Recognition of Nutrition Intake Using Time-Frequency Decomposition in a Wearable Necklace Using a Piezoelectric Sensor

Nabil Alshurafa, *Student Member, IEEE*, Haik Kalantarian, *Student Member, IEEE*, Mohammad Pourhomayoun, *Student Member, IEEE*, Jason J. Liu, *Student Member, IEEE*, Shruti Sarin, *Student Member, IEEE*, Behnam Shahbazi, *Student Member, IEEE*, and Majid Sarrafzadeh, *Fellow, IEEE*

***Abstract*— Food intake levels, hydration, ingestion rate, and dietary choices are all factors known to impact the risk of obesity. This paper presents a novel wearable system in the form of a necklace, which aggregates data from an embedded piezoelectric sensor capable of detecting skin motion in the lower trachea during ingestion. The skin motion produces an output voltage with varying frequencies over time. As a result, we propose an algorithm based on time-frequency decomposition, spectrogram analysis of piezoelectric sensor signals, to accurately distinguish between food types, such as liquid and solid, hot and cold drinks, and hard and soft foods. The necklace transmits data to a smartphone, which performs the processing of the signals, classifies the food type, and provides visual feedback to the user to assist the user in monitoring their eating habits over time. We compare our spectrogram analysis with other time-frequency features, such as matching pursuit and wavelets. Experimental results demonstrate promise in using time-frequency features, with high accuracy of distinguishing between food categories using spectrogram analysis and extracting key features representative of the unique swallow patterns of various foods.**

***Index Terms*— Nutrition monitoring, wearable necklace, spectrogram analysis, piezoelectric sensor, machine learning, classification.**

I. MOTIVATION AND BACKGROUND

HEALTHY eating is associated with reduced risk for many diseases, including several of the leading causes of death: heart disease, some cancers, stroke, and diabetes [1]. The development and the incorporation of wireless technologies has the potential to address our ultimate goal of enabling healthier lifestyle choices and behavior modification needed to prevent obesity and obesity-related diseases. Much of the wireless technology developed and used in the market, however, focuses primarily on exercise and physical activity [3], [4], [6], [7], [15], [32]. In this paper, we describe a novel system that attempts to infer eating patterns from a device disguised as a necklace.

Automatically and accurately recognizing the type of food in a non-intrusive manner has been for the most part an unaddressed challenge. Most of the current technologies for eating pattern detection are either inaccurate or exhibit low rates of adherence to using the technology, due to one or more of these shortcomings: 1) they infer eating indirectly from, for example, hand movements or food images [10], [37]; 2) they require manual data entry or user involvement in capturing data [19]; or 3) they are non-wearable, bulky, invasive, or semi-invasive [11]. There is a need for a system that is non-invasive and detects individual's eating patterns, and provides necessary guidance and feedback to the user (see Figure 1). Such a system represents a significant advance in researchers' ability to evaluate the combined impact of adherence to dietary guidelines.

Current systems either have low accuracy in detecting swallows and distinguishing food types, or must be uncomfortably worn around the neck, which renders continuous use impractical [8]. In this paper, we focus on a piezoelectric-based design of a necklace that is not worn tightly around the neck, but rather hangs loosely and falls more naturally right above the sternum [5], [20], [21], [23], [24].

Equally critical to detecting eating episodes is determining whether calories are consumed in solid or liquid form. Studies show reduced ability of the body to compensate properly when calories are consumed in liquid form compared to solid form [30], with the result that some health recommendations now explicitly recommend restrictions on liquid calories consumed (e.g., 2007 report by the World Cancer Research Fund [16]. The proposed system will enable individuals to track the amount of solid vs. liquid consumed throughout the day.

This paper is organized as follows. Related works are described in Section II. In Section III we describe the nutrition monitoring necklace. Then we describe our method of differentiating food types using spectrogram-based feature extraction and classification algorithm in Section IV.

Manuscript received November 27, 2014; accepted January 19, 2015. Date of publication February 11, 2015; date of current version May 19, 2015. The associate editor coordinating the review of this paper and approving it for publication was Prof. Zeljko Ignjatovic.

