Technologies for Developing Ambulatory Cough Monitoring Devices
Justice Amoh and Kofi Odame
Analog Laboratory, Thayer School of Engineering, Dartmouth College.
address: 14 Engineering Drive, Hanover, NH 03755
tel: +1 (603) 513 - 8411, +1 (603) 646 - 9156
{justice.amoh.jr, kofi.m.odame}@dartmouth.edu

ABSTRACT
Cough is a prevailing symptom in most lung diseases. While cough sounds themselves can be very instrumental in the diagnosis of certain diseases, their intensity and frequency also infers the intensity of the particular illness. There is an imperative need for a robust system for identifying and analyzing cough sounds. In implementing such systems, researchers are confronted with technical challenges such as the choice of sensors and methods of signal acquisition, the real time analysis of the acquired signals, and the accurate identification of cough events, distinguishing them from similar sounds such as speech, laughing, throat clearing and sneezing. Previous approaches have employed external environmental sensing methods to achieve more accurate detections at the expense of mobility, scalability and real-time cough sensing. Alternative approaches have proposed wearable cough sensing methods, which, while mobile, can often face challenges in terms of robustness and obtrusiveness. In this paper, we explore the strengths and shortcomings of the various techniques that have been proposed for automatic detection and analysis of cough sounds. We also suggest the next steps in furthering the state of the art.

I. INTRODUCTION
Patients with asthma and chronic obstructive pulmonary disease (COPD) are at a high risk of suffering an exacerbation of their conditions. These episodes account for more than 700 million hospitalizations and emergency department visits per year in the United States, which places a significant burden on the patients and the health care system as a whole. From the patients’ perspective, admission to hospital carries a serious risk of harm due to medical error or infection. For the health care system, the financial cost of asthma and COPD hospitalizations is approximately $23 billion per year. In addition, emergency care diverts resources away from elective healthcare and causes disruptions in patient waiting lists.

Self-management plans have been shown to dramatically reduce the incidence of hospitalization. For example, it has been shown that self-management by asthma patients can lower the incidence of hospitalization by as much as 40 %. Self-management plans involve monitoring one’s
symptoms (cough, wheeze, dyspnea), the frequency of use of medication and physiological measures like FEV and PEF, as well as implementing a pre-determined action plan to address exacerbations.\textsuperscript{17}

Unfortunately, patients’ adherence to a self-management regime is poor, with symptoms often going unreported and physiological measurements sometimes being fabricated.\textsuperscript{18} For this reason, there is growing interest in developing new tools that will allow patients to accurately monitor their management of the disease in a way that is convenient, minimally disruptive and encourages compliance.\textsuperscript{19-21} One of these tools is an ambulatory cough monitoring device, which maintains an automatic and continuous record of the frequency of the patient’s cough symptoms.

In addition to its importance in asthma and COPD, cough is also a common symptom of other airway diseases such as pneumonia and tuberculosis.\textsuperscript{22} In some instances, it is in fact the only observable symptom of the disease.\textsuperscript{23} The frequency and intensity of coughs in a patient correlates with the severity of their illness.\textsuperscript{24} Also, the characteristics of coughs seem to vary across diseases and hence can be useful in the diagnosis of certain medical conditions.\textsuperscript{25}

Given the clinical relevance of coughs, there is a need for robust systems for automatically identifying and characterizing them. In this paper, we will review the technology trends that have been driving the development of ambulatory cough monitoring devices.

\textbf{II. SENSORS}
\textbf{A. Types Of Sensors}

The cough reflex results in a characteristic sequence of measurable changes to the flow rate of air into the lungs, lung volume, intrathoracic pressure and vocal sounds.\textsuperscript{22,26} During the first, inspiratory phase of the cough, the patient draws air into the lungs, causing a negative flow rate and making the lungs increase in volume. Following this is the compressive phase, during which the glottis is closed off and the flow rate is essentially zero. Intrathoracic pressure builds up to a maximum during this phase, as the chest wall contracts against the closed glottis. The final phase is the expulsive one, when the glottis is forced open with the expiration of air. This rush of expelled air causes the flow rate to reach a positive maximum level. Also, the intrathoracic pressure drops from its compressive phase maximum, and the lung volume drops as well. It is during the expulsive phase that the familiar cough sound is created.
The rich variety of physiological changes that characterize a cough means that it can be captured by any one of several different types of sensors, from microphones to flow meters and balloon catheters. Each of these sensors presents some challenges when applied to an ambulatory, continuous monitoring setting.

