A Mobile-Cloud Service for Physiological Anomaly Detection on Smartphones

Dezhi Hong
Computer Science
University of Virginia
dh5gm@virginia.edu

ABSTRACT

There is a growing number of examples that use the microphones in phone for various acoustic processing tasks as mobile phones become increasingly computationally powerful. However, there is no general-purpose physiological acoustic anomaly detection service available on smartphones. To this end, we propose a physiological acoustic anomaly detection service which implements a complete dedicated pipeline to detect irregularity and anomalies in lung sounds and notifies the user. The service is assisted with the cloud in the sense that users can contribute to the repository of sound clips in the cloud and offload the training of classifiers to the cloud. We also build a portable accessory that can be easily used to collect lung sounds and perform a preliminary evaluation of our system. Our service is able to identify the anomalous periods in asthmatic lungs with 98.2% accuracy.

Categories and Subject Descriptors

C.3 [**Special-Purpose And Application-Based Systems**]: Real-time and embedded systems

General Terms

Performance, Experimentation

Keywords

Physiological Sound, Anomaly Detection, Mobile Service

1. INTRODUCTION & RELATED WORK

Mobile phones are prevalent enough to reach nearly everyone in the world and a large portion of them are smartphones. Meanwhile, as smartphones become increasingly computationally capable and powerful, m-Health has been in rapid development in the past years. Mobile health platforms is able to track the condition of a user on a 24×365 basis that is something a real doctor would not be able to accomplish. Such a platform could also free the users from frequenting doctors as well as enable remote diagnosis, especially in the third world countries where medical condition is far from being satisfactory. Daily acoustic footprints of a user's heart and lung can be referred to in later medical consultations.

Smartphones have noticeably begun to use their microphones for various acoustic processing tasks, some state-of-the-art examples include: emotion and stress detection [6, 9], cough detector [5], lung function monitoring [4], heart beats counter [8]. However, these are generally limited to voice and environmental sounds detection, and specifically, MusicalHeart [8] only delves into music's impacts on heart beats. To the best of our knowledge, currently there is no general-purpose physiological acoustic anomaly detection service on smartphones. To this end, we propose a physiological acoustic anomaly detection service [3] based on our mobile-cloud platform – Auditeur [7] (a French word meaning "Listener"). Auditeur is a general-purpose acoustic event detection platform that provides APIs to developers to enable their apps to register for and notify the users on a wide variety of acoustic events. The platform is efficient in communication, computation, and energy consumption since the phone will connect to the cloud only once to download the classification plan. The acoustic computational units (e.g., FFT) in the phone will be dynamically instantiated and wired up – so that only necessary components run within the phone. An energy-aware acoustic processing plan to be executed on the phone will be generated at the cloud side to increase the lifetime of the device while maintaining high accuracy.

However, there is no specific classification plan for physiological sounds in the current version of Auditeur, therefore we leverage the platform and extend it to include classifiers used for detecting irregularity and anomalies in lung sounds and alerts the user. Such a service is believed to provide promising opportunities for early pulmonary problems detection. The main contributions of this work are:

- First, we implement a service that contains a complete pipeline to accomplish the task of anomaly detection in asthmatic lung sounds.
- Second, we build a cheap, portable and easy-to-use accessory with commodity components to collect lung sounds and we verify the effectiveness and reliability of the tool.
- Third, we evaluate our service with 3 diseased lung sounds and show that our service is able to identify the adventitious parts with 98.2% accuracy.

2. LUNG SOUNDS BACKGROUND

In this section, we give a brief introduction to a few features of typical normal and abnormal lung sounds.

2.1 Normal Lung Sounds

Normal lung sounds can be heard from different positions on body and the sounds vary a lot in terms of intensity, pitch and inspiration-expiration temporal ratio.

- **Tracheal**: Heard over the trachea as harsh, high-pitched and discrete sounds. This originates from turbulent air flow in the upper airways of the body and cover a wide range of frequency from 100 Hz to 1500Hz. In general, the expiration phase is slightly longer than inspiration.
- **Bronchial**: Mostly heard over the manu-brium or upper part of the sternum and usually consist of high-pitched and large amplitude sounds similar to the sound of air blowing through a tube. There is a brief pause between expiration and inspiration, where expiration is longer than inspiration.
- **Bronchovesicular**: Heard near the main stem bronchi, i.e., the median upper chest. They contain pitch and intensity characteristics between vesicular and bronchial sounds. They can be heard during the inspiratory and expiratory pass, which each lasts for approximately the same period and do not have a pause in between them.
- **Vesicular**: Heard over majority of the lung during the shallow breathing of normal respiration. They possess low-pitched and soft sounds which are generated as a result of changing airflow patterns in the lungs. There is no pause in between the two phases and the inspiration is longer than expiration.

