KiMS: Kids' Health Monitoring System at Day-Care Centers using Wearable Sensors and Vocabulary-based Acoustic Signal Processing

Abhishek Basak, Seetharam Narasimhan and Swarup Bhunia
Electrical Engineering and Computer Science Department
Case Western Reserve University
Cleveland, OH, USA
E-mail: {axb594, sxn124, skb21}@case.edu

***Abstract*– Wearable sensors for healthcare and wireless health monitoring are rapidly becoming ubiquitous. They enable remote, accurate and low-cost health monitoring and can provide personal healthcare with timely detection of health issues. In this paper, we present a novel integrated system for monitoring children at day-care centers in order to facilitate proper care of health issues and overall wellbeing, including early detection of symptoms for various diseases, post-treatment monitoring as well as encouraging healthy habits and activities. The proposed "Kids Health Monitoring System", referred to as *KiMS*, is built around a wearable acoustic sensor with embedded digital signal processing capabilities in order to detect various audio signals of interest, such as coughs, sneezes, and cries. It is also equipped with wearable body temperature and pulse rate sensors, along with on-site processing and a Bluetooth unit for communicating alerts and activity on a timely basis. The record of a child's activities can be used by daycare specialist, parents or the healthcare provider for understanding the probable cause or time of onset of symptoms and encouraging healthy habits. This paper also presents a signal processing framework for feature detection and classification of various audio signals, under varying Signal to Noise Ratios (SNR).**

***Keywords*–** *Smart Health, Wearable Sensors, Acoustic Signal Processing, Health Monitoring.*

I. INTRODUCTION

Healthy children are the first step to a healthy adult population. In today's fast-paced world, it becomes difficult for many parents to continuously monitor the health and wellbeing of their children, who spend most of their day at a day-care center. This is especially true for children belonging to the age group of 2-5 years when they are unable to express themselves clearly and tend to become sick easily. As more and more children are sent to day-care centers, even the day-care service providers cannot adequately monitor each child individually. When the children return home, their parents are mostly unaware of the events which took place at the day-care facility. Events such as high rates of coughing and sneezing, frequent crying, vomiting, high periods of low activity, extended duration of sleep, low fluid consumption, and high body temperature might signify the initial stages of the contraction of an infection and hence must be brought to the notice of the day-care specialist as well as parents. In the absence of timely detection of such events and proper treatment, a child's health can deteriorate significantly.

Similarly, there are scenarios when children are not given enough time to recuperate, following treatment of an illness. They are sent back to the day-care within a short period, before complete recovery, often leading to a relapse of the disease and further complications. The time between ages 2-5 is the formative time when a child learns about healthy habits and needs to develop a sense of personal hygiene. In the absence of continuous parental guidance, a child may become prone to bad habits like not washing hands properly, not flushing toilets, drinking less fluids, etc. These habits are detrimental to their long-term health and wellbeing. Hence there is a need to have a simple yet accurate automated monitoring system for providing healthcare information about the kids. Some facts and figures for significant health issues, prevalent among kids in many countries over the world, are presented in Table I.

With growing advances in pervasive computing and development of miniature wearable sensors for monitoring health issues, one can envision an integrated system consisting of wearable non-invasive sensors with associated signal processing capability which can detect simple health-related events and record their rate and time of occurrence. Information about potential health-related issues can be

TABLE I. COMMON HEALTH ISSUES IN DAYCARE CENTERS

Health Issues	Symptoms & Activities	Facts & Figures
Ear infection (Otitis media)	Crying, sore throat, sneezing, diarrhea, fever, feeding and sleeping problems	About 4.65 million children in USA suffer from ear infection every year [2]
Diarrhea	High fever, bloody stools, vomiting, dehydration	About 9000 hospitalizations per year in USA for diarrhea [1]
Influenza	Mild fever, bouts of cough, sneezing, excess sleep, symptoms of fatigue	20-30% of children in USA contract influenza every year [1]
Diabtetes (both Type I & Type II)	Low activity, higher fatigue leading to excess sleep, excessive urination, high food intake	Each year over 13000 children in USA are diagnosed with Type I diabetes (Juvenile diabetes) [4]
Obesity	High food intake, low activity rates, high sleeping rates	By 2000, 22% of preschool children were overweight and 10% obese [4]
Whooping cough (Pertussis)	Mild (2-3days) followed by severe cough, whoop sound while inhaling, often vomiting	In 2010, 4017 cases of whooping cough reported in California, leading to death of 9 children [3]
Rotavirus Diarrhea	Fever, vomiting and diarrhea leading to dehydration	It is a leading cause of dehydrating diarrhea among kids in USA [1]

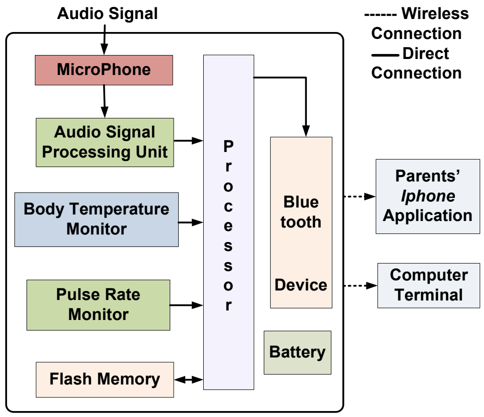


Figure 1. Overall view of the proposed KiMS system showing interconnections between the constituent components.

