Automated Ingestion Detection for a Health Monitoring System

William P. Walker, *Member, IEEE*, and Dinesh K. Bhatia, *Senior Member, IEEE*

***Abstract*—Obesity is a global epidemic that imposes a financial burden and increased risk for a myriad of chronic diseases. Presented here is an overview of a prototype automated ingestion detection (AID) process implemented in a health monitoring system (HMS). The automated detection of ingestion supports personal record keeping which is essential during obesity management. Personal record keeping allows the care provider to monitor the therapeutic progress of a patient. The AID-HMS determines the levels of ingestion activity from sounds captured by an external throat microphone. Features are extracted from the sound recording and presented to machine learning classifiers, where a simple voting procedure is employed to determine instances of ingestion. Using a dataset acquired from seven individuals consisting of consumption of liquid and solid, speech, and miscellaneous sounds, $> 94\%$ of ingestion sounds are correctly identified with false positive rates around 9% based on 10-fold cross validation. The detected levels of ingestion activity are transmitted and stored on a remote web server, where information is displayed through a web application operating in a web browser. This information allows remote users (health provider) determine meal lengths and levels of ingestion activity during the meal. The AID-HMS also provides a basis for automated reinforcement for the patient.**

***Index Terms*—Health monitoring, obesity management, patient empowerment, patient monitoring.**

I. Introduction

OBESITY is an epidemic spreading throughout the world [1]. With obesity and overweight comes an increased risk of a myriad of chronic diseases such as type II diabetes, hypertension, coronary heart disease, and sleep apnea [2], [3]. Along with these and other comorbidities, obesity and overweight results in a reduced quality of life [4] and imposes a financial burden on the health care system [5].

Principal to obesity and overweight management is the concept of energy balance, consisting of energy expenditure (EE) and energy intake (EI) [6]. An imbalance between EI and EE will result in a change in weight. Many sources contribute to EE such as resting metabolic rate (RMR), thermic effect of food (TEF), and activity energy expenditure (AEE) [7]. The effect of TEF is minimal when compared with RMR and AEE. The RMR varies between individuals and involves factors out of the immediate control of the individual such as age, gender, and body composition [6]. Various formulas exists that can provide an estimate to an individual's RMR [8]–[10].

Personal record keeping is an important aspect to weight loss and weight maintenance [11]. For self-reporting of AEE, the current market is saturated with respect to devices to monitor AEE from pedometers [12] to heart rate monitors [13] to multisensor armbands [14], which require the individual to log the information captured, either manually or through an interface provided with the device.

Presently, self-reporting for EI consists of maintaining a food diary or the completion of a food frequency questionnaire. There are a few drawbacks to these methods of self-reporting, where individuals may have differences in motivation to complete and maintain an accurate food diary as well as a lack of awareness of food intake (the exclusion of reporting snacking for example) [15]. For energy balance monitoring, new, more accurate, methods of measuring EI are needed [16].

A new direction in EI measurement is currently under way and is taking the form of electronic sensing devices. The ultimate goal behind this new direction is automated dietary monitoring, where sensors are used to detect and annotate instances of EI, as well as the substance being consumed, thus replacing the need for self-reporting [15]. This goal, however, is not realistic given the current level of technology [17]. What is possible with current sensor technology, is to detect possible instances of ingestion, creating an automatic activity log. This activity log may be used to "jog" the memory of the individual for further information, as proposed by Amft *et al.* [17], [18].

There are three areas of focus for automated EI detection using sensors, chewing, swallow, and eating gesture detection [15]. Chewing detection can be done with piezoelectric strain sensors and microphones [15], [17], [19]–[22], swallow detection can be done with surface electromyography (SEMG) and microphones [21], [22], and gesture detection is typically performed using accelerometers [15], [18].

EI is solely performed through ingestion of solids and liquids, and is generally consumed orally [6]. In other words, every substance that contributes to total energy input is swallowed. For this reason, and due to the simplicity of implementing an external throat microphone [23], detection of ingestion via swallow sounds is pursued in this study.

The authors envision the following four possible application areas for ingestion detection from swallow sounds. The first three are areas to supplement an existing obesity management therapy and the last is a possible area of analysis unrelated to obesity/overweight management:

1) Full day monitoring;
2) Meal length monitoring;

Manuscript received February 4, 2013; revised May 11, 2013 and August 8, 2013; accepted August 13, 2013. Date of publication September 9, 2013; date of current version March 3, 2014.

The authors are with the Embedded and Adaptive Computing Group, the University of Texas at Dallas, Richardson, TX 75080 USA (e-mail: will.walker@mathworks.com; dinesh@utdallas.edu).

Digital Object Identifier 10.1109/JBHI.2013.2279193

3) Liquid/fruit intake therapy adherence monitoring;
4) Monitoring for stroke/dysphagia afflicted individuals.

Full day monitoring addresses one of the failings of a food diary, omitted instances of ingestion [15]. The sensing platform (external throat microphone) is worn during the day, detection time periods where the individual may be consuming food or liquid that may normally go unnoticed. The goal for this area is to provide feedback to the individual, reminding them of these ingestion periods, so that proper recording may be performed.

One of the major drawbacks to the full day monitoring application is user acceptance. As was found during the creation and implementation of the BodyMedia SenseWear armband [14], [24], [25], placement and size of the sensing device plays a crucial role for viable implementation. Sazonov *et al.* mention that the neck microphone used for swallow detection could be incorporated "as a medallion on a neck band" [20].

Meal length monitoring is less restrictive than the full day monitoring. Meal length monitoring can provide information as to how long the individual's meals lasted. This may be helpful for dieticians associated with the individual's therapy, as synching meal length with the individual's recorded intake can help to identify an approximate quantity consumed. For example, if the individual registers intake of a "salad" and the sensor detected the meal lasting 30 min with a high level of ingestion activity (few breaks), then an indication of size of the salad or possibly omitted foods is given.

After reviewing various interventions for weight loss/management, Sharma [26] advises increased fruit intake and water consumption, which can increase levels of satiety without a high amount of caloric intake [27], [28]. From experimentation, swallows of a liquid (or solid with a large amount of liquid such as fruits) have been observed to produce stronger sound characteristics than swallows of dry solids. The strong signals can be an indicator as to adherence to a therapy involving increased water and fruit/vegetable intake.

Although not explicitly the focus of this study, a myriad of studies have been performed with respect to swallow sound analysis/detection for dysphagic individuals [29]–[35]. It may be possible to take the swallow sound detection process presented here and implement it in a system that helps to identify dysphagic individuals as well as help monitor recovering stroke victims, who may have difficulty with the swallowing mechanism.

The primary goal for the study presented here is to provide a framework that offers an increase in the level of communication between individual and health professional. The increased levels of feedback and user interaction, even Internet-based interaction, can increase the efficacy of a weight loss/management therapy [36]–[40].

Presented here is an overview of the entire system for the detection and monitoring of ingestion activity, termed AID-HMS. The remainder of this paper is organized as follows. The following section introduces previous works in swallow detection and analysis. The next section offers a brief overview of the AID-HMS designed. The following section describes the components and implementation of the AID process. This paper is concluded with a discussion of the results and future work.

II. Related Works

There are three phases to the physiological process of swallowing, oral preparation, pharyngeal, and esophageal [41]. The oral preparation phase is where the food is prepared into a swallowable bolus. In the pharyngeal phase, the food bolus travels through the pharynx using a sequence of muscle activations. In the esophageal phase, the food bolus is propelled toward the stomach.

The oral preparation phase typically involves chewing, and can be addressed using the chew detection sensors described previously. The pharyngeal and esophageal phases are associated with the properties of swallowing. It is not possible to perform sensing during the esophageal phase in a noninvasive manner. Therefore, the pharyngeal phase is ideal for performing noninvasive swallow detection [21], [31].

One note needs to be made concerning the three phases. Although monitoring the pharyngeal phase is ideal for externally located sensors (microphones, accelerometers, SEMG, etc.), it was found during the research presented here that sounds emitted during the oral phase for food preparation via chewing emitted sounds with very similar characteristics to swallow sounds. In this sense, the oral phase may be used to determine ingestion activity as well.