2.2 Adventitious Lung Sounds

Adventitious sounds usually signify a pulmonary disorder or disease in patients. Here are several adventitious lung sounds most commonly encountered in clinical auscultation:

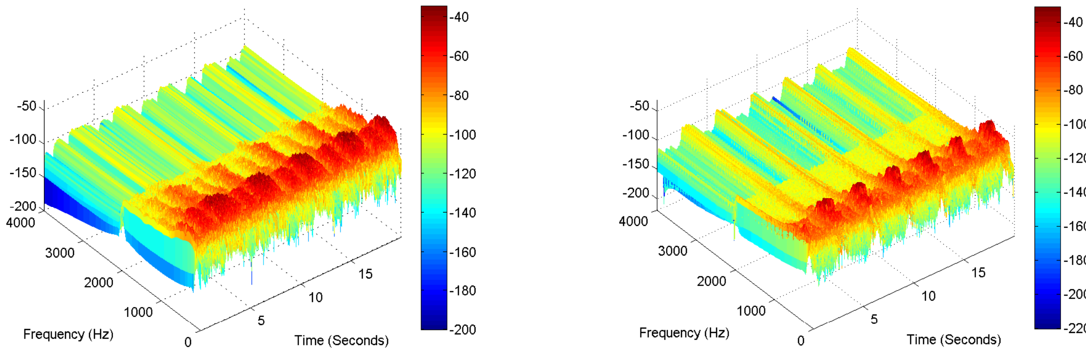
- **Crackles**: Are discrete and explosive popping sounds and generally heard during inspiration. They are caused by air passing through moist airways and alveoli. Short duration (5-10 ms duration) crackles are called fine crackles and long duration (20-30 ms) ones are called coarse crackles. They are associated with lung disease such as pneumonia as well as with heart disease.
- **Wheezes**: They are high-pitched sounds that are more prominent in expiration with a duration longer than 250 ms. They are caused by air flowing across obstructed passages and create a few sharp spectral peaks around 400 Hz. These sounds indicate diseases such as pneumonia, asthma and emphysema.
- **Rhonchi**: They are continuous and low-pitched sounds with causes similar to those for wheezes. They occur at frequencies lower than 300 Hz and usually come with bronchitis and chronic obstructive pulmonary disease (COPD).

3. APPROACH

In this section, we sketch our analysis both in the time domain and in the frequency domain. In the time domain, we concentrated on respiratory cycle per minute (cpm) and amplitude ratio of inspiration to expiration (I/E). For the frequency domain, we first discuss an observation of the data and then based on that we present the complete pipeline of feature extraction. In general, each sound clip is segmented into short frames and each frame is expressed with the peak frequency in spectrum and the bandwidth of the peak.

3.1 Time Domain

To identify the respiratory cycle which is a basic feature of lung sounds, several stages of processing includes: 1) down sample and apply a band-pass filter to reserve the envelope of the signal, 2) de-noise and further down sample, 3) apply a low-pass filter, 4) take the derivative of the signal from 3). The band-pass filter in step one has a cut off frequency of 100-2000 Hz which reserves the interval we need. The twice downsamplings reduce the number of data points and relieve the computational task for smartphone. The low-pass filter in



(a) a normal lung **(b)** an asthmatic lung

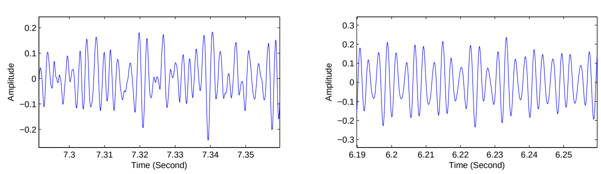
Figure 1: Spectrogram of lung sounds: an asthmatic lung has sharper hotspots than a normal lung.

step 3 has a cut off frequency of 0.6 Hz because usually the cpm index of a human being cannot exceed 30, which is 0.5 Hz. The zero-crossing points in the derivative help locate the peaks in the original signal and the interval between two adjacent peaks is used to calculate the respiratory cpm.

