown in the Table 1.

Table 1. The test results for the event extraction algorithm

TP	FP	TN	FN
1,091	2,661	8,898	27

As the cough event extraction finds out 3,752 cough “candidate” events which are used for detecting cough from GMM-UBM models proposed in Sect. 3. An event can contain from a few to several hundreds of sliding windows. We do two evaluations: subject-dependent and subject independent protocols.

Under the subject dependent protocol, we use training and testing data from the same subject. For each subject, we divide the data into two equal portions; one is used for training, and the other is for testing; and the two portions are permuted; after that the results are averaged. We repeat the process for all subjects and the results are aggregated. It is noticed that the subject dependent implementation is particularly useful for cough detection systems that need to be adaptive to the users.

Under the subject independent, we use data from 9 subjects for training, and left out another subject to test. Then we repeat the process for all subjects and the results are aggregated. It is noticed that the testing data is not included into the training data and comes from other subject. The subject independent protocol is useful for cough detection systems acquiring pre-trained models (the cough model is trained from other person's coughs).

The evaluation results are represented in the Fig. 4. Subject-dependent result (91 %) is significantly higher than subject-independent (81 %). This is reasonable as test and train data are the same person for subject-dependent evaluation. This is highly recommended that models trained for cough detection and monitoring systems on personal devices coming from the owner would improve the detection accuracy. With the precisions and recalls are as high as 81 % and 91 % for subject-independent and subject-dependent respectively, MobiCough demonstrates that the detection of coughs using low-cost mobile devices is feasible. These results are very potential for heal-care applications that acquire cough information for diagnose and treatment cough-related diseases. The evaluation results are comparable to other works [3, 4, 6, 7, 9, 16] while we mainly focus on the detection of cough in real-time which is very crucial for opportune interventions, diagnosis, cure, treatment and even emergency aids made by the doctors and remarkably more convenient use.

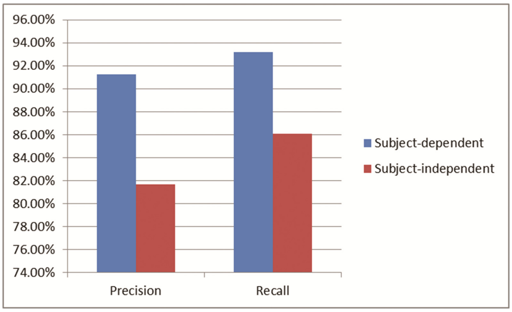

Fig. 4. Precision and Recall of cough detection evaluations

5 Conclusion

We present MobiCough, a method and system for cough detection and monitoring on low-cost mobile devices in real-time. Our proposed method combines predictive GMM and UBM models to enhance the detection performance. The proposed method is rigorously evaluated over a dataset consisting of more than 1000 cough events and a significant number of noises. With the detection accuracies are more than 91 % precision and recall for subject dependent training, and over 81 % precision and recall for subject independent training, the results are very comparable to other works while being feasible for real-time processing, low-cost and pretty convenient use. MobiCough is very promising for heal-care applications that acquire cough information for diagnose and treatment cough-related diseases.

References

1. WHO's pneumonia fact sheet. http://www.who.int/mediacentre/factsheets/fs331/en/
2. KidsHealth. http://kidshealth.org/parent/infections/lung/pneumonia.html
3. Larson, E.C., et. al.: Accurate and privacy preserving cough sensing using a low cost microphone. In: Proceedings of UbiComp, pp. 375–384. Beijing (2011)
4. Birring, S.S., et al.: The leicester cough monitor: preliminary validation of an automated cough detection system in chronic cough. Eur. Respir. J. **31**(5), 1013–1018 (2008)
5. Schappert, S., Burt, C.: Ambulatory care visits to physician offices, hospital outpatient and emergence. Vital Health Stat. **13**, 1–66 (2006)
6. Drugman, T., et al.: Audio and contact microphone for cough detection. In: Proceedings of INTERSPEECH, pp. 1303–1306. IEEE Press, Portland (2012)
7. Vizel, E., et al.: Validation of an ambulatory cough detection and counting application using voluntary cough under different conditions. Cough **6**(3), 1–8 (2008)
8. Kraman, S.S., et al.: Comparisons of lung sound transducers using a bioacoustics transducer testing system. J. Appl. Physiol. **101**(2), 169–176 (2006)
9. Masto, S., et al.: Detection of cough signals in continuous audio recordings using hidden Markov models. IEEE Biomed. Eng. **53**(6), 1078–1083 (2006)
10. Zheng, S., et al.: CoughLoc: location-aware indoor acoustic sensing for non-intrusive cough detection. In: International Workshop on MobiSys (2011)
11. Pham, C., et al.: The ambient kitchen: a pervasive sensing environment for situated services. In: Proceedings of ACM Conference on Designing Interactive Systems, Newcastle, UK (2012)
12. Pham, C., et al.: A wearable sensor based approach to real-time fall detection and fine-grained activity recognition. J. Mobile Multimedia **9**, 15–26 (2013)
13. Drugman, T., et al.: Assessment of audio features for automatic cough detection. In: Proceedings of 19th European Signal Processing Conference, pp. 1289–1293 (2011)
14. Mark, S., Hyekyun, H., Mark, B.: Automated cough assessment on a mobile platform. J. Med. Eng. **2014**, 1–9 (2014)
15. Praat. http://www.fon.hum.uva.nl/praat/
16. Shin, S.H., et al.: Automatic detection system for cough sounds as a symptom of abnormal health condition. Trans. Inf. Tech. Bio. **13**(4), 486–493 (2008)