The authors are with the Department of Computer Science, Wireless Health Institute, University of California at Los Angeles, Los Angeles, CA 90095 USA (e-mail: nabil@cs.ucla.edu; kalantarian@cs.ucla.edu; mpourhoma@cs.ucla.edu; jasonliu@cs.ucla.edu; ssarin@cs.ucla.edu; bshahbazi@cs.ucla.edu; majid@cs.ucla.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSEN.2015.2402652

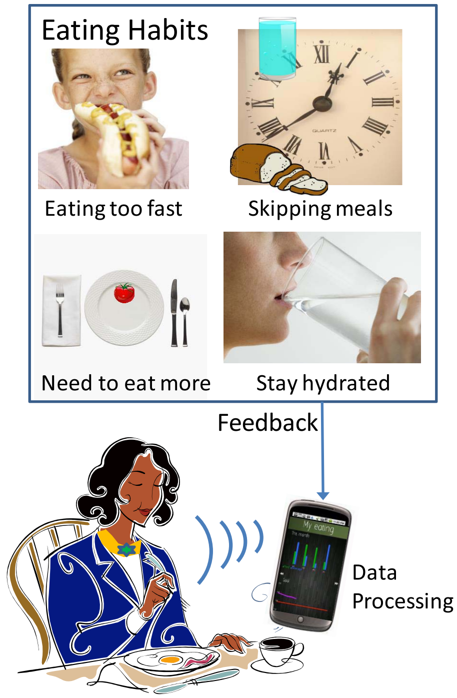

Fig. 1. Monitoring eating habits is essential to promoting healthy lifestyle behavior. In a fast-paced society, some common bad eating habits include eating too fast, skipping meals, not eating enough, and being dehydrated. Having a non-invasive automated method of detecting bad habits can be a means of changing negative habits into positive ones.

We present our experimental setup in Section V, and results in Section VI. We conclude in Section VII.

II. Related Work

Various sensors have been employed to identify the volume of food being consumed, and among the most popular methods is acoustic detection [22]. Several systems have identified chewing and swallowing acoustically by placing a microphone near the throat, and using signal processing techniques for classification. For example, Sazanov et al. [35] uses acoustic data acquired from a small microphone placed near the bottom of the throat. However, their system is coupled with a strain gauge placed near the ear which is not practical for daily use. Similarly, Nagae et al. [33] attempts to distinguish between swallowing, coughing, and vocalization using wavelet-transform analysis of audio data. Though results are promising, this technology is targeted towards those who suffer from dysphagia, and identifying the volume or characteristic of food intake is not the focus of their work.

Aboofazeli et al. [2] present another approach to acoustic swallow detection, achieving basic classification between swallows and breath sounds using a feedforward neural network classifier. A manual inspection of their classification results is performed using a spectrogram, which is a basis for the feature extraction technique for food classification used in our work. Makayev et al. apply spectrograms for swallow detection using machine learning algorithms [27], though once again, no classification is performed, and their analysis is limited to identifying swallows. Ultimately, acoustic detection of food intake is promising, but suffers from several serious drawbacks including the interference of background noise, a lack of uniformity between individual eating styles, and no prior work validating the feasibility of classification between different types of food.

Several other methods for detecting swallows have been explored. Amft et al. [9] performs detection of eating and drinking by identifying associated arm gestures using accelerometers and gyroscopes. For example, the use of cutlery, spoon, hand, and cup can be identified based on the gestures associated with food intake using these objects. However, the eating style does not necessarily reveal the volume of food intake, which severely limits the usefulness of this approach. Other works place electrodes on the neck and perform an EMG to identify deglutition, but the hardware is cumbersome and the system is limited to a clinical environment [26].

Piezoelectric sensors, which produce an output voltage corresponding with the mechanical stress applied to the body of the sensor, are used in countless applications. Recently, they have been applied to problems in the medical domain, such as identifying individual heart beats and respiration [25]. Very few works describe attempts to use piezoelectric sensors for monitoring food ingestion, with several exceptions [13], though evaluation of dysphagia symptoms is the primary objective of their work.