3.2 Frequency Domain

3.2.1 Observation

Wheezing lung sound are believed to possess a different pitch from a normal lung sound, and if we zoom into a random short period of 50 ms frame of the expiration phase in both a normal lung and one with asthma accompanied by wheezing, clear difference is observed as illustrated in Figure 2. The one from an asthmatic lung on the right contains a component with a certain frequency while no obvious component or harmonics in the waveform of a normal lung is seen.



(a) a normal lung **(b)** an asthmatic lung

Figure 2: A 50 ms long window in respiratory recording: a certain frequency component is observed in asthmatic lung while a normal lung is not.

Based on our observation, we make the conjecture that there should be a few hotspots in the spectrogram as well as spectrum of asthmatic lung sounds, which would correspond to the high pitch of wheezing. The conjecture is partially verified by the spectrogram of two signals as shown in Figure 1. The one on the right from a diseased lung has more sharp and dominant hotspots in the spectrogram than a normal lung.

3.2.2 Feature Extraction

In this work, our efforts are mainly focused on wheezing detection in lung sounds. We apply a window-based FIR bandpass filter to the audio clips and preserve only part of the sounds that fall into the frequency band of 100-2000 Hz. The size of the window is 1/12 of the sample frequency and recall the description in previous section, wheezing is a period as minimal as 250 ms in duration, so we at first break each clip into 300 ms-long sliding frame with 50% overlap between each two for feature extraction. We then look into the power spectral density (PSD) of the signal. Specifically, the PSD is computed with an auto-regressive (AR) model of the original time-series signal, assuming x_t is the value of signal at time t and the AR is described as:

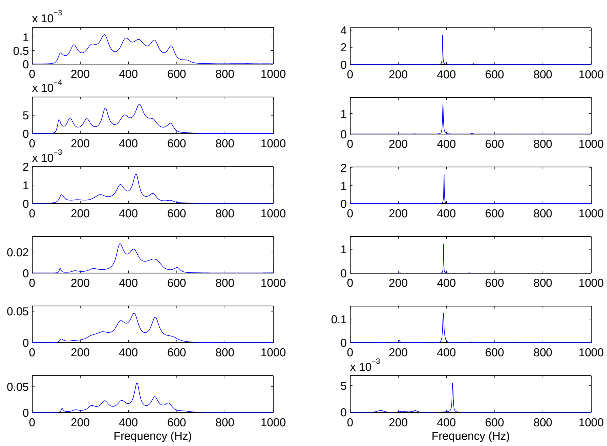
$$x_t = \sum_{i=1}^{p} \varphi_i x_{t-i} + \varepsilon_t$$

Where p is the order of the AR model, φ_i is the ith AR coefficient, ε_t is the white-noise signal at time t, σ^2 is the variance of the white-noise signal. By applying the AR model on the time-series signal, we take into account the correlation between points and the signals in frames containing transition between wheezing and non-wheezing would be smoothed. The PSD of the AR signal can be obtained with:

$$P_i(w) = \frac{\sigma^2}{|1 + \sum_{i=1}^{p} \varphi_i e^{-jwi}|^2}$$

Where the amplitude of PSD reflects the power at frequency w. PSD estimates the power of each single frequency content in the frequency domain which is just the square of the amplitude of each frequency in spectrum. And there is an FFT unit in Auditeur [7], to simplify the logic of the feature extraction, we just do FFT on each 300 ms AR time window and the outputs of FFT are then squared to get the PSD of every single frequency content. Figure 3 gives a sense of the PSD of signals from normal and diseased lungs, where each subplot is the result of a 300 ms frame and the model order p is set to 100 which gives us a clear-cut spike in the spectrum.

The spikes located around 400 Hz in the right figure match the special high pitch in the wheezing sounds, and these spikes are what we need to use as a basis of anomaly detection. We search for peak f_p with the highest amplitude in PSD spectrum and compute the bandwidth B_p of the peak, where B_p is defined as the difference of the higher and lower boundary of the peak, i.e., $(f_h - f_l)$, and f_h, f_l of a peak are required to fall below 5% of the peak amplitude. With the peak frequency and bandwidth, a frame is represented as a vector (f_p, B_p).

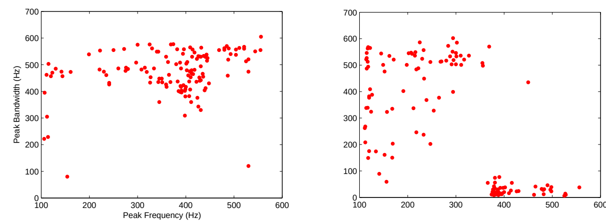


(a) a normal lung **(b)** an asthmatic lung

Figure 3: Power spectral density of an expiration phase: the power of a normal lung is broadly distributed while an asthmatic lung falls upon a certain frequency.