From a large dataset of manually annotated swallow sounds, Sazonov *et al.* [19], [20] found a relation between swallowing event frequency (rate between successive swallows) and the type of substance being ingested is liquid, solid, or nothing at all. The ingestion of a liquid typically resulted in a smaller time between successive swallows (rate of 15 to 20 swallows per minute) than for solids (5 to 10 swallows/min) and no substance (around 2 swallows/min). Furthermore, using the measurements taken of average bolus size for each individual, a rough estimate of the amount of substance ingested was made. The results indicate a possibility of differentiating substance types and amount based on the frequency of swallows.

The general approach toward identifying swallow sounds is to take a feature and perform some decision method such as thresholding or classification. Amft *et al.* [21], [22] combined swallow sounds and an SEMG signal. Classification was performed using the Naïve Bayes and k-Nearest Neighbor classifiers using a set of time and frequency-based features such as peak count, spectral power, spectral fluctuation, spectral roll-off, and band energy. Also, explored was classification of bolus size and viscosity. A fair level of accuracy and precision was reported, where accuracy is the measure of true positives (swallows) detected and precision is a measure of false positives.

Makeyev *et al.* [42] use the 2-D plots of the wavelet decomposition-based scalogram and the short-time Fourier transform (STFT)-based spectrogram over windows of swallow sounds as inputs to a limited receptive area neural classifier (LIRA), which is a four layer perceptron used in image recognition. Very high swallow detection accuracy is reported, however, the number of neural network nodes needed is very large (5000 to 8000).

Sazonov *et al.* [43] implemented a support vVector machine (SVM)-based classifier using the wavelet packet decomposition (WPD) and mel-scale Fourier transform (msFT) as feature inputs. The WPD was found using the Coiflet 4 wavelet and evaluated using the unbiased entropy. The msFT was found by applying a series of triangular critical band filters on the spectral output from the Fourier transform, where the spacing of the filters is based on the mel-scale. The time series signal was segmented into overlapping epochs, where each feature is calculated and the results classified.

In their exploration of detecting swallow sounds using the subsonic range ($\leq$5 Hz) and sonic range (20–2500 Hz), Fontana *et al.* [44] removed unwanted vocal sounds using a second, ambient microphone. Over ranged when the ambient microphone detected noises, the signal from the throat microphone was canceled out. Thersholding was performed on the resulting signal after filtering, to detect instances of swallows.

Nonlinear-based approaches have also been applied to the problem of swallow detection [29], [30], where an attempt at modeling the underlying tracheal system was made via delay embedding using Takens method of delays [45]. The features used were based on the recurrence plots and waveform fractal dimension (WFD). The classification of these features using simple thresholding and classification using hidden Markov models (HMM), to detect the states of the swallow sounds, was explored. Another method explored is thresholding of a decomposed signal obtained from the discrete wavelet transform, with a high level of reported accuracy [33].

Factors inherent to the nature of swallow sounds make comparison between methods of detection difficult. Unlike the electrocardiogram (ECG) signal, a standard database of swallow sounds does not exist, as the signal is highly dependent on the method of capture (placement of microphone/accelerometer, preamplification, etc.). The signal may also vary by the type (liquid or solid) [34], viscosity, and volume [21], [22] of the substance swallowed, making the accurate detection of swallows difficult, especially when attempting to discern swallows from other sounds such as voice.

The inclusion of nonswallow sounds is crucial for examining the effectiveness of an automated ingestion detection process. Some of the previous works mention the inclusion of nonswallow sounds such as voice; however, the relative amount of noise included generally remains ambiguous. During the design of the AID-HMS, specific focus was given toward the inclusion of noningestion related noises, such as coughing, clearing of the throat, movement, and speech.

III. AID-HMS Overview

The AID-HMS is a prototype, where an AID process is implemented over an HMS framework. The HMS framework supports the automated acquisition, transmission, storage, and display of biometric information obtained from sensing devices. The process of data capture, transmission, storage, and display from sensors such as a pulse oximeter, blood pressure sensor, and digital weight scale has been previously demonstrated [46]–[49].

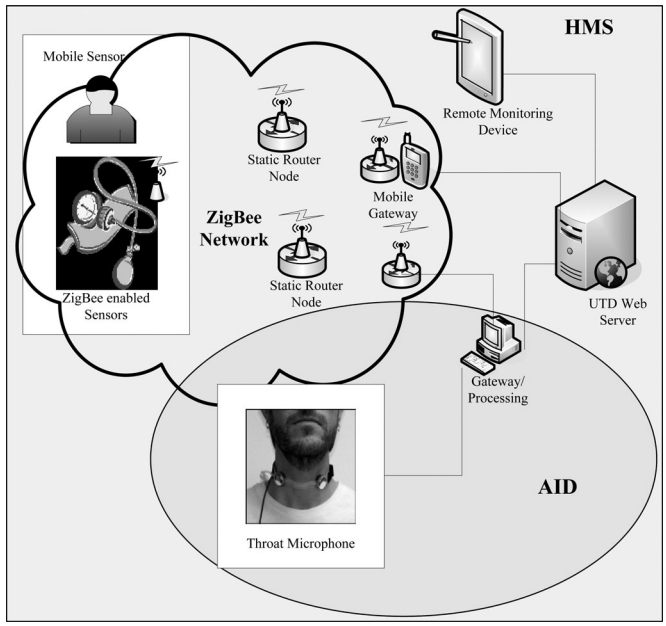

Fig. 1. AID-HMS structure. The AID process, ellipsis, is implemented over the HMS. Zigbee enabled sensors allow mobility throughout the wireless network. The gateway device is any device with the ability to form a TCP/IP Socket connection with the remote web server.

The sensors interface with small-form processing devices termed "Biote," designed in-house. The Biotes forward data through a Zigbee wireless network to a network coordinator. The network coordinator interfaces with a gateway device, where the sensor data are forwarded to a remote server. The server houses an MySQL database as well as a web application. The web application provides remote monitoring by displaying data acquired through a web browser. This structure is depicted in Fig. 1.

The gateway device may be any device with access to the internet. A TCP/IP socket is used for communicating with the web server. The AID process is comprised of the iASUS NT3 Throat Mic System (TMS) [23] connected to an integrated sound card on the gateway/processing PC. The future iterations of the AID process will connect the TMS to a mobile gateway device.

All information obtained from the sensors (blood pressure, pulse oximeter, digital weight scale, and TMS) is associated to an individual based on an account ID and password, maintained on the remote server through the web application. Any authorized user may view stored biometric information from an individual, which allows monitoring of trends and promotes feedback. Authorization is sharing based, where an individual may grant other users access to their data. An example would be an individual sharing their data with their primary care physician. The AID process and integration in the HMS are the focus of this study, and are discussed in the following section.

IV. AID Process

The AID process is comprised of five stages: data acquisition, event detector, feature extraction, classification, and ingestion response, depicted in Fig. 2. The input to the AID process is a time series recording obtained from the TMS, sampled at

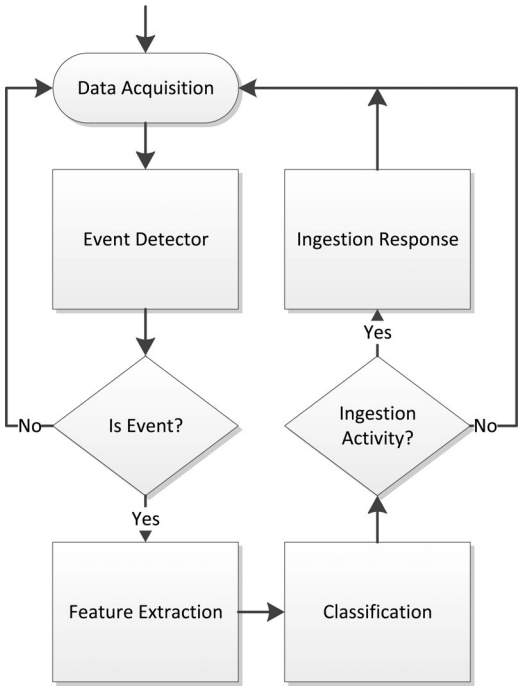

Fig. 2. AID process flow chart, containing the five components of Data Acquisition, Event Detector, Feature Extraction, Classification, and Ingestion Response. The input to the AID process is a time series recording obtained from the TMS, sampled at 44.1 kHz. The output of the AID process is an indicator of ingestion activity, stored on the HMS database.