In this paper, we compare the accuracy of classification techniques from a piezoelectric sensor worn around the neck using statistical features extracted in the time domain to a novel spectrogram-based approach which considers time and frequency-based components in tandem. A spectrogram, often used for speech recognition and other countless applications, is a visual representation of the frequency spectrum over time generated using a short-time Fourier transform (STFT) with a fixed window size, the squared magnitude of which yields the spectrogram. Spectrograms are used to visually represent changes in the frequency spectrum over time, have been applied to countless research problems pertaining to the analysis of acoustic signals. Examples include speech recognition, the identification of animal sounds such as whale vocalizations, and pattern recognition in genome sequences [12], [31], [36]. However, their utility in analyzing piezoelectric sensor data has not been adequately explored. The primary novelty of our work is the application of spectrograms for analysis of piezoelectric sensor data in the realm of detection and classification of food ingestion. In this paper we collect statistical features on a spectrogram of swallows to distinguish between solid and liquid foods.

III. Nutrition Monitoring Necklace Design

Our nutrition monitoring system comprises two main components: piezoelectric-based sensor technology, and a smartphone application that performs data processing, user guidance, and feedback. The smartphone application performs swallow detection, feature extraction and classification to detect swallows. This section describes the sensor technology and user guidance and feedback (See Figure 5). Section IV will further discuss the classification algorithms implemented on the smartphone. Figure 2 provides a component overview of the system.

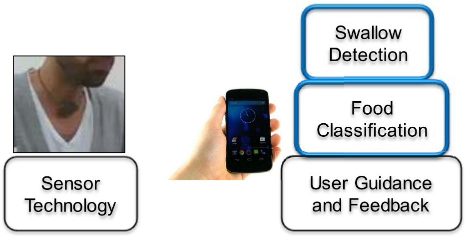

Fig. 2. The essential components of the system. Besides the sensor technology, the main challenges in designing the system involve swallow detection and food classification. User guidance and feedback is also part of the system and has the potential in being of great benefit to the user regarding improving health.

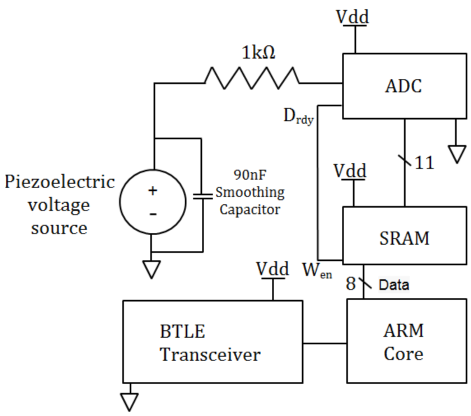

Fig. 3. Circuitry diagram for the system.

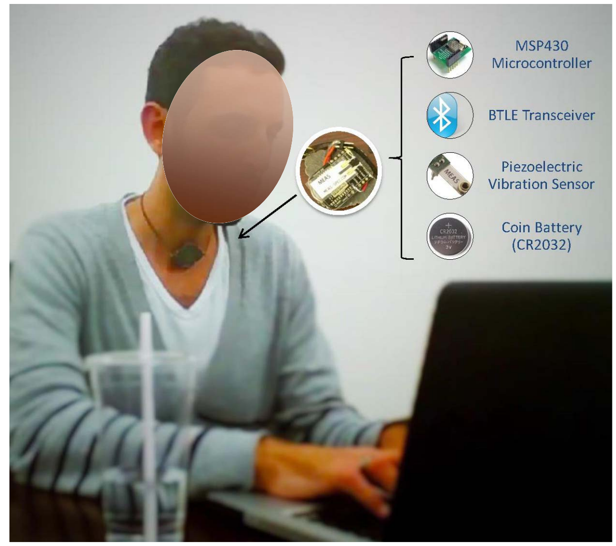

Fig. 4. This figure shows a subject wearing the necklace while working. The necklace comprises the piezoelectric sensor, coin battery, MCU/RF board and a fashionable cover, where the sensor is designed to be in contact with the skin.