conveyed to the parents daily by sending a digest of events to them possibly via their smart-phone. Such information can also be used by the doctor or healthcare provider for gaining insight about onset or aggravation of certain symptoms. The system can also alert the day-care specialist upon detection of severe symptoms in order to take timely action. Our envisioned Kids' Health Monitoring System (or KiMS) is depicted in Fig. 1. It consists of some wearable sensors for detecting audio signals, and for monitoring body temperature and pulse rate. The system is scalable to incorporate other sensors and contains an integrated processor for detecting events of interest along with a non-volatile memory for storing relevant information. The Bluetooth transceiver conveys health related information periodically. The system contains a small rechargeable battery and should be capable of operating at ultralow power budget. It should occupy small area to be integrated into a wrist-band type device.

The proposed 'KiMS' system is beneficial for healthcare of kids at day-care centers, in the following four areas as shown in Fig. 2.

i) Early detection of infectious diseases leading to prevention of future complications and fatalities, e.g. if ear infection can be detected early, a kid can be saved from possible permanent damage of the middle ear lobe.

ii) Post-treatment monitoring of a child who has just recovered from an infection. This helps the parents and day-care specialists to detect relapse of infectious diseases like pneumonia, urinary-tract infection, malaria, etc.

iii) Monitoring practice of healthy hygienic habits by the kid at the day-care, e.g. proper hand washing after meals and after using the toilet, toilet flushing, drinking sufficient fluids, etc. Hand washing is claimed to be most effective for preventing infections that are passed from person to person.

iv) Detection of any chronic health issue in the child, e.g. general tendencies of coughing and wheezing, especially after high activity periods may indicate the possibility of the kid suffering from asthma. Besides Type I diabetes may be suspected if a kid has a general low activity and high sleep patterns over the day.

Building such a system involves several challenges, some which are addressed in this paper. The main contributions of this paper are:

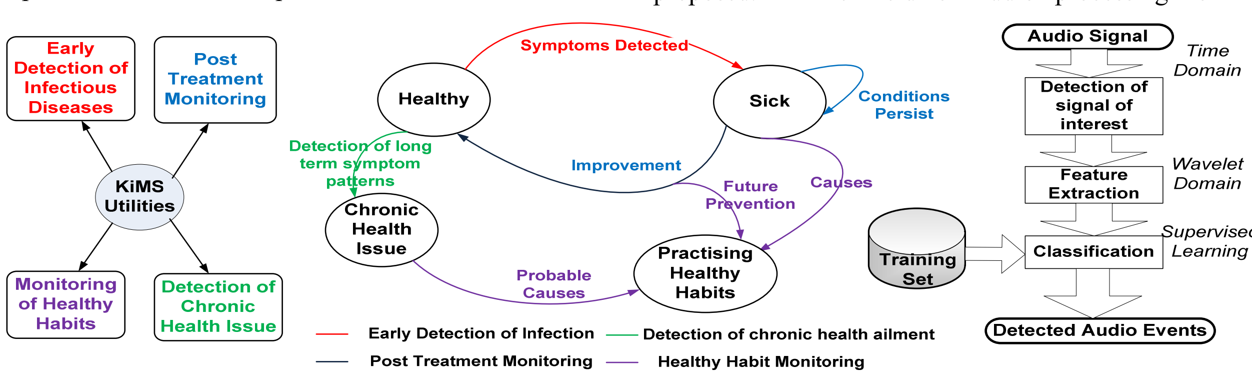
1) It presents an integrated scalable framework for monitoring the health and activity of kids at a day-care center. It proposes the use of non-invasive sensors for audio signal detection along with other parameter monitoring like temperature, pulse rate, and activity level for interpreting the overall health and wellbeing of the kid.

2) It presents a novel audio signal processing algorithm which can detect various events of interest using multi-resolution wavelet analysis. The wavelet domain feature extraction algorithm is shown to perform reliably under low Signal-to-Noise Ratio (SNR) scenario to detect various audio signals of interest. The vocabulary-based encoding enables effficient storage and wireless communication using minimum power and bandwidth resource.

In Section II, we present a background on related work for audio signal processing for smart health. In section III, we describe the KiMS framework and discuss the audio signal processing in details. The simulation results are presented in section IV and we conclude in section V.

II. Background

Research has been conducted in the past in the area of kids monitoring. Until now, mainly video monitoring systems (employing surveillance cameras) have been used for this purpose [5]. Besides, sensing mechanisms employing proximity sensors [6], wrist tags, as well as blood glucose level monitoring for diabetic children [7] have been proposed. In the field of audio processing for health



(a) (b) (c)

Figure 2. (a) Classification of KiMS system utilities into four broad areas. (b) A flow diagram illustrating the uses of the KiMS system in monitoring different states of health and wellbeing. (c) Audio signal processing algorithm.

monitoring, short-speech/distress/emergency signal detection has been employed for tele-monitoring of patients and elderly people [8]. Similar audio signal processing techniques are used for speech activity detection (SAD) [9] and speech recognition from non-speech signals [10], and denoising of speech signals for speech enhancement [11]. Among non-speech signals mainly coughing and snoring sounds have been analyzed [12]. In particular, frequency-based approaches, wavelet-based thresholding for denoising, hidden Markov and Gaussian mixture-based classification approaches have been pursued. In this paper, we propose a monitoring system primarily based on acoustic signal processing. It combines acoustic analysis with information from few other sensors to accomplish effective monitoring of children health issues and healthy habits in a day-care setting. We show that personalized online recording and analysis of audio signals can provide valuable health information, while being more attractive in memory requirement, system cost and privacy than video monitoring.