To measure intra-thoracic and intra-abdominal pressures changes, balloon-tipped catheters can be inserted via one nostril and placed in the esophagus and stomach, respectively. The balloon is inflated with several milliliters of air, and it transmits any pressure changes down the length of the catheter to an external pressure transducer, which converts the signal into a voltage. The catheter system is able to capture the maximal pressure and pressure drop off that are characteristic of the compressive and expulsive phases. However, the insertion of the catheter requires medical expertise, and its invasive nature can cause complications such as infection, tissue trauma, and accidental insertion into the trachea. Also, the potential for patient discomfort (gagging, vomiting) makes this an impractical solution for long-term automated cough monitoring.

The chest volume changes that accompany a cough can be measured with a plethysmography system. One approach is optoelectronic plethysmography, where the motion of the subject’s abdomen is captured by a video recorder. Special reflective markers that are placed on specific sites on the abdomen allow image processing software to calculate the volume of the abdomen. Statistical analysis of the volume data can be used to detect the occurrence of a cough. Unfortunately, the video capture requirements of optoelectronic plethysmography make it suitable only for a laboratory setting. A feasible alternative for ambulatory applications is a chest impedance belt. The impedance measured across the chest is related to the amount of air inside – and hence volume of – the chest. So far, the chest impedance solutions that have been reported have been bulky and obtrusive, and are not very accurate predictors of cough. However, advances in electronic textile technology might provide some new solutions to these issues.

A flow sensor can be used to measure the respiratory air flow rate. A commonly used flow sensor is based on an instrumented nasal cannula, which is kept in place with a harness that rests on the ears. Integrated into the cannula is a thermistor or a thermocouple, which measures the nasal air flow rate by sensing the temperature changes that are caused by inhalation and exhalation. Unfortunately, this device does not capture oral air flow (that would require a face mask, which is impractical for continuous, mobile monitoring). Since most of the air expulsion of a cough is done through the mouth, the nasal flow sensor has a very limited ability to detect cough events.
The cough sound can be measured with one or more audio sensors that are placed on, or within the vicinity, of the subject. The audio sensors capture the cough sound’s sequence of distinct phases, each of which has a characteristic spectral signature. This provides multiple dimensions for analysis and identification of the cough. As such, the most reliable cough monitoring devices are based on audio sensors.

B. Limitations of Audio Sensors
There are three main approaches to implementing an audio-based cough monitoring system, and they each present a different trade-off between performance and restrictions on the patient's physical location.

Some monitoring systems use a large network of audio sensors that are installed in the patient's environment, such as in the home or a hospital ward. The multiple distributed sensor nodes perform collaborative acoustic sensing, allowing the system to capture and localize -- via spatial filtering -- any cough sounds that occur within the instrumented environment. On the other hand, the system is useful only if the patient remains within the instrumented environment.

A body-worn sensor allows the patient to move about freely, but its necessarily-small size creates other challenges in terms of system performance. For instance, the sensor can only be powered with a small battery, which limits its power budget. This in turn constrains the amount of computation that the sensor can perform. To address this problem, some sensors compress the audio signals and wirelessly transmit them to another device, (e.g. a server) for further analysis and processing. Environmental noise poses another challenge for body-worn sensors. Typically, spatial filtering techniques would be able to suppress extraneous sounds, but they can only be used to limited effect with a body-worn sensor, given that it can only accommodate a small array of microphones.

Piezotransducers are a way to avoid the environmental noise problem that is faced by small arrays of microphones. When the patient coughs, some of the sound energy is transmitted through the chest and to the surface of the body. These vibrations can be detected by a piezotransducer that is placed in contact with the patient's thorax or throat. Figure 1 shows a piezotransducer sensor and a casing that we have fabricated for cough analysis in our lab. Because the piezotransducer is not a free-air microphone, it will not measure any environmental sounds. However, the disadvantage is that poor contact between the piezotransducer and the patient’s skin will result in signal attenuation. Also, the patient's natural movements can create
motion artifacts that will corrupt the measured signal.\textsuperscript{38,39} There is a lot of active research devoted to addressing contact quality and motion artifact, as these are issues that are faced by most types of on-the-skin sensors for ambulatory applications.