A. Sensor Technology

A piezoelectric sensor, also known as a vibration sensor, produces a voltage when subjected to physical strain. By placing a piezoelectric sensor against the throat, the muscle contraction and motion of the skin during a swallow is represented in the output voltage of the sensor, when sampled at frequencies as low as 5 Hz. Our necklace features a thin, lightweight piezoelectric vibration sensor attached to the inside of the necklace, along with a small microcontroller board capable of sampling the sensor and transmitting the data to a mobile phone via Bluetooth. The hardware is powered by a lightweight lithium-polymer battery. Figure 4 depicts a subject wearing the necklace, and further illustrates each component.

Figure 3 models the piezoelectric sensor as a nonlinear voltage source when subject to mechanical excitation. The data is smoothed using a 90nF filtering capacitor, and a small resistor brings voltage levels to the valid input range of the Analog/Digital converter unit (ADC). After sampling is complete, the data is buffered into SRAM memory, processed, and transmitted. The on-board ADC has a resolution of 10 bits and can convert data at a rate of up to 257 kHz. The offset error, gain error, and absolute error ratings are 1, 3, and 3 LSB volts, respectively. The resolution of the ADC was therefore approximately 15 mV, based on the supported input voltage range. This was sufficient for the purposes of a nutrition monitoring application, as swallows were typically associated with a voltage spike of 50 mV or more. The hardware platform supports an input voltage range of 1.8-3.6 volts. The lithium-polymer battery used to power the device is a 3.7 volt unregulated voltage source with a capacity of 170 lmAh and a maximum discharge current of 1 Ampere at room temperature.

The necklace is available in several varieties including a sportsband suitable for athletes and other active individuals, and another targeted towards a more fashion-conscious audience. Because the hardware components of the necklace are very small and lightweight, they can be embedded in several different form factors.

The microcontroller board samples the voltage of the vibration sensor at a rate of 20Hz, converting the voltage to a digital signal using the on-chip A/D converter. The data is then buffered and transmitted to a mobile phone. This Arduino-compatible board features a Bluetooth 4.0 LE transceiver on-board, based on the RFD22301 SMT module. The embedded processor is an ARM Cortex M0 with 256kB of flash memory and 16kB of RAM.

B. Device Battery Lifetime

The device battery lifetime depends on many parameters including sample rate, battery capacity, Bluetooth connection interval, and various algorithm parameters such as window size and sample rate. A CR2032 coin-cell battery typically has a capacity of approximately 235mAh. Our experimental simulations reveal that a window size of 10 and a sample rate of 20Hz results in a power consumption of .07 mW, using the Nordic nRF simulation software and a low-power MSP430 microcontroller. This would correspond with a device lifetime of over 6 months. However, the hardware platform used in this paper is not optimized for energy-efficient applications, as the focus is aggregating data for offline processing to evaluate our algorithms.

C. User Guidance and Feedback

This system includes a mobile phone application for data reporting and visualization (see Figure 5). The application displays the estimated liquid and solid volume of the current

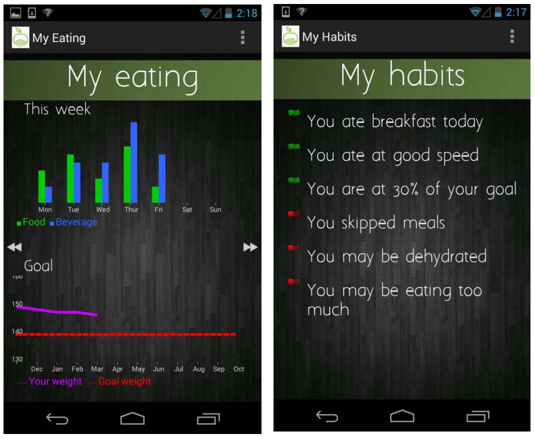

Fig. 5. Mobile application screens. The snapshot on the left shows a summary of the users weekly food and beverage consumption. The snapshot to the right provides a summary of habits for the user, some positive and others negative.

meal, as well as the daily and monthly total. A reporting tool displays alerts to the user.