III. KIDS HEALTH MONITORING SYSTEM

KiMS is designed around a set of sensors – primarily audio sensors – for monitoring of various events and body parameters, which can be used as to analyze different health issues in children. The system can be attached around the wrist of a child in the form of a wrist-band. The major components of such an integrated system, as shown in Fig. 1, are:

1) *Microphone for recording acoustic signals.*
2) *An audio signal processing unit for detecting signals of interest, extracting features and classifying them into a set of relevant signals as cough, sneeze, cry, vomit, toilet flush and hand washing sounds.*
3) *A digital thermometer for periodic monitoring of body temperature.*
4) *Integrated pulse rate monitor.*
5) *A general-purpose microcontroller, which takes input from the audio signal processor, thermometer, and the pulse rate monitor and outputs probable health issues based on a collective decision-making algorithm.*
6) *A non-volatile memory to store the occurrence of various detected events of interest.*
7) *A battery for providing power to the system components.*
8) *A Bluetooth device to send the information to the parent's smart phone or a central host computer.*

Next, we list the major challenges for designing the proposed system and describe possible solutions for each:

i) Audio signal processing algorithm to identify unique sounds pertaining to cough, sneeze, cry, hand washing, toilet flushing, etc. from the background noise.
ii) Algorithm to detect and record various events relating to the health and wellbeing of the kid, based on the sensed signals.
iii) Method to efficiently store and transmit relevant information regarding the kid's health condition, based on the detected events.
iv) Introducing tunable parameters in the system which can be adjusted to the variability from kid to kid and as well as the environmental and temporal variations.
v) Scalability of the proposed system to incorporate more sensors to expand the scope of application.

A. Audio Signal Processing

The recorded audio signals have to be processed in order to identify different events of interest. The audio signal processing task has three main steps, as shown in Fig. 2c:

1)*Detection of 'signals of interest' in the time domain:* We search for 'signals of interest' in the recorded time domain signals, based on the real time detection of a 'collective burst of peaks'. Certain parameters need to be properly chosen for optimal detection performance. These parameters are as follows:

1) The amplitude threshold '**T**', above which a sample is detected as a peak (high thresholds leads to rejection of artifacts but may also miss detection of signal of interest whereas low thresholds lead to detection of two or more signals of interest as a single one, inclusion of artifacts, etc.); 2) The maximum interval '**d**' between two peaks for them to belong to the same 'signal of interest' (or, the minimum duration **'d'** between peaks of two consecutive distinct signals of interest); and 3) duration '**D**' below which a burst of peaks can be classified as an artifact or disturbance (reversely minimum duration for a burst to be recognized as a signal of interest). A real time recording of a sequence of cough signals with intermittent artifacts is shown in Fig. 3, with an illustration of these parameters.

For our set of audio data (all with a sampling frequency of 44.1 KHz), the results which lead to the selection of threholds for these parmeters are shown in Fig. 4. The value of k in the amplitude threshold T = (m + k*s) is varied in Fig. 4(a), where 'm' and 's' are mean and standard deviation of the absolute sample values over a time window of 4 sec. The minimum error in detection occurs at k=2 for our set of signals. The performance metric for detection is the number of bursts detected, with their durations and positions in comparison with the actual values. For the value of the duration 'd', the probability distributions of inter- and intra-peak intervals are compared. As seen from Fig. 4(b), there is

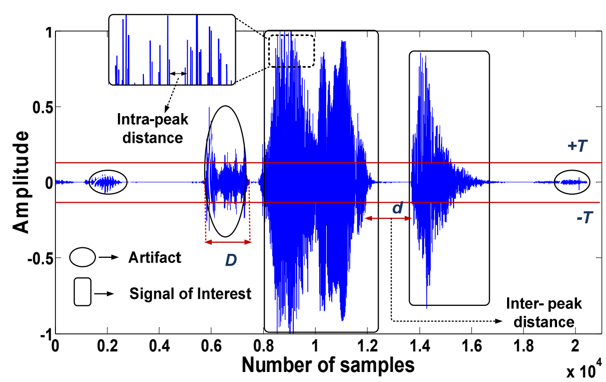


Figure 3. Real-time recorded cough signals in time domain.

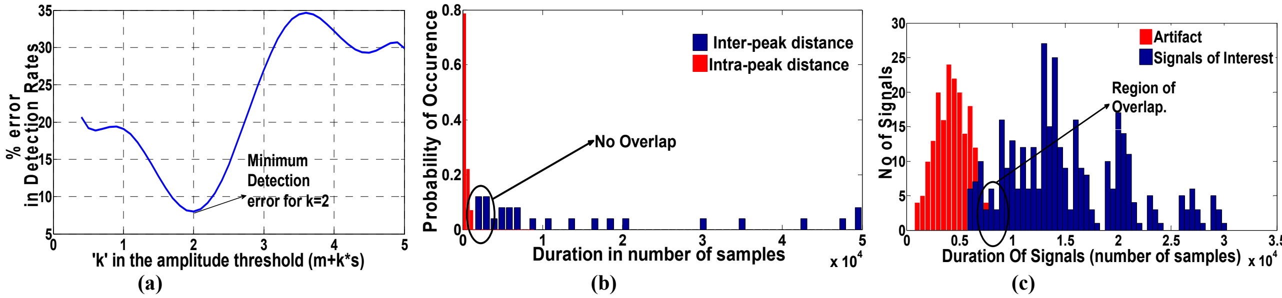


Figure 4. (a) Variation of detection rate with varying amplitude thresholds. (b) Probability distribution of intra-peak distance and inter-peak distance. (c) Histogram of artifact duration and duration of signal of interest.