The work of Larson \textit{et al} also highlights a problem that is common to all continuous acoustic sensing systems: privacy and security.\textsuperscript{44} Since the microphone can pick up voice and other sound signals, there is the need for ways to only register the cough signals so as not to invade the speech privacy of the patient. Also, transmitting such data wirelessly to a server or substation calls for some security measures to secure patient's health information. They address these challenges by first decomposing the acoustic signals to matrices of eigenvectors, and then only transmitting these eigenvectors to the server. They later on reconstruct the cough data on the other server side for further processing. The decomposition and reconstruction are done in a way such that only cough data can be reconstructed with reasonable fidelity. While their approach is promising, it still does not guarantee other forms of audio information privacy such as the identity and location of the user. Hence, privacy and security remain real challenges in the design of cough monitoring systems.

\section*{III. COUGH SOUND DETECTION}
\subsection*{A. Modeling the Cough Sound}
Speech, laughter and throat-clearing can sometimes have similar characteristics as cough sounds. To avoid false positive detections, it is important to form an accurate model of the cough sound.

One approach to modeling the cough sound is to characterize the signal by its tussiphonogram. This is the total energy of the cough given by the integral of the sound intensity.\textsuperscript{45} Korpas \textit{et al}
and Salat et al have discussed how the tussiphonogram of cough sounds can be distinctive of diseases and hence can be diagnostically useful. Korpas et al and Abaza et al also add that the tussiphonogram can inform about pathological changes in the airways since the nature of the airway walls greatly influence the intensity of cough sounds produced. Therefore, not only can the tussiphonogram be useful in diagnosing some respiratory diseases, but it can also be instrumental in determining the severity of the disease or the effectiveness of a particular treatment.

A number of studies have also modeled the cough sound signal as consisting of three phases. Thorpe et al define these as the "initial burst", "middle phase" and "final burst". The first phase corresponds to the opening of the glottis and the burst of sound energy that is observed as the first cough sound. Then follows the longer second phase which is a relatively steady flow of air. The third phase, which is not always present, corresponds to the second cough sound from a secondary glottal closure. These cough divisions are illustrated in four cough events in Figure 2. Olia et al inform more that there are variations in the energy and frequency components across the three phases. They mention that though the first and final phases have highly variable frequency content, the middle phase has a more stable frequency content. Thorpe et al also add...
that the features of the first and middle phases especially, differ between asthmatics and non-
asthmatics. And so, the subdivision of the cough signal into three phases can also be helpful in
screening diseases.

B. Cough Sound Features

Once we have decided on a model for characterizing the cough event, we ought to identify the
audio features of interest in analyzing the cough event. These features are the parameters of the
acoustic signals that can be used to verify the occurrence of a cough event and to describe the
nature of the cough. We can broadly categorize the features into two: the temporal and the
spectral features.

The temporal features refer to those parameters that are obtained from the time domain
waveform of the cough sound. The waveform of two cough events is shown in comparison with
a speech waveform in Figure 3 (a). A simple temporal feature that can be used is the duration of
the cough event. Coughs are generally noted to have an average length of about 350 ms. For
applications that employ flow-meters alongside microphones, a related measure of cough
duration used is a ratio of the length of the flow rate signal and the acoustic cough signals. Another temporal feature commonly used is some averaging of the intensity of the cough sound.
This could refer to the mean, or the ratio of the maximum sound wave pressure from the cough to the minimum. Some studies also include skewness and kurtosis, both of which are a description of the probability distribution of the signal, as temporal features of the cough sound.

The spectral features, on the other hand, refer to those notable parameters of the frequency domain of the cough signal. The spectrogram of cough and speech signals in Figure 3 (b) depicts that the spectral content of cough and speech signals differ essentially. The dominant frequency of the cough sound spectrum is one useful spectral feature. It refers to the frequency with the highest energy or amplitude in the cough sound. Through generation of spectrograms, a number of studies have observed the frequency distribution of cough sound signals and have sometimes identified fundamental frequencies in the phases of cough events. Korpas et al add that in pathological cough sounds, the fundamental frequencies deviate from the expected. Chung et al generalize this notion by expressing that the distribution of power amongst the frequencies in the spectrum can reveal more about the nature of the cough event. They add that such analysis can characterize audible changes in cough sounds from different conditions. A related spectral feature is the energy distribution in the spectrogram. Chatzarrin et al demonstrate that the energy observed in cough phases is descriptive of the airway pathology. Also, the energy of the cough sound has been mentioned to differ between male and female subjects. Generally, energy measurements are mostly instrumental in verifying the presence of audio activity. However, the most common spectral feature used is perhaps the Mel Frequency Cepstral Coefficients (MFCCs). The Mel-frequency cepstrum represents the short-term power spectrum of a signal. The mel frequency scale itself is a logarithmic frequency scale that closely corresponds to human