44.1 kHz. The output of the AID process is an indication of ingestion activity, stored on the HMS database. The following subsections discuss each component.

A. Data Acquisition

The TMS from iASUS is used for acquiring throat sounds. The TMS consists of an elastic band containing two surface microphones. The TMS is placed around the throat over the laryngeal prominence and is held in place by a magnetic clasp. All sound signals presented were obtained by connecting the TMS to the microphone input of an integrated sound card on a personal computer (PC) running Windows 7. Microphone levels were set to a volume of "5" and a boost of +20 dB. Recording was performed at a sampling rate of 44.1 kHz at 24 bits per sample.

A graphical user interface (GUI) was designed in MATLAB to perform the recording. The GUI performs mapping of the user identifier, for storage and retrieval in the database, and defines the method of processing. Currently, layered recording is implemented, alternating between two recording processes. When the AID process begins, a timer is set for 1 min and the first recorder starts recording. When the timer fires, after 1 mine, the second recorder starts recording. The first recorder is then stopped, and the minute long recording is processed. This procedure repeats until the application is stopped.

The time length for each recording (1 min) was arbitrarily chosen. The AID process is not dependent on the length of the time-series recording, within reason (long enough to contain one or more swallows). Future iterations will include implementation on a small-form processing platform. For such an implementation, testing is required to find a suitable sound card and interface that offers similar results to those achieved using the PC. Processing is performed by passing the minute long sound signal to the proceeding components.

B. Analysis Dataset

For determining the most effective parameters, methods, features, and classifiers presented in the following sections, a dataset from seven individuals was acquired. A GUI was designed in MATLAB (separate from the one used during the data acquisition component) to obtain recordings. The term "event" is used to indicate a type of sound present in the sound recording. During recording, events are manually annotated, using the GUI, as belonging to one of five event typessuch as Swallow of a Solid (So), Swallow of Liquid (Li), Swallow of No-Substance (Ns), Vocalization (Vo), and Miscellaneous (Ms), during the recording process using the GUI. The Vo event type contains all instances where the vocal cords are activated for speech. The Ms event type consists of sounds such as a cough and clearing of the throat.

The swallows in the dataset consist of ad-libitum consumption of a peanut butter and jelly sandwich portion, ad-libitum drinking of a carbonated beverage, consumption of an apple and orange, and drinking of water. In all instances, the bolus size was up to the discretion of the individual, allowing for variation in bolus sizes. Across the seven individuals, a total of 39 min of recordings were made, with 262 annotations of swallow events and 25 Ms events. Annotations were not made of individual Vo events, instead, an indicator of start and end of each Vo segment is created. The reason for not annotating each event is the length of the event varies between each individual, where words may meld into one another, creating longer events. A recording of speech was taken from each individual.

C. Event Detector

The event detector is a light weight processing method that condenses a long time-series recording into zero or more discrete events of much smaller length. The event detector does not distinguish between event types; it simply highlights the portions of the recording which need further analysis. By reducing the number of data samples examined within each recording segment, more computationally expensive methods can be performed to place the events into the appropriate types.

A sixth event type was created during the analysis portion, False Event (Fe). The Fe event type is a detected event that does not match an annotation. It was observed that most Fe events occurred during the ingestion process, typically resulting from noises produced during the oral stage of the swallowing process. These events have very similar characteristics as swallow sounds, and as such, are very difficult to separate. Further, depending on the individual, single or multiple cleanup swallows were observed, where the entire bolus is not transmitted during the initial swallow. The cleanup swallows were not annotated during the recording process, and are thusly detected

as Fe events. Approximately 200 Fe events were present in this dataset.

Unlike the ECG signal, it is possible that an event may not occur during the recording segment. For this reason, an adaptive threshold method is not employed. Instead, a modified version of the Pan and Tompkins QRS detection method [50] is used. The detected events are further reduced based on the typical duration of swallow events, where swallows do not generally exceed 0.8 s in length [31], [51]. The exclusion of events longer than 0.8 s negates some vocal segments, where spoken words may form longer events. Also based on observations, a window for the event energy is enforced, as many swallow events fall within a small distribution. The event energy is found by taking the aggregate of the squared signal. The implementation of the energy window negates some strong Ms events such as coughing.

Parameters for the Event Detector were chosen such that the most of the annotated swallow events (98%) are detected and passed on to the next component and the number of detected nonswallow events is minimized with respect to the examined parameter choices.

D. Feature Extraction

The feature extraction stage takes each individual event (from the event detector) as input and outputs a sequence of mapped values corresponding to selected features used during classification. During analysis, many of the time/frequency domain features described in the previous works were examined, as well as some features available in the Voicebox toolbox for MATLAB [52]. Additionally, various information cost functions [53], [54] were evaluated over decompositions based on the WPD from 51 distinct wavelets [55].

The time/frequency domain features were examined with varying parameters for signal segmentation and feature creation. Parameters tested for segmentation were for segment size, segment overlap, and window function applied. Each feature had varied options for creation, such as returned coefficient count, type of critical band filter to apply, etc. For the WPD- based features, the parameters varied consisted of eight cost functions applied over 11 levels of decomposition.

Based on the resulting output distribution from each feature and combination of parameters, 198 feature parameter sets over four separate feature types, cepstral coefficients (CC), frequency band energy (FBE), spectral autocorrelation (SPacorr), and spectral autocovariance (SPacov), and 17 wavelet decomposition (WD)/cost functions were selected. The SPacorr and SPacov are the spectral analysis of the autocorrelation and autocovariance, respectively. The result was a set of 215 parameter sets, where a parameter set consists of a feature, segmentation parameters, and creation parameters, or in the case of the WPD feature, a wavelet, cost function, and decomposition level. The selected parameter sets gave the best separation between annotated swallow events and nonswallow events, using the above dataset from seven individuals.

Figs. 3 and 4 depict resulting distributions for two parameter sets, from the CC and WPD feature types, respectively. In Fig. 3, each event is mapped to a single scalar value based on the particular parameter set. Evident is the relatively close distribution of swallow event types compared with the distinctly different distribution of Vo events. The Fe events are nearly indistinguishable from the swallow events. The solid line depicts a threshold, where at least 90% of swallow events and no more than 18 Vo events exist on the same side. The threshold line is only used as a metric for analyzing the distribution.

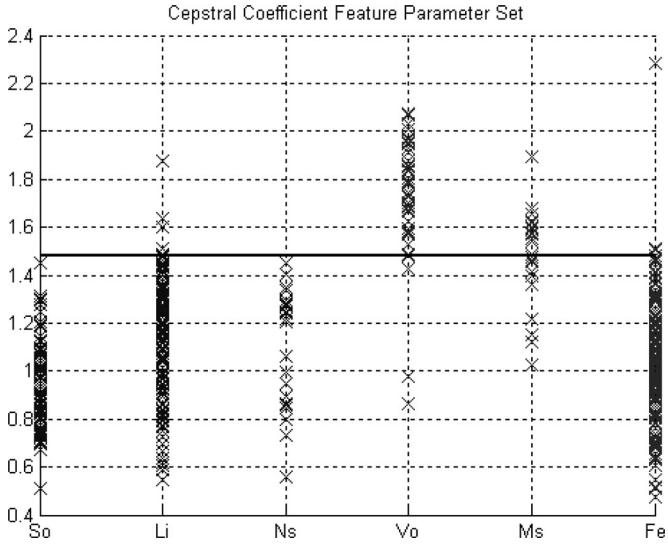

Fig. 3. Resulting distribution from a single parameter set for the Cepstral Coefficient feature type. Each event from the analysis dataset is depicted as a plot point, ordered by event type (horizontal axis). The solid line depicts a threshold where at least 90% of swallow events and no more than 18 Vo events exist on the same side. The threshold line is only used as a metric for analyzing the distribution.