3.2.3 Clustering

Figure 4 shows an example of the distribution of frames as feature vectors from two sound clips in the 2D space. We see that on the left, which is the one for a normal lung, few points fall into the lower half of the space meaning the peaks in the PSD spectrum have wide bandwidth generally while on the right there is a bimodal distribution, and the small dense cluster in the right-lower area correspond to the wheezing frames. Therefore we exploit the k-means clustering algorithm (k is set to 2) to accomplish the classification task because of the distinct characteristic in feature distribution of normal lung sounds and wheezing ones.



(a) a normal lung **(b)** an asthmatic lung

Figure 4: Distribution of frame feature vectors

4. PRELIMINARY RESULTS

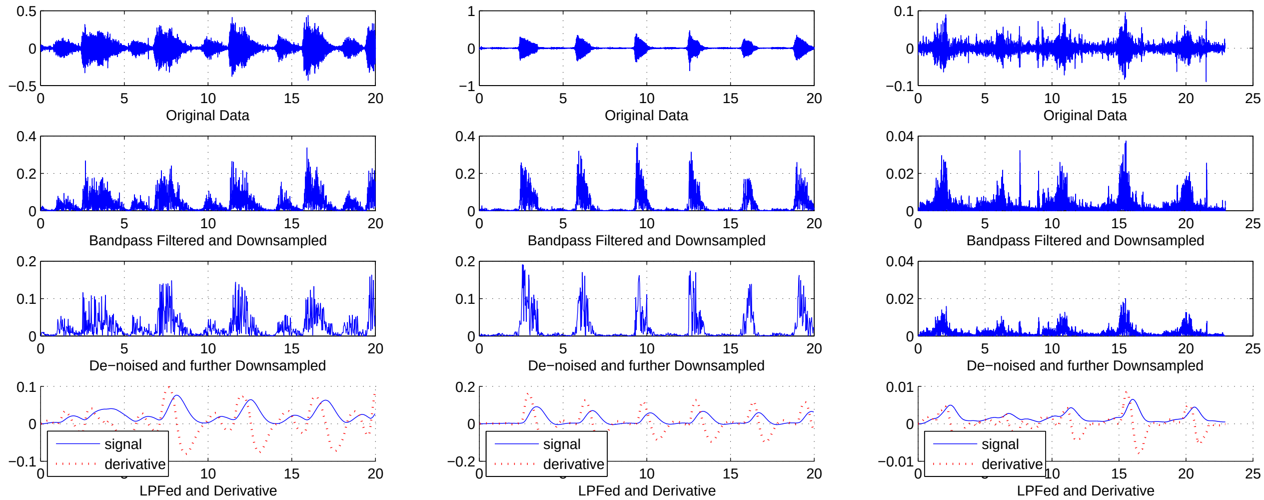
4.1 Experimental Setup

We perform a preliminary evaluation of our approach on 6 sound recordings obtained from [1, 2], which consists of 2 normal lung sounds and 4 wheezing sounds collected by digital stethoscopes and 3 recordings collected with the device we built. The duration of the clips ranges from 10 to 20 seconds and the sample frequency varies from 16 kHz to 44.1 kHz. The ground truth of a frame whether it is a wheezing one or not is manually labeled by human, so it's non-trivial to label some frames containing transitions from non-wheezing phase to wheezing phase and this is the main cause of classification error as will be discussed later in this section.

The device we built is comprised of a commodity electret microphone and a conventional mechanical stethoscope. We don't need the earbud at the end of the tube of stethoscope so we remove it. The tube is also cut off to minimize the attenuation in sound along the tube and the microphone is placed at the opening end of the tube to collect data. To use the device for data collection, one can just plug in the microphone into the audio jack of a smartphone and place the stethoscope on the body.

4.2 CPM and I/E

We analyze some recordings and the result goes in Figure 5. From the result we see that: (a) and (b) show the effectiveness of the processing in the time domain to extract respiratory cycle; the cpm increases from normal lung (a) to diseased (b); the I/E ratio also varies from (a) to (b); (c) is a sound clip collected with our device and demonstrates the usefulness of our tool. Parameters such as cpm and amplitude ratio of I/E can act as indicators of how well a user's lung functions on a daily



(a) A normal lung

(b) A lung with asthma

(c) A normal lung

Figure 5: Samples of time domain analysis on lung sounds: (a) and (b) are from online database, (c) is collected with our device.

basis and these footprints can be referred to in our service. However, we still lack the expertise on how to use these features in the time domain and we are exploring, so we skip further evaluation at this time.