The mobile application uses the Bluetooth 4.0 LE protocol to receive data from the necklace while maximizing battery life. The data is then processed for swallow identification, classification, and analytics. To ensure user compliance, the application is able to detect when the necklace has been removed, based on the observation that the average individual swallows saliva periodically. Lastly, all collected and processed data is uploaded to a secured cloud server for patient tracking and statistical analysis.

IV. Algorithm

A. Swallow Detection

Figure 6 provides a summary of the algorithm implemented on the mobile phone, which is used to detect swallows based on data acquired from the vibration sensor and received via Bluetooth. The data is buffered locally until a sufficient number of samples have been acquired. Subsequently, a sliding window is applied to generate a new waveform representing the standard deviation of the original data. The swallows are represented in the resulting waveform as peaks, while they may correspond to either peaks or troughs in the original data.

The algorithm then proceeds to smooth the waveform by applying a Savitzky-Golay convolution filter to increase the signal-to-noise ratio without distorting the signal, which yields clearly visible peaks representative of each original swallow, while removing noise from the signal [34]. Subsequently, the number of swallows can be identified by counting the number of peaks, provided there is sufficient spacing between swallows.

B. Spectrogram

Once a swallow is detected, a spectrogram is generated centered around each swallow. The spectrogram is calculated from the time signal $x(t)$, as shown in Equation 1 using the short-time Fourier transform (STFT).

$$STFT\{x(t)\} \equiv X(n, \omega) = \sum_{t=-\infty}^{\infty} x[t]\omega[t-n]e^{-j\omega t}. \quad (1)$$

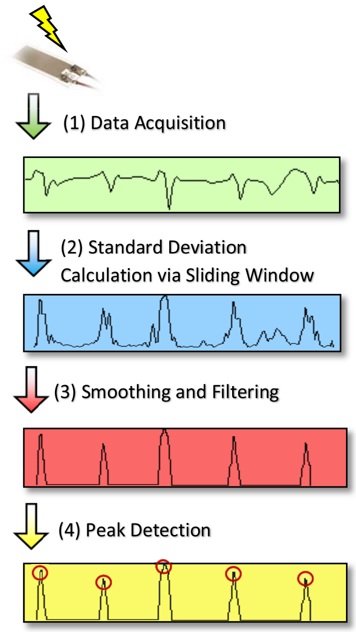

Fig. 6. This figure shows the swallow detection process, whereby a sliding window is applied to the signal to generate a waveform representing the standard deviation of the signal, after smoothing, peaks are detected and identified as potential swallow regions.

$x(t)$ is multiplied by a window function for a short period of time. The data is divided into frames F_i, which overlap. Each frame is Fourier transformed, and the result is added to a matrix that records the magnitude and phase of each point in time and frequency.

The spectrogram is the resulting 3-dimensional plot of the energy of the frequency content of a signal as it changes over time [14]. For our window function, we used different values; for the Hamming window we tried lengths of $w = 32, 64$, and 128, with an FFT length of $nfft = 32, 64$, and 128, and an of overlap of 25%, 50%, 75%, and no overlap. We set the dynamics range to 50dB. Figure 7 provides an illustration showing a sample swallow spectrogram for three food types (water, chips, and sandwich). Each spectrogram is defined by a matrix $P \in R^{m \times k}$, where m is the number of bins in the time domain, and k is the number of bins in the frequency domain. P represents the power spectral density.

The distinguishing attributes of these piezoelectric signals are visible. For example, chips and sandwich swallows contain more high frequency components than water swallows due to the effect of chewing. Distinguishing between chips and sandwich swallows, though significantly more challenging, is captured by our statistical feature extraction methodology.

C. Feature Extraction

Once a spectrogram for each swallow is generated, we found an optimal division of the spectrogram images into 14 bins along the frequency domain and another 16 bins along the time domain, for a total of 30 bins. We then calculate statistical features on each bin, to generate a feature vector V_i for each swallow. Table I lists the main features that were calculated for each bin, which generates a total of $s = 360$ features per