no overlap in the two distribution plots. Hence 'd' is chosen to be the minimum duration between the first and last peaks of two consecutive signals of interest, which, in this case, is 2000 samples. For getting the optimized value of 'D', we compare the distribution plots of artifact duration and the duration of signals of interest. We observe from Fig. 4(c) that there is some overlap in the plots. As a result, we choose value of 'D' as minimum duration of a burst of peaks = 5300 samples, so that we do not miss any signals of interest. Any artifact included thereby in the detection phase would be distinguished from the other signals during the feature extraction and subsequent classification process.

2) *Extraction of features of the detected signals:* We perform feature extraction of the detected audio signals in the wavelet domain. Noise corrupts many of the features of the signals of interest in the time domain. The influence of noise on a particular feature of a cough signal, with decreasing SNR is shown in Fig. 5. In most of the cough signals, the majority of the burst of peaks occur in the first quartile of the signal duration, where a peak is a sample with ampliude greater than (T=m+2*s). We observe from Fig. 5 that with increasing noise levels and using the variable amplitude threshold T, bursts of 'peaks' are more uniformly distributed over the signal length. Hence, the feature used for identification of cough signals gets degraded with noise. Besides the frequency content of some of our signals of interest are similar and overlapping. Very minute differences in spectral content exist in different frequency bands, as observed from the spectrum of a cough and a sneeze signal in Fig. 6. Hence simple frequency-domain filtering cannot be used for extraction of unique features.

So the joint time-frequency i.e. the wavelet domain is

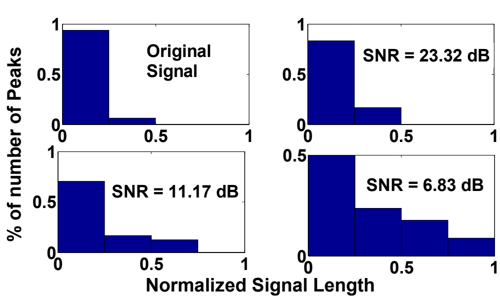


Figure 5. Loss of time domain feature (bursts of peaks only in initial part of signal) with increasing noise levels in a cough.

used for extracting features of the detected signals of interest. The wavelet transform retains both time and frequency information while reducing the number of samples at each level of wavelet decomposition. This helps us identify distinct hyperclusters corresponding to different features of these signals which can be easily computed in real-time using limited computational resources while allowing distinction between these signals on the basis of small set of features as described below. Detection of 'signals of interest' may be done in the wavelet domain to nullify the effect of noise on detected features, but that would require wavelet decomposition of the entire signal duration. This would require huge computational power and storage requirement from the system point of view. So detection is done in the time domain and the signals of interest are sent to the wavelet engine for feature extraction.

To obtain the optimal number of levels, a time domain feature and a frequency domain feature were tested for increasing levels of decomposition, corresponding to different signals of interest. The minimum of the sum of the errors of the two feature values (with respect to the feature values of original noiseless signal) occurs for the 3rd level of approximation. As shown in Fig. 7, the 3rd level of approximation provides a good trade-off between denoising (getting rid of time-domain noise superimposed on the desired signal) and preservation of original signal features. An approximation of the signal at a higher level of decomposition leads to smoothening of the signal (hence denoising), along with a loss of spectral content of the original signal corresponding to higher frequency bands. We denote the 3rd level approximation coefficient vector as a3 and extract various features such as:

i) The distribution of the number of peaks over the signal duration (a3), where a peak is defined as a sample whose absolute value exceeds (T=m+2*s), where 'm' and 's' are the mean and standard deviation of the a3 coefficients.

ii) The distibution pattern of the peak values with respect to the amplitude threshold, over the duration of a3 coefficients.

Fig. 8 illustrates these two features for the dataset consisting of cough, sneeze, cry and toilet flush sounds.The two features in conjugation also gives an accurate view of the energy distribution over the duration of the signal. High count of high amplitude peaks signify high energy in that duration band like the initial band of cough signals in Fig. 8

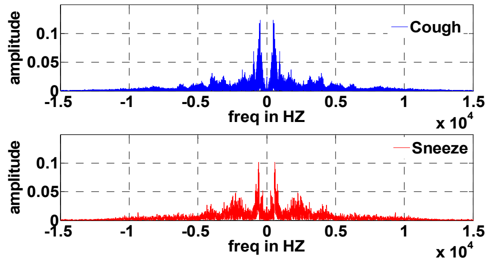


Figure 6. Frequency spectrum of a cough and a sneeze signal showing high overlap in spectral contents.

where as low count of low amplitude peaks correspond to comparatively lower energies like that of sneeze signals over the later half duration. High and low amplitude peaks may be defined as those whose absolute values are greater than 1.5 times and less than 1.2 times 'T', respectively. A small burst of high amplitude peaks and a large burst of comparatively low amplitude peaks may correspond to the same energies over that duration. Besides an estimation of the mean absolute sample value may be obtained from these two features. Bursts of peaks scattered over the entire duration of a3 signifies high mean value of samples like that in cry signals in Fig. 8.