4.3 Clustering

We evaluate the clustering accuracy on 4 wheezing lung recordings and each recording is split into two subsets to do the classification with 2-fold cross-validation: one half is used to train the clusters and the other is used to test the clustering accuracy. The clustering result is shown in Table 1. On average, we achieve 98.2% and 90.1% accuracy of wheezing period and normal period detection respectively. We also manually check the misclassified frames and they all locate at the transition joints from non-wheezing to wheezing periods or the opposite, which is hard to determine whether it's a wheezing frame or not by human.

	non-wheezing	wheezing
non-wheezing	254	28
wheezing	2	108
accuracy	90.1%	98.2%

Table 1: Confusion matrix of the clustering result

5. CONCLUSION AND DISCUSSION

In this work, we propose and implement a service with a complete classification plan as an extension to Auditeur on smartphones. The service can be used to detect potential irregularity in lung functions on a daily basis and alert the user to the presence of anomalies. On average, the system is able to identify the wheezing parts with 98.2% accuracy by exploiting k-means clustering. The reason why we refer to k-means rather than choosing support vector machines (known as SVM, which is good at classifying high dimension data vectors with a hyperplane) is because the frames sitting at the transition between wheezing and non-wheezing make the boundary of data distribution less clear-cut thus degrading the performance of SVM. We also build our own device for data collection and prove the usefulness as well as effectiveness of the tool.

In future work, we plan to test the tool on real patients with pulmonary diseases to see if the tool is able to pick up the specific frequency component in sound and collect our own abnormal lung sounds for further evaluation. We will explore more on how the cpm or I/E relates to activity levels or disease severity, and how the peak frequency in PSD spectrum of wheezing varies from day to day of a specific person. We would like to also extend the category of adventitious sounds detected by our service to crackles as well as other positions on body such as stomach and intestine.

6. ACKNOWLEDGEMENT

This work is part of a larger project which is focused on producing a holistic platform for daily human-wellness monitoring and is supported, in part, by NSF grants CNS-1035771 and EECS-901686.

7. REFERENCES

[1] StethoGraphics. `http://www.stethographics.com/`.

[2] The R.A.L.E Repository. `http://www.rale.ca/Recordings.htm/`.

[3] D. Hong, S. Nirjon, J. A. Stankovic, D. J. Stone, and G. Shen. Poster abstract: a mobile-cloud service for physiological anomaly detection on smartphones. In *Proceedings of the 12th international conference on Information processing in sensor networks*, IPSN '13, 2013.

[4] E. C. Larson, M. Goel, G. Boriello, S. Heltshe, M. Rosenfeld, and S. N. Patel. Spirosmart: using a microphone to measure lung function on a mobile phone. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, UbiComp '12, 2012.

[5] E. C. Larson, T. Lee, S. Liu, M. Rosenfeld, and S. N. Patel. Accurate and privacy preserving cough sensing using a low-cost microphone. UbiComp '11, 2011.

[6] H. Lu, D. Frauendorfer, M. Rabbi, M. S. Mast, G. T. Chittaranjan, A. T. Campbell, D. Gatica-Perez, and T. Choudhury. StressSense: detecting stress in unconstrained acoustic environments using smartphones. UbiComp '12, 2012.

[7] S. Nirjon, R. F. Dickerson, P. Asare, Q. Li, D. Hong, J. A. Stankovic, P. Hu, X. Jiang, and G. Shen. Auditeur: A Mobile-Cloud Service Platform for Acoustic Event Detection on Smartphones. In *Proceedings of the 10th international conference on Mobile systems, applications, and services*, MobiSys '13, 2013.

[8] S. Nirjon, R. F. Dickerson, Q. Li, P. Asare, J. A. Stankovic, D. Hong, B. Zhang, X. Jiang, G. Shen, and F. Zhao. MusicalHeart: a hearty way of listening to music. SenSys '12, 2012.

[9] K. K. Rachuri, M. Musolesi, C. Mascolo, P. J. Rentfrow, C. Longworth, and A. Aucinas. EmotionSense: a mobile phones based adaptive platform for experimental social psychology research. Ubicomp '10, 2010.