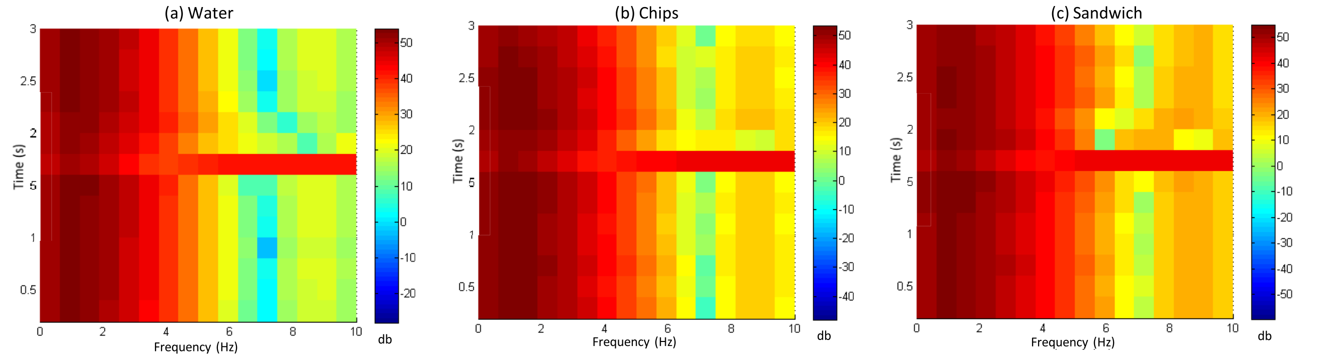

Fig. 7. (a) Shows the spectrogram over a 3 second window of water. (b) and (c) Shows the spectrogram for chips and sandwich. The distinction between water and solid is captured in the spectrogram. However, the difference between chips and sandwich is more challenging, and the statistical features collected from the spectrogram images is capable of learning the difference.

TABLE I
FEATURE TABLE

Mean	Geometric Mean	Std. Dev.
Skewness	Mean of Standardized Z-Scores	IQR
Kurtosis	Harmonic Mean	Rank Corr.
Range	Median Absolute Deviation	Partial Corr.

spectrogram swallow. We generate a matrix B which contains all n samples for each spectrogram P_i.

$$B = [V_1, V_2, ..., V_i, ..., V_n] \\ = \begin{bmatrix} \upsilon_{11} & \cdots & \upsilon_{1s} \\ \upsilon_{21} & \cdots & \upsilon_{2k} \\ \vdots & \ddots & \vdots \\ \upsilon_{n1} & \cdots & \upsilon_{ns} \end{bmatrix} \tag{2}$$

where $\upsilon_{i,j}$ represents the j^th feature of the i^th sample.

D. Feature Selection and Classification

The conventional feature selection algorithms usually focus on specific metrics to quantify the relevance and/or redundancy to find the smallest subset of features that provides the maximum amount of useful information for prediction. Thus, the main goal of feature selection algorithms is to eliminate redundant or irrelevant features in a given feature set. Applying an effective feature selection algorithm not only decreases the computational complexity of the system by reducing the dimensionality and eliminating the redundancy, but also increases the performance of the classifier by deleting irrelevant and confusing information.

The two well-known feature selection categories are the filter and wrapper methods. Filter methods use a specific metric to score each individual feature (or a subset of features together), and are usually fast and much less computationally intensive. Wrapper methods usually utilize a classifier to evaluate feature subsets in an iterative manner according to their predictive power [17]. We applied the wrapper method, testing on multiple combinations of feature subsets and classifiers including: kNN, Bayesian Network, Random Forest. We reduce the dimensionality of the features from $s = 360$ to l, where l depends on the feature selection algorithm used, however we limit it to $l = 30$. The optimal feature extraction, feature selection and classification combination is selected to run in real time. We then divide each data sample into training and testing samples. Figure 8 provides an illustration of the system architecture, where an optimal feature subset and classifier is trained to distinguish between food types.

V. EXPERIMENTAL SETUP

Two experiments were performed to validate the efficacy of our algorithm in accurately detecting swallows and recognizing eating patterns. The first experiment involved 10 subjects, and the second experiment included an additional 10 subjects (for a total of 20 subjects). The two experiments are described in this section.