Finally, we take inspiration from wavelet packet decomposition and modify traditional pyramid-based multi-resolution wavelet decomposition [13] in order to extract different levels of low-frequency and high-frequency coefficients corresponding to the different spectral clusters, in order to identify frequency-domain differences between the different signals. In order to reduce the aliasing effects of superimposed noise while estimating the general signal trends rather than every minute detail, we apply the multiresolution based wavelet decomposition on a normalized thresholded signal. The thresholded version constitutes only the weighted distribution of peaks in the original signal. The weights are proportional to the normalized difference between the peak value and the amplitude threshold 'T'. A thresholded version of a cough signal is shown in Fig. 9. Pyramid-based multiresolution wavelet decomposition is performed on this thresholded signal over 3 levels, as shown in Fig. 10. The normalized energy of the resulting 8 frequency bands, denoted as x1-x8 give an estimate of the spectral distribution of the signal. Generally this distribution follows a unique pattern for each signal type, with some extent of overlap. The distribution of the normalized energy of the 8 frequency clusters is shown for a single cough, sneeze, cry, and toilet flush signal in Fig.

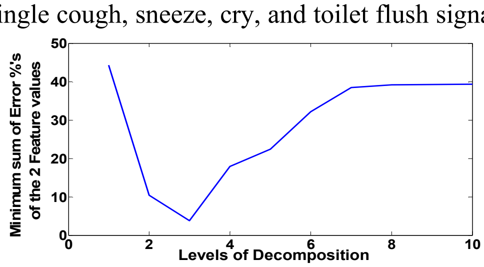
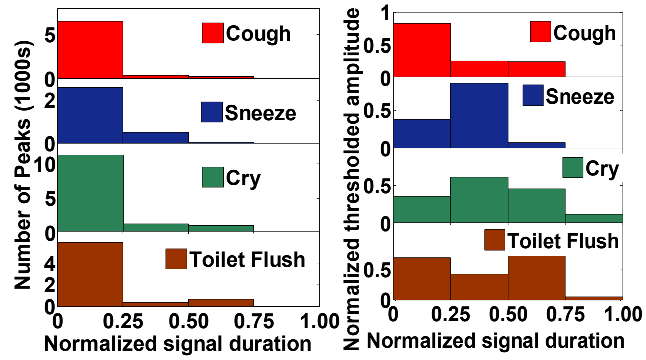


Figure 7. Variation of sum of errors of a time domain and a frequency domain feature with varying levels of decomposition.



(a) (b)

Figure 8. Distribution of (a) number of peaks, and (b) values of thresholded peaks, over the signal duration.

11. We have chosen the decomposition level as 3 because increasing it does not alter the classification results much, but increases the complexity and computational power from the hardware point of view. Reducing it to 2 leads to greater overlap among different signal types (only 4 frequency bands), thus detoriating the classification performance, as illustrated in Fig. 12.

3) Classification of the signals based on extracted features: Clssification is done through a non-linear classifier using the extracted features. The classifier is initially trained using a supervised learning network with a training set of features of the four different classes of data namely cough, sneeze, cry, and toilet flush signals. The features given as inputs to the network are as follows:

a) The number of peaks corresponding to each quartile over the total signal duration. This number is normalized with respect to the signal duration.

b) The normalized values of peaks, above the amplitude threshold, for each of the 4 quartiles of signal duration.

c) The normalized energy of each of the 8 frequency bands obtained by 3rd level wavelet decomposition of the normalized thresholded signal. The energy is normalized with respect to the total energy of the 8 bands combined.

We have used the Levenberg-Marquardt back-propagation training algorithm [14] as it provides good results in terms of performance (mean square error of classification and minimum false positive rates) as well as time for convergence for our data set. This algorithm uses an approximation of the Hessian matrix of the performance metric for updation at each iteration. It emulates steepest descent in the initial period (high performance error) and the

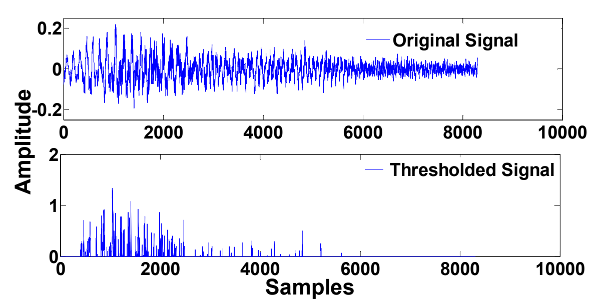


Figure 9. Original Cough signal and the thresholded version using a weighted distribution of peaks.