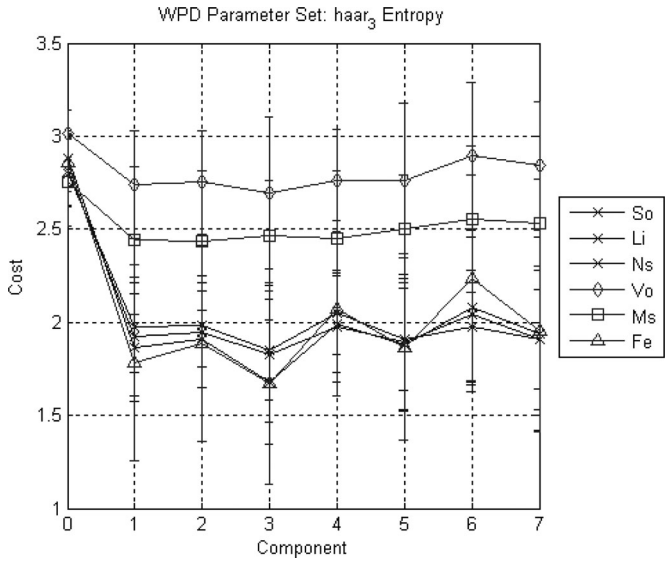

Fig. 4. Resulting distribution from a single parameter set for the WPD feature type. The wavelet used is the Haar wavelet [56]. The decomposition level offering the best distribution (separation between the swallow events and the Vo event type) was level 3. Evaluating the entropy [53] at this level of decomposition yields a vector of length 8, where each component corresponds to a frequency range. The values depicted are the mean $\pm$ standard deviation for each event type, evaluated for the parameter set. A large separation from swallow events and Vo events is visible.

The distribution depicted in Fig. 4 is for a WPD parameter set, where each plot line depicts the average $\pm$ standard deviation for each event type evaluated using the WPD parameter set. Similar to the output distribution depicted in Fig. 3, a large separation is seen between the swallow event types and the Vo event type. A modest separation between the Ms event type and swallow event types is also evident. The Fe event type is nearly indistinguishable from the swallow event types.

The performance of a classifier is dependent on the choice of feature, where the best classifier for one feature may not be the best choice for another feature [57]. Examining every possible parameter set and classifier was not technically feasible; hence, the reduction of parameter sets to 215. The output distribution from each parameter set gives an indication of suitability for classification. Parameter sets offering a greater distinction, may offer a stronger classification.

Based on classification performance (the next stage), these 215 selected parameter sets were further narrowed to a small collection, where each parameter set is matched with a classifier. As aforementioned, the output from this stage is a mapped version of the input event. The output may be a vector sequence or a scalar, depending on the parameter set.

E. Classification

In the classification stage, the output from the feature extraction stage is matched with a machine learning classifier. During analysis, 16 classifiers belonging to the six families of Naïve Bayes (NB), discriminant analysis (DA), support vector machine (SVM), k-nearest neighbor (kNN), regression tree (Reg Tree), and Bayes net (BN), were examined for each of the selected 215 parameter sets. The classifiers used are available in the bioinformatics [58], statistics [59], and BNET [60] toolboxes for MATLAB. These classifiers were selected based on the availability and ease of implementation, other methods of classification may offer improved results and exploration of these are saved for future work.

1) Individual Classification: Binary classification was selected, due to the ease of implementation, where events were classified as either *Ingestion* or *Noningestion*. Classification performance was measured based on correct *Ingestion* event classification, true positive rate (TPR), and incorrect *Noningestion* event classification, false positive rate (FPR), and overall classification accuracy, resulting from 10-fold cross validation [61]. Classification accuracy is the number of correctly classified events over the total number of classified events. A feature parameter set and matched classifier are referred to as a parameter set/classifier combination within this study.

In the ideal case, a parameter set/classifier combination would be able to detect each separate event type of So, Li, Ns, Vo, Ms, and Fe. This classification proved extraordinarily difficult, especially given the similarities between Fe and annotated swallow events So, Li, and Ns. As aforementioned, the Fe events were observed to coincide with ingestion activity, including cleanup swallows, where a bolus required multiple swallows to completely clear the oral and pharyngeal phases.

Naturally, to determine the performance of the classification process, the events in the dataset must be distributed into one of the two classification categories of *Ingestion* and *Noningestion*. This placement dictates training and allowed the authors to ascertain correct/incorrect classification during testing. The authors explored four ways of placing the events into the two classification categories. The first (Ca_1) placed So, Li, and Ns in the *Ingestion* category and Vo, Ms, and Fe in the *Noningestion* category. The second (Ca_2) placed So and Li in the *Ingestion* category and Ns, Vo, Ms, and Fe in the *Noningestion* category, which is the most restrictive of the four event placements explored. The third (Ca_3) placed So, Li, Ns, and Fe in the *Ingestion* category and Vo and Ms in the *Noningestion* category. The Ca_3 placement resulted from the observation that the Fe events occurred mostly during the ingestion process. The fourth (Ca_4) placed So, Li, Ns, Fe, and Ms in the *Ingestion* category and Vo in the *Noningestion* category. The Ca_4 placement is the least restrictive of the four, and only serves as a means of differentiating vocalizations from all other sounds picked up by the TMS.

Each of the 16 classifiers was matched with each of the 215 feature parameter sets, yielding 3440 parameter set/classifier combinations. The parameter set/classifier combinations were evaluated using each of the four event type placements $Ca_1 - Ca_4$, resulting in a total of 13 760 classification sessions. The performance of a classification session was measured based on the results from 10-fold cross validation, achieved by first splitting up the entire dataset into ten disjoint subsets using the MATLAB function *crossvalind*. This function provided a roughly equal distribution of the event types among each of the subsets. The ten iterations of training and classification were performed, where each of the ten subsets is left out exactly once, to be used for testing, and the remaining nine are used for training. The performance of a classification session is the mean performance over the ten iterations, measured on the test sets.

The performance of each classification session was measured by classification accuracy, TPR, and FPR, which were found using the MATLAB function *classperf*. Fig. 5 depicts the receiver operating characteristics (ROC) curve for the top performing classifiers for each of the 215 feature parameter sets and each of the four event type placements $Ca_1 - Ca_4$. The top performing classifier was selected based on an empirically derived linear combination of classification accuracy, TPR, and FPR, where slight preference was given to the TPR results. In the ROC plot, values toward the upper left corner indicate better performance. Correct classification was determined based on how the classifier was trained using one of the four event type placements $Ca_1 - Ca_4$. Evident in Fig. 5 is the impact of the event type placement. The reduced TPR experienced by the Ca_1 and Ca_2 event type placements is due to the confusion introduced by training the classifiers to place the Fe events into the *Noningestion* classification, as the Fe events were observed to share similar characteristics with the So, Li, and Ns events.

2) Multiple Classifier System: The top performing parameter set/classifier combinations, depicted in Fig. 5, were grouped to form a multiple classifier system (MCS). The MCS style implemented is a simple parallel structure with a majority voting

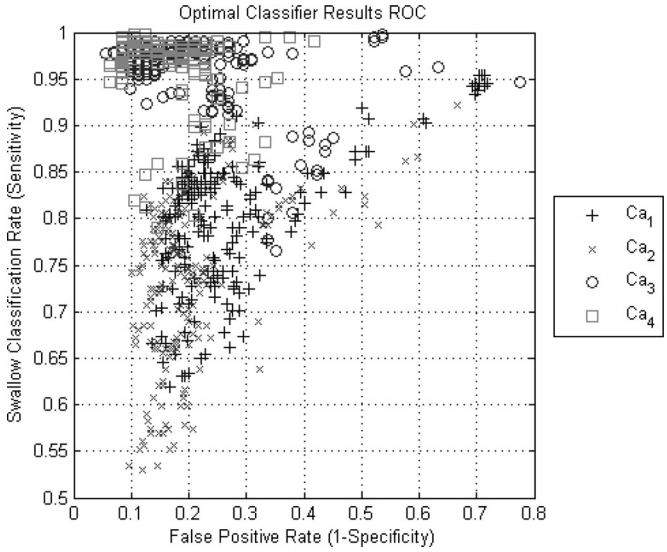

Fig. 5. Receiver operating characteristics (ROC) curve depicting TPR (rate of correctly classified *Ingestion* events) versus FPR (rate of incorrectly classified *Noningestion* events) for the top performing classifiers for each of the 215 feature parameter sets and each of the four event type placements $Ca_1 - Ca_4$. The top performing classifier for a feature set and event type placement is the classifier that offers in the best score using a linear combination of classification accuracy, TPR, and FPR.