Once a swallow is detected, we test the ability of our system to classify swallows using features collected from three time-frequency based algorithms including: Matching Pursuit with dictionaries of Gabor functions [29], Morlet Wavelet (as it is closely related to human perception, both hearing and vision) with statistical features collected from its corresponding scaleogram [28], and statistical features collected from a spectrogram.

To prevent bias in the classification results between each class label in the training set, we randomly select an equal number of swallows across categories. We also perform 10-fold cross validation and report the results. We test each classifier's ability to distinguish between different food types.

A. Experiment 1

In the first experiment data was collected on ten subjects, two female and eight male with ages ranging between 20 and 40 years of age. We placed the necklace around their neck so that the sensor was loosely touching the skin. The necklace tightness was adjusted such that each subject was comfortable wearing the device. We placed the necklace centered between their right and left clavicle right above the sternum.

Each subject consumed two types of food: a tuna, egg, or chicken sandwich on white bread, and a few pieces of Pringles potato chips. Each subject selected which food type to start eating or drinking first. The subjects were asked to consume an 8 oz glass of water at room temperature. We ensured that the portion sizes were identical from one subject to another. The subjects were then asked to push a button every time they swallowed; this helped us further annotate the data in order to provide truth labels for the dataset.

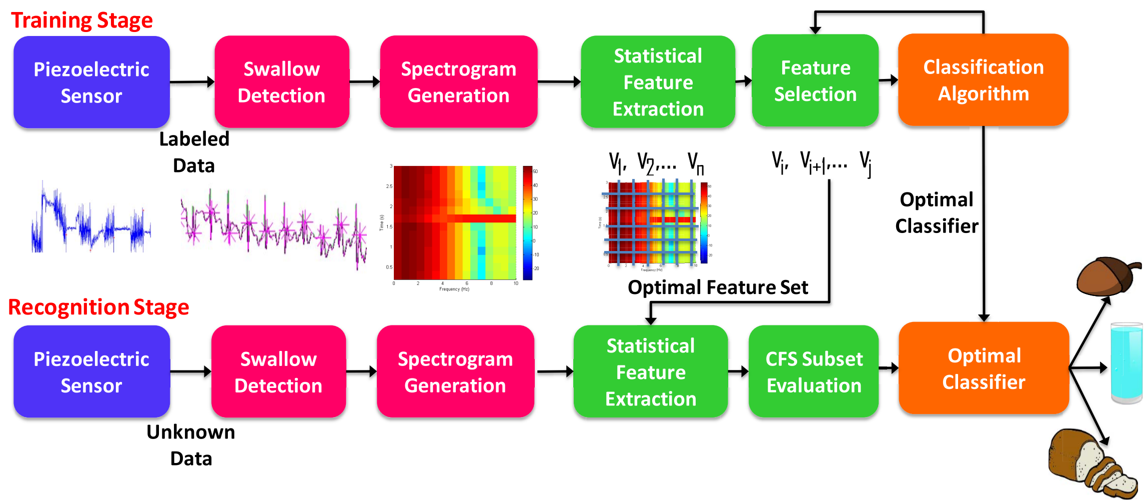

Fig. 8. System architecture.

B. Experiment 2

In the second experiment we increased the number of subjects to twenty, eight female and twelve male, ages 20 to 40 years. We also added hot tea as another food type of liquid form. This enabled us to distinguish between hot and cold drinks. The subjects each consumed 8 ounces of room temperature water and 8 ounces of hot tea. They also consumed two fun-sized snicker bars (Chocolate), a meat-like veggie patty (Patty), and a handful of mixed nuts (Nuts).

In this experiment we compare the classifiers' ability to distinguish between liquid and solid as well as different textures, temperatures, and consistencies.

VI. Results and Discussion

The feature selection algorithm that consistently performed best in combination with the classifiers was the Correlation-based feature subset selection algorithm [18]. This method evaluates the worth of a subset of attributes by considering the individual predictive ability of each feature as well as the redundancy between them.

A. Experiment 1

According to our classification results, using spectrogram-based features on a signal from a piezoelectric sensor can distinguish between liquid and solid swallows with higher accuracy than Matching Pursuit and Wavelets (See Table II) using the Random Forest Classifier (with n=100 trees), which yielded the optimal results for all three experiments. Best results were achieved using a window size of 32, an FFT length of 32, and an overlap of 50%. Spectrogram-based features consistently outperformed other methods and for this reason we focus on spectrogram-based feature results.