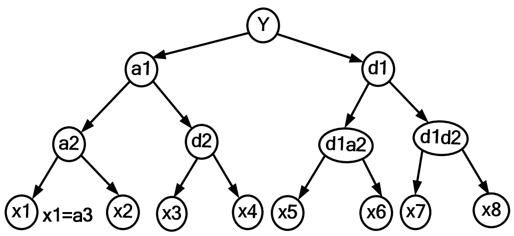


Figure 10. **Three levels of pyramid-based multiresolution wavelet decomposition of the thresholded signal.**

Newton's method near the error minimum. In the training set, 70% of samples were used for training, 15% for validation and the remaining 15% as a test set. Post-training, the trained classifier was used to test a completely independent test set of the 4 types of signals. For deciding upon the convergence of the training operation, two parameters were used namely the the performance error (m.s.e) and the consecutive number of validation failures. Empirically derived thresholds of 10^{-12} and 4 respectively were used for these two parameters.

B. Sensors

Apart from a miniature microphone which records audio signals, two other sensors are also used to monitor the following body parameters.

a. *Body Temperature*: A digital thermometer is integrated in the wrist band style device which will record the body temperatures at regular intervals. The temperature recordings coupled with inputs from the audio signal processing unit, can lead to the probable detection of various infections, e.g. patterns of higher body temperature coupled with periodic coughing, sneezing and extended sleep probably signals the onset of an influenza in the kid.

b. *Pulse Rate Monitor*: A pulse rate monitor is also integrated in the wrist-band device for measuring the pulse rate at regular intervals. The pulse rate information provides insight into general activity level, sleep patterns, etc. of the child concerned. A sharp decrease in pulse rate (usually to about 50-60, which can vary with individual subjects) over an extended duration, coupled with no major audio signals being recorded (only background sound) during that time, indicates a period of sleep. A pulse rate of more than 140 over some duration indicates a period of high activity.

C. Detection of Health Issues and Symptoms

Sudden increase in the rates of various audio signals of interest like cough, sneezes, crying, vomiting etc. could indicate potential contraction of an infection by the kid and hence the requirement of proper steps of diagnosis and care. Besides, these audio signals can assist in interpreting the progress of the kid in terms of recovery from an infection. Information regarding relapse of the disease can be formulated by observing the audio signal patterns over the day. Periodic body temperature monitoring can lend vital data for diagnosis and treatment. Besides, monitoring the pulse rate as well as frequency of speech signals would enable one to understand the activity and sleep patterns of a

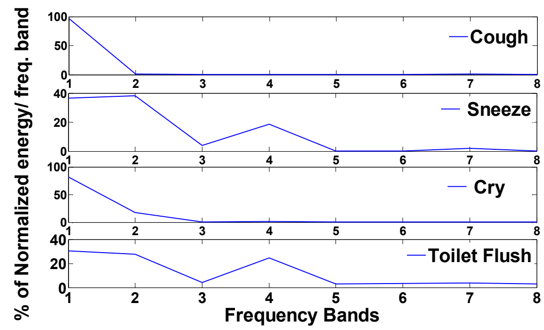


Figure 11. Distribution of normalized energy in 8 different frequency clusters for 4 types of signals.

child. Activity monitoring is important for obese and diabetic kids, for whom higher activity should be encouraged. Sleep patterns provide an insight into the general well-being of the child. Healthy habits like proper washing of hands, flushing of toilets, etc. can be monitored by measuring associated audio signals like distinct sounds of flowing water, toilet flushes, etc. Fig 2.b shows a general flow of KiMS system utilities in detecting different states of health and wellbeing.

D. Event Storage and Transmission

We use a vocabulary-based encoding scheme [15] to store and transmit the occurrences of events. As the number of events we are detecting ranges from 8-10, we use 4 bit encoding to store the occurrences of events, e.g. we encode occurrence of cough as '0000' and sneeze as '0001', decrease of sleep as '0010' etc. On detection of an event of interest, the corresponding code is stored in the flash memory along with its particular time of occurrence. The time of occurrence can also be encoded according to different durations of the day. 3 bits may be used to identify 8 different time periods of the day. Often in many cases, the general time trends (duration of the day) of occurrences of events like increase of body temperature, increase in sneeze rates, etc. are more useful for clinicians for diagnostic purposes, rather than the exact time of occurrence. Hence this methodology of encoding of time of occurrence of events serves our purpose, with the advantage of low storage, low power requirement etc. Similarly frequency of occurrence within pre-specified duration is also encoded.

A flash memory stores the counts of coughs, sneezes, crying as well as the body temperature and general activity of the kids over the past 2-3 days. This history information is coupled with the currently detected event data to help in identifying symptoms of various health-related issues. For example, increase in number of coughs and increase in sleep indicates onset of influenza. When required for transmission via Bluetooth, the record of events is transmitted using a random key-based scrambling for data security over the wireless channel [16]. This entire encoding scheme is very efficient from the perspective of transmission bandwidth reduction, leading to low-power operation, as well as reducing the requirement of storage space, which can help reduce the size, weight and cost of the system. As an example, detection of an event like a cough in absence of insitu signal processing and vocabulary based encoding would entail, sending the whole data set (size of order of magnitude 4) over the transmission channel, thus requiring huge bandwidth and power requirement. In contrast, in presence of the above features, only 7 bits is required (4 for the event and 3 for the time of occurrence) for transmission. In addition, the system can raise an alert to the day-care specialist if urgent care is required. These signals can be sent via Bluetooth to the phone of the day-care specialist or enable a beeper, associated with unique ID for each kid.