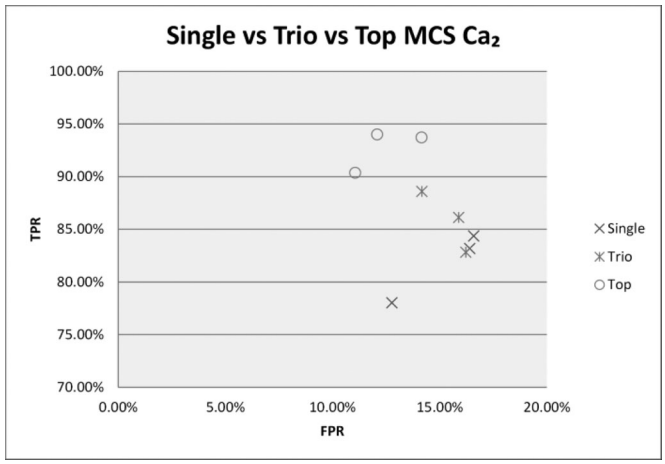

Fig. 6. ROC plot for MCS groups with Ns, Vo, Ms, and Fe considered *Noningestion* (Ca_2). The comparison is made between MCS groups consisting of a single parameter set/classifier combination (*Single*), three parameter set/classifier combinations of the same feature type (*Trio*), and the MCS group without restriction on group size (*Top*). The three *Single* plot points depicted consist of one parameter set/classifier combination with the CC feature type and two with the WD feature type. All three *Trio* plot points depicted consist of parameter set/classifier combinations with the WD feature type. The MCS groups without size restriction, *Top*, offered an improved performance over both *Single* and *Trio* and are listed in Table I.

rule for classification decision [57]. This structure was chosen due to the simplicity of implementation. The Future iterations of the AID-HMS will pursue other means for creating the MCS such as bagging or boosting, which may improve classification performance. The goal for an MCS is to combine the results from as independent as possible classifiers to achieve a higher level of classification performance [57].

A voting procedure was implemented, where the decision threshold (number of votes needed to classify an event as *Noningestion*) was varied to determine the threshold offering the best classification performance. A total of 267 MCS groups were examined. An MCS group is a collection of one or more parameter set/classifier combinations, where each combination submits a vote as to the class of an unknown event. The method of MCS group creation was varied, where parameter set/classifier combinations with unique feature parameter sets are used, repeated feature parameter sets are used as long as the classifiers belong to differing families.

MCS group creation was implemented as follows. The features discussed previously were split into five categories, CC, FBE, SPacorr, SPacov, and WD. All combinations of one, two, three, four, and five feature categories were formed (ignoring order and with no repeated feature categories), where for each category the top performing one, two, and three parameter set/classifier combinations were selected, where the same feature type was used as long as the classifiers belonged to different families. This process resulted in the 267 MCS groups mentioned previously.

From the four ways of event type placement mentioned previously $(Ca_1 - Ca_4)$, the placements Ca_2 and Ca_3 were used for the creation of the MCS groups and for the measurement of MCS group performance. Event type placement Ca_2 has the effect of measuring the ability to discern annotated swallow events with a bolus (solid or liquid) from all other sounds, and event type placement Ca_3 has the effect of measuring the ability to discern annotated swallow sounds and Fe events from Vo and Ms events. The top performing parameter set/classifier combinations when using Ca_2 were not necessarily the same when using Ca_3; thus, the 267 MCS groups were formed based on Ca_2 and a separate 267 MCS groups were formed based on Ca_3.

3) MCS Performance: Fig. 6 depicts the ROC of MCS groups consisting of a single parameter set/classifier combination (*Single*), three parameter set/classifier combinations of the same feature type (*Trio*), and the MCS group without restriction on group size (*Top*) when using event type placement Ca_2 for training and classification. The results were obtained using tenfold cross validation, discussed previously. Each of *Top*, *Single*, and *Trio* were selected based on the scores obtained using linear combination of TPR and FPR (TPR added to one minus the FPR, where a higher result is deemed better).

The MCS groups without size restriction offered better TPR and FPR over the single parameter set/classifier combination. The *Top* plot points in Fig. 6 correspond to the MCS groups depicted in Table I, which lists the feature type and the corresponding classifier. The decision threshold was 0.45. The cost function implemented for the WD feature type is listed as well. The cost functions listed are the compression number (CN), compression area (CA), L_p norm (L_pN) [54], and the minimum description length (MDL) criterion [62], [63]. SVM-RBF is the SVM classifier implementing the radial basis function, QDA is quadratic DA, D-QDA and D-LDA are QDA and LDA, respectively with a diagonal covariance matrix, M-DA is DA using the Mahalanobis distance, NB-N implements the Gaussian dis-

TABLE I
TOP MCS GROUPS FROM FIG 6 WITH DECISION THRESHOLD 0.45

MCS_A		MCS_B		MCS_C	
Feature	Classifier	Feature	Classifier	Feature	Classifier
WD/MDL	BN	WD/MDL	QDA	WD/MDL	BN
WD/CN	BN	WD/MDL	kNN-10	WD/CN	BN
CC	SVM-RBF	WD/CN	kNN-10	CC	SVM-RBF
CC	kNN-10	CC	kNN-10	CC	kNN-10
CC	M-DA	FBE	kNN-5	SPacorr	NB-K
SPacorr	NB-K	SPacorr	NB-K	SPacov	NB-K
SPacov	NB-K	SPacorr	M-DA		

TABLE II
TOP MCS GROUPS FROM FIG 7 WITH DECISION THRESHOLD 0.3

MCS_D		MCS_E		MCS_F	
Feature	Classifier	Feature	Classifier	Feature	Classifier
WD/CA	kNN-5	WD/CA	kNN-5	WD/CA	kNN-5
WD/L_pN	kNN-10	WD/L_pN	kNN-10	WD/L_pN	kNN-10
WD/CN	NB-K	CC	D-QDA	CC	D-QDA
CC	D-QDA	CC	kNN-5	CC	NB-K
CC	NB-K	CC	NB-K	SPacov	D-LDA
FBE	NB-K	FBE	NB-K	SPacov	NB-N
FBE	kNN-10	SPacorr	M-DA	SPacov	SVM-C
SPacov	D-LDA				
SPacov	NB-N				
SPacov	kNN-10				

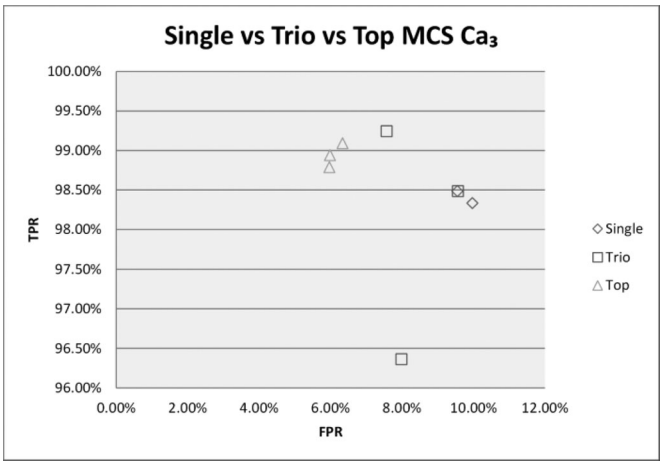

Fig. 7. ROC plot for MCS groups with only Vo and Ms considered *Noningestion* (Ca_3). The comparison is made between MCS groups consisting of a single parameter set/classifier combination (*Single*), three parameter set/classifier combinations of the same feature type (*Trio*), and the MCS group without restriction on group size (*Top*). All three *Single* plot points depicted correspond to individual parameter set/classifier combinations of the CC feature type. Two of the *Trio* plot points correspond to parameter set/classifier combinations with the CC feature type and the third with the SPacorr feature type. The MCS groups without size restriction, *Top*, offered an improved performance over both *Single* and *Trio* and are depicted in Table II.