![Fig. 4. A 3D scatter plot of three Mel Frequency Cepstral Coefficients (MFCCs) for cough (red) vs speech (cyan) events. The MFCCs are widely used in speech recognition and audio processing applications. They are good for spectral characterization of cough and are a very discriminative feature set for cough identification. Normally, 12 coefficients are used but we only show three of these here.](image_url)
auditory perception. Hence, the MFCCs are widely used in speech and other audio processing applications. In cough analysis, many studies have found the MFCCs to be a very discriminative feature. In Figure 4, three MFCCs associated with several instances of cough (red) and speech (cyan) data are visualized in a 3D scatter plot to illustrate the discriminative nature of MFCCs in cough detection. In addition, Itticaichareon et al mention that it's relatively easy to implement software algorithms for extracting MFCCs, making it an even more desirable feature for cough processing.

Besides the temporal and spectral features, we can also consider another set of audio features. This third group refer to those features that describe how noisy the sound signal is. Drugman et al and Thorpe et al both use the Zero-Crossings rate as one such feature. The Zero-Crossings Rate tells how many times the signal crosses the zero line. The noisier the signal, the more times it crosses the zero line and hence the Zero-Crossings Rate is a measure of noise in the signal. In the frequency domain, a feature that is a measure of noise is the Harmonic to Noise Ratio (HNR). The HNR is a description of the energy distribution between the harmonics and the noise in the signal. In fact, Drugman et al demonstrate that the HNR is a good feature for measuring the amount of noise in the signal and can be useful in identifying and characterizing cough sounds.

C. Classification

After the processing and analysis of the acoustic event, the next and most crucial step, is to decide whether or not the registered event is a cough. This decision making, based on the extracted acoustic features, is a classification problem and requires a machine learning classification algorithm. To understand the performance of any classifiers used, we have to first consider the challenges in identifying cough events in software. Matos et al mention the two main problems: 1) cough sounds vary between cough subjects and 2) cough sounds can be very similar to other ambient sounds like speech, sneezing and throat clearing. So the classifier should have a high sensitivity to be able to detect cough signals that may vary significantly from those used in training. Secondly, the classifier should have a high specificity, in order to better identify sound events that are not coughs. Researchers have used a number of different classifiers for cough detection. The most popular and efficient ones used are Decision Trees, Artificial Neural Networks (ANN), Hidden Markov Models(HMM) and Gaussian Mixture Models(GMM).

The Decision Tree algorithm uses a tree-like model of decisions to classify a behavior. Using a multitude of such trees called a random forest, Larson et al have achieved high true positive rates
in cough detection - a mean of 92% sensitivity.\textsuperscript{44} Martinek \textit{et al} also express that decision trees can be very effective in distinguishing between voluntary coughs and non-cough sounds in healthy patients, with an 86% sensitivity and 91% specificity.\textsuperscript{43} However, they add that trees are not as accurate in patients with respiratory diseases. In such cases, sensitivity drops to as low as 28\%.\textsuperscript{43} Their work suggests that artificial neural networks are a better classifier for cough sound events than decision trees.

Artificial Neural Network (ANN) is a mathematical model and a method of classification which employs a similar structure as the biological neural network. ANN uses an interconnected group of artificial neurons which enables it to handle nonlinear classification and recognition problems such as speech and pattern recognition.\textsuperscript{52, 55} ANN is useful in cough processing because it is able to model complex relationships between the audio features and whether or not they pertain to a true cough event.\textsuperscript{52} By modeling the differences in spectral patterns, ANN can be used to classify sounds as cough events or not.\textsuperscript{51} Using an ANN classifier, Martinek \textit{et al} obtained an 82% sensitivity and a 96% specificity in detecting cough sounds.\textsuperscript{43} Drugman \textit{et al} have also demonstrated how increasing the number of neurons in an ANN classification improves the performance of the classifier. With an ANN configuration of 64 neurons, the study obtained a sensitivity of 94%. ANN therefore seems to work fairly well in identifying cough events.