E. System Calibration

There is a provision for initializing and calibrating the KiMS system according to a child's body and activity patterns, e.g. the normal pulse rate, pulse rates during eating and sleeping, time of the day when the child usually takes his/her meal (after which he/she is supposed to wash hands properly) can be registered before usage or in a periodic manner during a short. Based on this calibration step, the tunable parameters of the algorithm for event detection and alert notification are set. Periodic calibration over the lifetime of the device can help to adjust for sensor drift due to aging and environmental variations.

F. Scalability

The KiMS system is scalable in the sense that new sensors can be added to it to extend its monitoring capabilities and the application domain. For monitoring the amount of food and liquid intake, a weight sensor, preferably as a pressure sensor at the shoe sole, can be added to the system. Slight increase in body weight after food intake (till digestion sets in) may be measured by the pressure sensor and recorded along with the time of measurement. The sensor can coordinate its measurement with the main system via Bluetooth. A weight sensor can help monitor the feeding patterns of obese and diabetic children. Information about a child's food and liquid intake patterns may help in diagnostic purposes as well. The KiMS system can also incorporate a dust or pollen sensor to monitor the pollen levels in the day-care atmosphere, thus helping to diagnose allergy-related issues in kids. An application scenario where the KiMS system might come in handy is the early detection of a contagious infectious disease for preventing outbreak and spread of the infection in the day-care center. This includes diseases like influenza, rotavirus diarrhea, hand foot and mouth disease, and chicken pox. The framework of monitoring health and wellbeing in kids can also be extended to home care of elderly people. The information from the KiMS system can specially be helpful to the in-house healthcare providers or nurses for monitoring the elderly people. In developing countries, where the incidence of health-related problems is relatively higher, this low cost, simple and accurate monitoring system can be highly beneficial to both the people and the healthcare agencies.

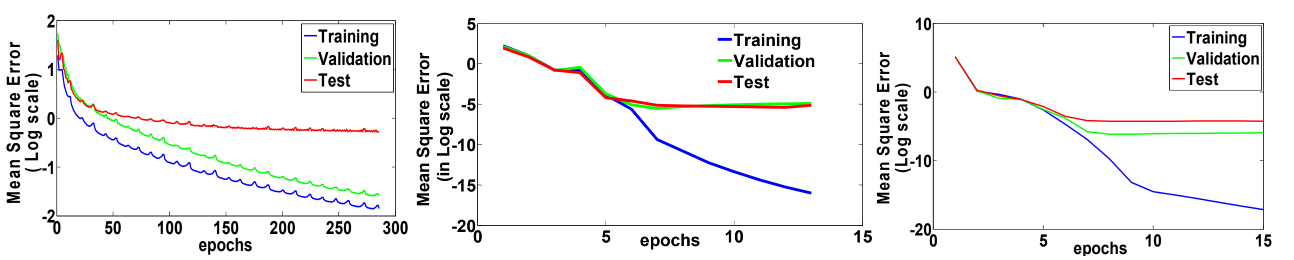
IV. RESULTS

A. Data Set

Our data set comprises of signals of 4 non-speech classes namely cough, sneeze, cry, and toilet flush, respectively. Some of these signals were collected from available sound databases of kids in the internet like the Audio Micro Stock library [17], Universal Soundbank etc. Others were recorded in real-time using wireless microphones attached to the body or close to the body at various distances. In total, they comprise of 142 cough signals, 78 sneezes, 89 cry signals and 61 toilet flush signals. Besides 22 different physiological noises like the falling of a pencil, talking in the background, distant noises of moving vehicles, etc. were recorded as artifacts, for testing the performance of the algorithm.

B. Simulation Results

We test our algorithm in the detection and classification of the 4 types of signals. The classifier is a feed forward back-propagation neural network, with one hidden layer of 30 neurons. The nonlinear classifier is trained using the Levenberg-Marquardt training algorithm. 70% of each of the 4 different types of signals was used for the training phase, while the remaining for the test phase. Each data vector consisted of 16 features (8 time-domain and 8 spectral-domain). The training results are presented in Fig. 12(b). As observed from the figure, the training is satisfactory with low mean square error for the training set. Furthermore, the validation and test performance curves follow a similar pattern i.e. they achieve their minimum around the same iteration point, thus eliminating the possibility of an over fit. The trained classifier was tested with the independent test set (30% of the data, which was not used in the training phase) for overall performance, the results of which are presented in Fig. 13. The classification error was less than 9% for sneeze, cry, and toilet flush sets and for cough it was slightly higher at 11%. Here, a signal is correctly classified if the classifier output is within a tolerance band of 10% of the actual target output. Classification error for a class is the percentage of



(a) (b) (c)

Figure 12. Training performance of the non-linear classifier with (a) 2, (b) 3, and (c) 4 levels of wavelet decomposition. The performance is poor for 2 levels and improves with increase in levels. There is minimal improvement in performance of the classification algorithm from 3 to 4 levels.

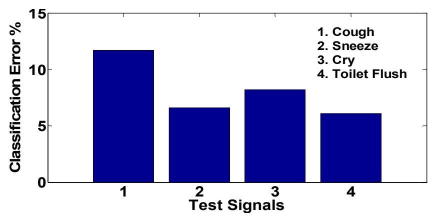


Figure 13. Classifier performance for independent test set.