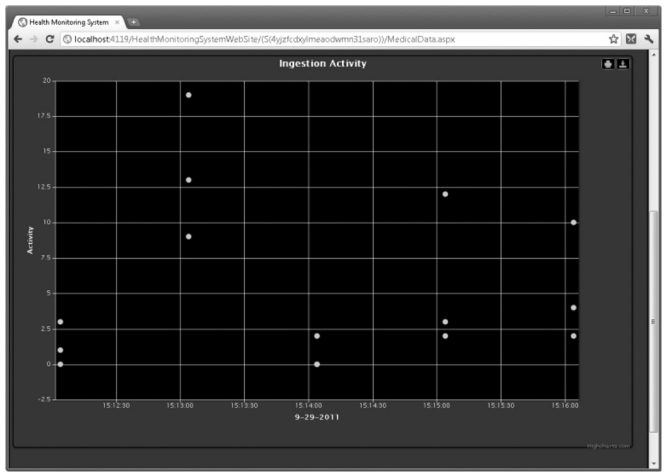

Fig. 8. Ingestion activity of an individual over a 5 min period. Three MCS groups (MCS_A, MCS_B, and MCS_D from Tables I and II) are used to perform an estimate of ingestion activity over each minute in the 5 min period. Over the 5 min period, the individual was speaking throughout the first minute, drinking only, during the second minute, speaking during the third minute, and eating an apple during the final two minutes.

tribution, and NB-K implements the kernel smoothing density estimate.

Fig. 7 depicts a similar ROC to that shown in Fig. 6, with the exception that placement Ca_3 was used for training and testing. As with Fig. 6, each of *Top*, *Single*, and *Trio* were selected based on the scores obtained using linear combination of TPR and FPR (TPR added to one minus the FPR, where a higher result is deemed better). The MCS groups without size restriction, depicted in Fig. 7 as the plot points for *Top*, are listed in Table II with each feature and corresponding classifier.

The MCS groups in Fig. 6 offer fair performance with respect to cancelling out all annotated nonswallow events, as well as Fe events, with approximately 10% to 15% incorrectly classified as *Ingestion*. The FPR is higher for each plot point in Fig. 6 than those in Fig. 7 due to the placement of Fe events. For the classifiers in Fig. 6, a classification of an Fe or Ns event as *Ingestion* is considered a false positive. For the classifiers in Fig. 7, a classification of an Fe or Ns event as *Ingestion* is considered a true positive. These results indicate a similarity exists between the annotated swallow events and the Fe events, as it is easier for classifiers to treat Fe events as *Ingestion* rather than *Noningestion*, hence the smaller FPR found in Fig. 7. The results depicted in Fig. 7 indicate a small performance gain using the MCS groups shown in Table II over the MCS groups consisting of three parameter set/classifier combinations of the same feature type (*Trio*). Approximately 6% to 8% of annotated Vo and Ms events are incorrectly classified as *Ingestion* using the MCS groups from Table II.

F. Ingestion Response

The final stage in the AID process is the ingestion response. The results from the Classification stage are passed to the HMS database, where the information may be remotely viewed. Presently, offline training is performed for the MCS groups. During processing, each unknown event is evaluated according to each feature in the selected MCS groups. Due to the differences in strengths of the MCS groups, more than one is implemented, offering a "spread" approach where ingestion activity is determined from different viewpoints. Fig. 8 is a screen shot of the web application depicting 5 min of recording of an individual, processed through the AID-HMS. The plotting tool used is the JavaScript library Highcharts [64].

Three MCS groups are used to perform an estimate of ingestion activity over each minute in the 5-min period. Over the

5-min period, the individual was speaking throughout the first minute, drinking only, during the second minute, speaking during the third minute, and eating an apple during the final 2 min. The three MCS groups used were MCS_A and MCS_B, listed in Table I and MCS_D listed in Table II. The selection of these MCS groups was based solely on observation, other combinations of MCS groups may yield better results; this exploration is left for future work.

During each of the speaking segments, the three MCS groups detect very few, if any instances of activity. Each MCS group detects a high level of activity during the liquid consumption segment. The authors observed that swallows of a liquid typically had much stronger sound characteristics than swallows of a solid. The low values reported for the consumption of the solid from MCS_A and MCS_B are due to the training of Fe events as *Noningestion*; thus, some of the actual swallows of solid are also classified as *Noningestion*, due to the reasons discussed previously. Furthermore, while consuming a solid, the frequency of swallowing is generally lower than for liquid consumption, due to the need for bolus preparation (chewing). MCS_D, however, detects a high level of activity during the solid consumption portion, which includes detection of some Fe events.

V. Discussion and Future Work

The plot given in Fig. 8 depicts an interesting result. If all three MCS groups detect a high level of activity, then the ingestion is likely liquid. If MCS_D detects a high level of activity and the two other MCS groups detect a modest level of activity, then the ingestion is likely solid. If none of the MCS groups detect a high level of activity, then there is likely no ingestion or a high amount of voice.

From experimentation, it was observed that the Fe events most resembled Ns events. Depending on the size of the bolus, Li and So events may tend toward characteristics of Ns events. With So events, if the individual employs multiple swallows to clear a single bolus from the mouth, then each swallow will consist of a smaller size. Therefore, So events may tend more toward characteristics of Ns events, and will be more likely to be excluded by MCS_A and MCS_B as seen in the above plot.

The higher level of ingestion activity for the liquid ingestion over the solid consumption portion corresponds with the results found in [19], [20]. Intuitively, the number of swallows of solid within some time frame should be lower than swallows of a liquid, as some amount of time is incurred due to the chewing and bolus preparation process.

It should be noted that the results of this process do not necessarily have a direct relation to caloric consumption. The intent is to capture periods of elevated activity to aid in the recording process. As an example use case, imagine a patient records the consumption of a sandwich during lunch. During this lunch period, the AID-HMS records a high level of ingestion activity over a period of 30 min. Reviewing this information may trigger further inquiries into the nature of the sandwich or if other items may have been consumed but not recorded. One of the envisioned end goals for this study is to help bridge the communication gap between patient and health care professional.

The entire process, from time-series recording and processing to storage on the remote server, is automated. The time required to calculate the classification using each of the three MCS groups for each minute long window is variable depending on the number of events found by the Event Detector. For recording segments with many events, the classification takes approximately 14 s total.

Naturally, the execution time will increase on a smaller processing platform. Processing is currently performed as a sequence of serial operations. The entire MCS process is highly parallelizable, however. Performing some of the operations in parallel using a platform such as the OMAP 5 Platform from Texas Instruments [65] may provide a suitable level of performance.

The results from the sound processing are immediately viewable from a remote location using a web browser. This allows the health professional to determine adherence to a prescribed therapy, with respect to ingestion instances. Furthermore, based on the observations about the characteristics of solid and liquid swallow detection, a rough estimate may be made concerning the type of food ingested. This may help in identification of increased water and fruit intake.

Presently, the AID-HMS requires a connection to the processing PC. Furthermore, steps are being made toward interfacing with mobile gateway devices such as a smart phone, which adds a higher level of mobility.

A portion of the AID process is currently implemented in the C programming language, to provide a lower execution time. The remaining components are also in the process of being implemented in the C and C++ languages for the conversion from processing performed on a PC to small-form processor.

This study was based on a very limited dataset, and the results are preliminary; however, this study provides a stepping stone for future attempts at the automation of detecting ingestion activity. Further testing is required, including the addition of a larger sample size of individuals, to determine applicability to a larger population. An adaptive classification process is being explored, where the offline training presented here is used to determine a suitable training set during operation, which will then be used to mold the classifiers toward the sounds emitted by the individual. Furthermore, methods for implementation of the MCS such as bagging or boosting may yield better results and will be explored in future iterations.

The ingestion activity plots in the HMS web application provide a tooltip describing the level of activity and the time recorded. Future iterations will offer the ability to add a food diary entry over a time span, allowing remote viewers access to both activity level and type of substance as recorded by the individual.

VI. Conclusion

A process for detecting ingestion is implemented as a prototype system, AID-HMS, and presented here. The process specific to ingestion detection, AID, is comprised of five components, Data Acquisition, Event Detector, Feature Extraction, Classification, and Ingestion Response. The data used to deter-

mine ingestion activity is captured using an externally placed throat microphone. The Event Detector is a lightweight processing method which finds events of interest in the recording. These events are sent to the next components in the AID process to perform identification. The method of identification is an MCS, where multiple features and classifiers are combined to vote on the classification of an unknown event.

From tenfold cross-validation testing, the top performing MCS groups detect ingestion activity correctly over 90% of the time. This offers an indication of the strength of the groups of features and classifiers to detect ingestion.