A Hidden Markov Model (HMM) is a statistical model which can be assumed as a finite state machine that can change from one state to another at any time depending on a transition probability.\textsuperscript{42} HMM enables a characterization of time varying patterns as a parametric random process.\textsuperscript{56} Since it can efficiently model temporal variations of signals, it has been widely used for pattern classification and speech recognition. In \textsuperscript{39}, HMM is used as a basis for keyword-spotting in detecting occurrences of coughs in continuous recording. Here, they use HMM to characterize the time-varying patterns of the cough sound so it can be detected from the audio stream. They are able to achieve an average sensitivity of 82\%. Using a similar HMM approach, Matos \textit{et al} attain a median sensitivity of 85.7% and a specificity of 99.9\%.\textsuperscript{42}

The Gaussian Mixture Model (GMM) is yet another statistical model and a classification method where a mixture of Gaussian distributions is used to approximate the probability of belonging to a class.\textsuperscript{52} In cough processing, Gaussian models are widely used to model cough and noise events. In one study, extracted features from both cough and non-cough events are used to train two gaussian models.\textsuperscript{40} The trained models are then later employed to classify new events as either coughs or non-coughs. As in the above mentioned Artificial Neural Network classifier, detection accuracy is observed to improve with increasing Gaussians in the mixture model.
Drugman et al find that a GMM with 16 gaussians yield a sensitivity of 95.2% and a specificity of 94.3%.\textsuperscript{52} They add that GMM outperforms ANN and in particular, require way less features to attain equivalent performance rates. A comparison plot of all the mentioned classifier performances in Figure 5 (a) agrees that GMMs do outperform the other classifiers. We also implement some of these classifiers on our own cough and speech data and generated Receiver Operating Characteristic (ROC) curves in Figure 5 (b) also show that GMMs are superior to NN and Trees in cough classification.

With time, studies have begun to explore using hybrid classification methods that consist of two or more of the already mentioned classifiers. Some employ GMM to model sound events as inputs to HMM for the actual classification.\textsuperscript{39,42} Other studies have also considered using the ANN to represent the spectral features of the sound event, and feed its output to an HMM for handling the temporal variations in the signal.\textsuperscript{43,51} Generally these hybrid models are more efficient but however, require more computational power. With advancing processor technologies, we should expect to see more of such approaches in cough event classification.

IV. DISCUSSION

Having just discussed the analysis and classification of cough events, it is important to mention that the principal challenge in actually analyzing cough data is the time required for processing. Real-time processing would mean an immediate cough inference system, a feasible continuous sensing system, and a reduction in the amount of data stored. This would translate to better security and privacy of patient data. Yet, while most approaches to analyzing and characterizing cough sounds can be implemented efficiently given ample time, it becomes extremely
challenging when we consider performing all the processing in real-time. And if we only pursue extreme processing speeds, we will end up with a huge system with enormous processing abilities and power requirements, which will be very expensive and probably not applicable to wearable, cough sensing methods. One group demonstrated that a computer with 3.0 Ghz CPU and 2GB of RAM can implement an efficient processing algorithm on a 1s cough signal in about 506ms. Given that cough sounds have a typical length of about 350ms, this speed is not fast enough to avoid the risk of missing cough events. Also, even with the current advances in microprocessors, it is very challenging to pack an equivalent processing power into a wearable continuous sensing device. The Apple iPhone 5s has one of the most powerful smartphone microprocessors, and yet its A7 dual-core processor can only run up to 1.3GHz. And even if we manage somehow to overcome the computational barriers, we are still confronted by the challenge in meeting the enormous power requirements for continuous real-time processing.

REFERENCES


(19) Ireland AM, Wiklund I, Hsieh R, Dale P, O'Rourke E. An electronic diary is shown to be more reliable than a paper diary: results from a randomized crossover study in patients with persistent asthma. Journal of Asthma 2012;49(9):952-960.


(52) Assessment of audio features for automatic cough detection. 19th European Signal Processing Conference (Eusipco11); 2011.


(54) Speech Recognition using MFCC. International Conference on Computer Graphics, Simulation and Modeling (ICGSM’2012); 2012.


(57) Lal Shimpi A. The iPhone 5s Review: A7 SoC Explained. 2013 September 17.