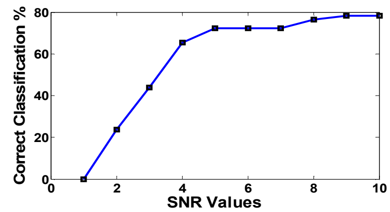


Figure 14. Percentage of correct classifications of cough with SNR.

misclassifications with respect to the total number of signals for that class. Those signals whose output values do not fall within the tolerance band of any target class value are classified as 'unclassified'. They probably represent the 'artifacts', included in the 'signal of interest' detection phase. The performance for each class is shown in Table II. Simultaneously, we also test the robustness of our algorithm i.e. test its performance in the presence of noise. To simulate this, we add Gaussian noise of increasing Signal to Noise Ratio (SNR= Signal Power/Noise Power). We illustrate the results of the varying detection rates of signals with increasing SNR values in Fig 14. It is observed that the algorithm performance is satisfactory (Correct Classification % > 75) up to an SNR value of 4.8, below which the performance degrades drastically.

V. CONCLUSION

We have presented a wireless health monitoring system for monitoring children health in day-care facilities. The device consists of a combination of sensors to collect information about meaningful events such as extent of coughing, sneezing, activity level, and amount of sleep, which can be used to predict health issues, diagnose symptoms, and monitor healthy habits. The device can potentially be useful for different parties, namely parents, doctors/nurses treating children, and the day-care managers. We have presented the sensing mechanism and the necessary signal processing algorithm to identify relevant events. Future investigations will include making the algorithm more robust, further reducing classification errors and including other acoustic signals for detection. Blind source separation techniques can be investigated for isolating an event from a mixture of simultaneously occurring events. An example would be when a child coughs and washes hands at the same time or when two nearby children cough and sneeze simultaneously. Independent component analysis can be used in such cases to separate individual sources. In the event of a kid coughing/sneezing, it will be recorded in a nearby kid/kids' monitoring system. However as acoustic intensity follows an inverse square distance relationship, the recorded cough/sneeze intensity of a nearby child will be low (below the amplitude threshold) and will not pass as a signal of interest in the initial phase of detection. Further investigation will also include building a prototype of the proposed system and its subsequent validation in-field.

TABLE II. % OF TRUE DETECTIONS AND MIS-CLASSIFICATIONS

Actual \ Detected	Cough	Sneeze	Cry	Toilet Flush	Unclassified
Cough	**88.3**	6.2	3.5	0	2
Sneeze	3.3	**93.4**	1.1	1.1	1.1
Cry	3.7	2.1	**91.8**	1.3	1.1
Toilet Flush	0	2.1	3.2	**93.9**	0.8

ACKNOWLEDGEMENT

The authors acknowledge valuable inputs on the technical content of the paper from Dr. Soumya Ray, Department of Electrical Engineering & Computer Science, Case Western Reserve University.

REFERENCES

[1] "Children's Health Topics-Infectious Diseases," http://www.aap.org/healthtopics/infectiousdiseases.cfm

[2] "Repeated Ear Infections seem to plague white kids, poor kids," The Laryngoscope, Vol. 120, No. 8, Aug 2010.

[3] "California has 4017 cases of Whooping cough and 9 deaths," http://www.medicalnewstoday.com/articles/201681.php

[4] "How to keep your kids Diabetes-free," http://diabetes.about.com/od/childrenanddiabetes/kidsandobesity.htm

[5] "Network cameras bridge the gap for working parents," http://byremote.net/clients/success/packerplayland

[6] " IonKids Child Monitoring System," http://www.diginetlink.com

[7] "Continous Glucose Monitoring for Diabetes," http://www.diabetes.webmd.com/continuos-glucose-monitoring

[8] Vacher M., Istrate D., Besacier L., Serignat J.F., Castelli E., "*Sound detection and classification for medical telesurveillance,*" 2nd International Conference on Biomedical Engineering, BIOMED Proceedings, ACTA PRESS, pp 395-398, Feb 2004.

[9] Shafiee S., Almasganj F., Vazirnehzad B., Jaffari A., "*A two stage speech activity detection system considering fractal aspects of prosody*," Pattern Recognition Letters, Vol. 31, No. 9, pp 936-948, 2010.

[10] Hung L.W., Kai L.Y., "*Classification of non-speech human sounds*," Proceedings of the 2009 IEEE International Conference on Systems, Man and Cybernetics, Oct. 2009.

[11] Meyer J., Simmer K., "*Multichannel speech enhancement in a car environment using weiner filtering and spectral subtraction,*" IEEE ICASSP,1997.

[12] Artyomov E., "*Manometry based cough identification algorithm*," International Journal of Information Theories and Applications, 2007.

[13] Mallat S., "*A Theory for multiresolution signal decomposition: The Wavelet representation*," IEEE Trans. Pattern Anal. Machine Intelligence, Vol. 11, No. 7, pp 674-693, 1989.

[14] Levenberg K., "*A method for the solution of certain non-linear problems in Least-Squares*," The Quarterly of Applied Mathematics.

[15] Narasimhan S., Zhou Y., Chiel H.J., Bhunia S., "*Low power VLSI architecture for neural data compression using vocabulary based approach,*" IEEE BioCAS, 2007.

[16] Narasimhan S., Wang X., Bhunia S., "*Implantable electronics: Emerging design issues and an ultra light weight security solution*," IEEE EMBC, 2010.

[17] "Audio Micro Stock Audio Library," http://www.audiomicro.com