From a combination of MCS groups, it was observed that when all groups detect a high level of ingestion activity, liquid ingestion is likely. When none of the MCS groups detect a high level of activity, then the segment likely contains noningestion sounds. When the MCS group trained to accept Fe events detects a high amount of activity, and the MCS groups trained to reject Fe events detect a low level of activity, swallows of a solid are likely.

The AID-HMS provides a means to supplement an obesity/overweight management therapy. The periods of ingestion that may be neglected by the individual with respect to record keeping may be highlighted. Furthermore, an indication of the length of meal periods can be made using the AID-HMS, and an estimate of type of substance (liquid or solid) ingestion may be made. The entire AID-HMS process is automated, requiring very little execution time compared to recording length. Future iterations will include an enhanced web user interface, and implementation on small-form processing platforms, allowing for greater mobility. Further testing on a wider range of individuals is also necessary.

REFERENCES

[1] S. Rossner, "Obesity: The disease of the twenty-first century," *Int. J. Obesity*, vol. 26, no. 4, pp. S2–S4, 2002.

[2] F. X. Pi-Sunyer, "The obesity epidemic: Pathophysiology and consequences of obesity," *Obesity Res.*, vol. 10, no. 2, pp. 97S–104S, 2002.

[3] A. E. Field, E. H. Coakley, A. Must, J. L. Spadano, N. Laird, W. H. Dietz, and G. A. Colditz, "Impact of overweight on the risk of developing common chronic diseases during a 10-Year period," *Archives Intern. Med.*, vol. 161, pp. 1581–1586, 2001.

[4] A. Anandacoomarasamy, I. D. Caterson, S. Leibman, G. S. Smith, P. N. Sambrook, M. Fransen, and L. M. March, "Influence of BMI on health-related quality of life: comparison between an obese adult cohort and age-matched population norms," *Obesity*, vol. 17, pp. 2114–2118, 2009.

[5] E. A. Finkelstein, J. G. Trogdon, J. W. Cohen, and W. Dietz, "Annual medical spending attributable to obesity: Payer and service specific estimates," *Health Affairs*, vol. 28, no. 5, pp. w822–w831, Jul. 2009.

[6] A. M. Sharma and R. Padwal, "Obesity is a sign—Over-eating is a symptom: An aetiological framework for the assessment and management of obesity," *Obesity Rev.*, vol. 11, no. 5, pp. 362–370, 2009.

[7] J. A. Levine, M. W. Vander Weg, J. O. Hill, and R. C. Klesges, "Non-Exercise activity thermogenesis," *Arterioscler. Thromb. Vasc. Biol.*, vol. 26, pp. 729–736, 2006.

[8] M. D. Mifflin, S. T. St Jeor, L. A. Hill, B. J. Scott, S. A. Daugherty, and Y. O. Koh, "A new predictive equation for resting energy expenditure in healthy individuals," *Amer. J. Clin. Nutrition*, vol. 51, pp. 241–247, 1990.

[9] D. Frankenfield, L. R- Yousey, and C. Compher, "Comparison of predictive equations for resting metabolic rate in healthy nonobese and obese adults: A systematic review," *J. Amer. Dietetic Assoc.*, vol. 105, no. 5, pp. 775–789, 2005.

[10] D. C. Frankenfield, W. A. Rowe, J. S. Smith, and R. N. Cooney, "Validation of several established equations for resting metabolic rate in obese and nonobese people," *J. Amer. Dietetic Assoc.*, vol. 103, no. 9, pp. 1152–1159, 2003.

[11] NHLBI Obesity Education Initiative. The practical guide: Identification, evaluation, and treatment of overweight and obesity in adults. (2011). [Online]. Available: http://www.nhlbi.nih.gov/guidelines/obesity/prctgd_c.pdf

[12] OMRON. (2011). [Online]. Available: http://www.omronhealthcare.com/product/1132–229-pedometers-hj-720itc

[13] Omron Health Care. (2011). [Online]. Available: http://www.omronhealthcare.com/products/hr-100 c/

[14] A. Teller, "A platform for wearable physiological computing," *Interact. Comput.*, vol. 16, pp. 917–937, 2004.

[15] O. Amft and G. Troster, "On-Body sensing solutions for automatic dietary monitoring," *Pervasive Comput.*, vol. 8, pp. 62–70, Apr./Jun. 2009.

[16] A. G. Ershow, J. O. Hill, and J. T. Baldwin, "Novel engineering approaches to obesity, overweight, and energy balance: Public health needs and research opportunities," in *Proc. IEEE Eng. Med. Biol. Soc.*, 2004, pp. 5212–5214.

[17] O. Amft, M. Stager, P. Lukowicz, and G. Troster, "Analysis of chewing sounds for dietary monitoring," in *Proc. Int. Conf. Ubiquitous Comput.*, 2005, pp. 56–72.

[18] O. Amft, H. Junker, and G. Troster, "Detection of eating and drinking arm gestures using inertial body-worn sensors," in *Proc. IEEE Int. Symp. Wearable Comput.*, 2005, pp. 160–163.

[19] E. Sazonov, S. A. Schuckers, P. Lopez-Meyer, O. Makeyev, E. L. Melanson, M. R. Neuman, and J. O. Hill, "Toward objective monitoring of ingestive behavior in free-living populations," *Obesity*, vol. 17, no. 10, pp. 1971–1975, 2009.

[20] E. Sazonov, S. Schuckers, P. Lopez-Meyer, O. Makeyev, N. Sazonova, E. L. Melanson, and M. Neuman, "Non-Invasive monitoring of chewing and swallowing for objective quantification of ingestive behavior," *Physiol. Meas.*, vol. 29, pp. 525–541, 2008.

[21] O. Amft and G. Troster, "Methods for detection and classification of normal swallowing from muscle activation and sound," in *Proc. Pervasive Health Conf. Workshops*, 2006, pp. 1–10.

[22] O. Amft and G. Troster, "Recognition of dietary activity events using on-body sensors," *Artif. Intell. Med.*, vol. 42, pp. 121–136, 2008.

[23] NT3 Throat Mic for two-way radios, wargamers, paintball, and airsoft enthusiasts. (2011). [Online]. Available: http://www.iasus-concepts.com/nt/nt-3.htm

[24] D. Andre et al. The development of the sensewear® armband, a revolutionary energy assessment device to assess physical activity and lifestyle. (2011). [Online]. Available: http://www.bodymedia.com/site/docs/papers/wp_accuracy_ee.pdf

[25] C. B. Liden *et al*. Characterization and implications of the sensors incorporated into the sensewear™ armband for energy expenditure and activity detection. [Online]. Available: http://www.bodymedia.com/Learn-More/Whitepapers/Characterization-and-Implications-of-the-Sensors-Incorporated-into-the-SenseWear

[26] M. Sharma, "Behavioral interventions for preventing and treating obesity in adults," *Obesity Rev.*, vol. 8, pp. 441–449, 2007.

[27] S. Klein *et al.*, "Clinical implications of obesity with specific focus on cardiovascular disease: A statement for professionals from the american heart association council on nutrition, physical activity, and metabolism," *Circulation*, vol. 110, pp. 2952–2967, 2004.

[28] S. Alinia, O. Hels, and I. Tetens, "The potential association between fruit intake and body weight—A review," *Obesity Rev.*, vol. 10, pp. 639–647, 2009.

[29] M. Moussavi and Z. Aboofazeli, "Comparison of recurrence plot features of swallowing and breath sounds," *Chaos Solitons Fractals*, vol. 37, pp. 454–464, 2008.

[30] M. Moussavi and Z. Aboofazeli, "Swallowing sound detection using hidden markov modeling of recurrence plot features," *Chaos Solitons Fractals*, vol. 39, pp. 778–783, 2009.

[31] M. Moussavi and Z. Aboofazeli, "Analysis of temporal pattern of swallowing mechanism," in *Proc. IEEE Eng. Med. Biol. Soc.*, New York, NY, USA, 2006, pp. 5591–5594.

[32] M. Moussavi and Z. Aboofazeli, "Analysis of swallowing sounds using hidden markov models," *Med. Biol. Eng.*, vol. 46, pp. 307–314, 2008.

[33] M. Aboofazeli and Z. Moussavi, "Automated extraction of swallowing sounds using a wavelet-based filter," in *Proc. 28th IEEE Annu. Int. Conf. Eng. Med. Biol. Soc.*, New York, NY, USA, 2006, pp. 5607–5610.

[34] A. Yadollahi and Z. Moussavi, "A model for normal swallowing sounds generation based on wavelet analysis," in *Proc. Can. Conf. Electr. Comput. Eng.*, Niagara Falls, ON, Canada, 2008, pp. 827–830.

[35] M. Suzuki and K. Nagae, "A neck mounted interface for sensing the swallowing activity based on swallowing sound," in *Proc. IEEE Eng. Med. Biol. Soc.*, 2011, pp. 5224–5227.
[36] C. B. Taylor, N. H. Miller, K. R. Reilly, G. Greenwald, D. Cunning, A. Deeter, and L. Abascal, "Evaluation of a nurse-care management system to improve outcomes in patients with complicated diabetes," *Diabetes Care*, vol. 26, pp. 1058–1063, 2003.
[37] D. F. Tate, R. R. Wing, and R. A. Winett, "Using internet technology to deliver a behavioral weight loss program," *JAMA*, vol. 285, no. 9, pp. 1172–1177, 2001.
[38] D. F. Tate, E. H. Jackvony, and R. R. Wing, "Effects of internet behavioral counseling on weight loss in adults at risk for type 2 diabetes," *JAMA*, vol. 289, no. 14, pp. 1833–1836, 2003.
[39] P. J. Morgan, D. R. Lubans, C. E. Collins, J. M. Warren, and R. Callister, "The SHED-IT randomized controlled trial: Evaluation of an internet-based weight-loss program for men," *Obesity*, vol. 17, no. 11, pp. 2025–2032, 2009.
[40] K. G. Volpp, L. K. John, A. B. Troxel, L. Norton, J. Fassbender, and G. Loewenstein, "Financial incentive based approaches for weight loss: A randomized trial," *J. Amer. Med. Assoc.*, vol. 300, no. 22, pp. 2631–2637, 2008.
[41] J. A. Logemann, "Evaluation and treatment of swallowing disorders," *NSSLHA J.*, vol. 12, pp. 38–50, 1984.
[42] O. Makeyev, E. Sazonov, S. Schuckers, P. Lopez-Meyer, T. Baidyk, E. Melanson, and M. Neuman, "Recognition of swallowing sounds using time-frequency decomposition and limited receptive area neural classifier," *Appl. Innovations Intell. Syst.*, vol. 16, no. 2, pp. 33–46, 2009.
[43] E. S. Sazonov, O. Makeyev, S. Schuckers, P. Lopez-Meyer, E. L. Melanson, and M. R. Neuman, "Automatic detection of swallowing events by acoustical means for applications of monitoring of ingestive behavior," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 3, pp. 626–633, Mar. 2010.
[44] J. M. Fontana, P. L. Melo, and E. S. Sazonov, "Swallowing detection by sonic and subsonic frequencies: a comparison," in *Proc. IEEE Eng. Med. Biol. Soc.*, 2011, pp. 6890–6893.
[45] F. Takens, "Detecting strange attractors in turbulence," in *Dynamical Systems and Turbulence*. vol. 898, New York, NY, USA: Springer, 1981, pp. 366–381.
[46] A. L. P. Aroul, D. Bhatia, and L. Estevez, "Energy-efficient ambulatory activity monitoring for disease management," in *Proc. 5th Int. Workshop Wearable Implantable Body Sens. Netw.*, 2008, pp. 201–204.
[47] A. L. P. Aroul, A. Manohar, D. Bhatia, and L. Estevez, "Power efficient multi-band contextual activity monitoring for assistive environments," in *Proc. 1st Int. Conf. Pervasive Technol. Related Assist. Environ.*, 2008, vol. 282, p. 19.
[48] W. Walker, T. Polk, A. Hande, and D. Bhatia, "Remote blood pressure monitoring using a wireless sensor network," presented at IEEE Sixth Annual Emerging Inf. Technol. Conf., 2006.
[49] W. Walker, A. L. P. Aroul, and D. Bhatia, "Mobile health monitoring systems," in *Proc. IEEE Eng. Med. Biol. Soc.*, 2009, pp. 5199–5202.
[50] J. Tompkins and W. J. Pan, "A real-time QRS detection algorithm," *IEEE Trans. Biomed. Eng.*, vol. BME-32, no. 3, pp. 230–236, Mar. 1985.
[51] S. R. Youmans, Increasing the objectivity of the clinical dysphagia evaluation: Cervical auscultation and tongue function during swallowing. (2003). [Online]. Available http://etd.lib.fsu.edu/theses/available/etd-09232003–010436/unrestricted/02_sry_text.pdf.pdf
[52] M. Brookes, Voicebox: Speech processing toolbox for matlab. [Online]. Available:http://www.ee.ic.ac.uk/hp/staff/dmb/voicebox/voicebox.html
[53] R. R. Wickerhauser and M. V. Coifman, "Entropy-Based algorithms for best basis selection," *IEEE Trans. Inform. Theory*, vol. 38, no. 2, pp. 713–718, Mar. 1992.
[54] C. Taswell, "Near-Best basis selection algorithms with non-additive information cost functions," in *Time-Frequency Time-Scale Anal., Proc. IEEE-SP Int. Symp.*, 1994, pp. 13–16.
[55] Mathworks. (2011). Wavelet Toolbox. [Online]. Available: http://www.mathworks.com/products/wavelet/
[56] M. Vetterli and C. Herley, "Wavelets and filter banks: Theory and design," *IEEE Trans. Signal Process.*, vol. 40, no. 9, pp. 2207–2232, Sep. 1992.
[57] R. Polikar, "Pattern recognition in bioengineering," in *Wiley Encycloedia of Biomedical Engineering*, M. Akay, Ed.. vol. 4New York, NY, USA: Wiley Interscience, 2006, pp. 2695–2716.
[58] MathWorks. (2011). Bioinformatics toolbox documentation. [Online]. Available: http://www.mathworks.com/help/toolbox/bioinfo/
[59] MathWorks. (2011). Statistics toolbox documentation. [Online]. Available: http://www.mathworks.com/help/toolbox/stats/
[60] K. Murphy, Bayes Net Toolbox for MATLAB. (2011). [Online]. Available: http://code.google.com/p/bnt/
[61] R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," in *Proc. int. joint Conf. Artif. Intell.*, 1995, pp. 1137–1145.
[62] N. Saito, "Simultaneous noise suppression and signal compression using a library of orthonormal bases and the minimum description length criterion," in *Wavelets in Geophysics*. New York, NY, USA: Academic, 1994, pp. 299–324.
[63] E. Y. Hamid and Z. -I. Kawasaki, "Wavelet-Based data compression of power system disturbances using the minimum description length criterion," *IEEE Trans. Power Delivery*, vol. 17, no. 2, pp. 460–466, Apr. 2002.
[64] Highcharts. (2011). [Online]. Available: http://www.highcharts.com/
[65] Texas Instruments. OMAP 5 applications processors. (2011). [Online]. Available: http://focus.ti.com/general/docs/wtbu/wtbuproductcontent.tsp?templateId=6123&navigationId=12862&contentId=101230

William P. Walker was born in Dallas, Texas, in 1980. He received the B.S. degree in computer science in 2005, and the M.S. and Ph.D. degrees in computer engineering from the University of Texas at Dallas, TX, USA, in 2007 and 2011, respectively.

From 2005 to 2011, he was a member of the Embedded and Adaptive Computing Group at the University of Texas at Dallas, focusing on health monitoring systems, obesity management, and wireless sensor networks. Since 2011, he has been a member of the Fixed-Point Designer team at Mathworks, focusing on model based design for embedded systems. His research interests include design automation, audio signal processing, and health monitoring systems.

Dinesh Bhatia (SM'09) received the Bachelors degree in electrical engineering and the Masters and Ph.D. degrees in computer science.

He is a faculty member in the Erik Jonsson School of Engineering and Computer Science at the University of Texas at Dallas. He directs research activities within the IDEA Laboratory, Computer Society, Engineering in Medicine and Biology Society, Circuits and Systems Society, Eta Kappa Nu, and has served on the editorial board of IEEE Transactions on COMPUTERs. He was IEEE Circuits and Systems societys distinguished lecturer for 2007 to 2008. He was the General Chair for the International Conference on Body Sensor Networks for 2011.