In this experiment the Liquid class label is represented by water, and the Solid class label is represented equally by chips and sandwich. Using spectrogram analysis, the classifiers that yielded the best results were using Bayesian Networks and Random Forest Classifier (with number trees set to 100). Using kNN, Bayesian Networks, and Random Forests we achieved weighted average F-measures (the harmonic mean of precision and recall) of 86.1%, 83.6%, and 91.2% respectively. We provide the Random Forest (n=100 trees) confusion matrix in Table III. The precision for Liquids is 91.7% which is higher than the precision for Solids, which is 90.1%, but the recall for Liquids is less than that of Solids.

TABLE II
Time-Frequency Decomposition Results

	Precision	Recall	F-Measure	AUC
Matching Pursuit	75.9	77.6	76.7%	86.5%
Wavelet	83.9	85.9	84.9%	94.3%
Spectrogram	91.2%	91.2%	91.2%	95.5%

TABLE III
Confusion Matrix for Experiment 1 Liquids Versus Solids Under Random Forest

	Predicted Outcome		
Swallow Type	Liquid	Solid	Recall
Liquid	75	10	88.2%
Solid	5	80	94.1%
Precision	93.8%	88.8%	

We further analyze the classification algorithms ability to distinguish between the two solid food types chips and sandwich. The results still favor the Random Forest Classifier with an F-measure of 75.9%, outperforming kNN and Bayesian Networks by more than 5%. Table IV provides the confusion matrix for the Random Forest Classifier, the majority of the misclassification occurs between the Chips and Sandwich class labels. The F-measure is approximately 76.6%. As can be seen from the results, distinguishing between solids is quite challenging, due to the large variation in swallow and chew behavior across subjects. To estimate the effects of this test run on distinguishing between solid and liquid food types, Table IV provides the resulting Liquid/Solid precision and recall values, which result in a 93.0% Solid precision and recall.

While it may be challenging to distinguish a single swallow of chips and sandwich, we wanted to find out if the

TABLE IV
CONFUSION MATRIX FOR EXPERIMENT 1 DISTINGUISHING EACH CATEGORY UNDER THE RANDOM FOREST CLASSIFIER

	Predicted Outcome				
Swallow Type	Water	Sandwich	Chips	Recall	Liq/Sld Recall
Water	43	3	4	86.0%	86.0%
Sandwich	2	39	9	78.0%	93.0%
Chips	5	12	33	66%	
Precision	86.0%	72.2%	71.7%		
Liq/Sld Precision	86.0%	93.0%			

TABLE V
CONFUSION MATRIX FOR EXPERIMENT 1 DISTINGUISHING EACH CATEGORY UNDER THE RANDOM FOREST CLASSIFIER

	Predicted Outcome				
Swallow Type	Water	Sandwich	Chips	Recall	Liq/Sld Recall
Water	10	0	0	100%	100%
Sandwich	1	9	0	90.0%	95.0%
Chips	0	2	8	80%	
Precision	90.9%	81.8%	100%		
Liq/Sld Precision	90.9%	100%			

classification algorithm could distinguish between solids given a fixed period in time. We calculated the spectrogram centered around a swallow with a 20 second time interval. The optimal results were provided using a window size of 128, FFT length of 64, and an overlap of 50%. The results show great improvement over single swallows, as shown in Table V, where the Liquid and Solid precision is now set to 90.9% and 100%, respectively, and the Liquid and Solid recall is 100%, and 95%, respectively.

B. Experiment 2

In the second experiment the classifier that provided the best results was also the Random Forest Classifier with number of trees set to 100 (we tested n=10, n=50 and n=100). The resulting confusion matrix is provided in Table VI. The results again show the ability of the algorithm to accurately distinguish between liquid and solids.

The Liquid class label precision and recall of 87.0% and 86.3%, respectively. The Solid class label produced a high precision and recall of 86.4% and 87.1%, which further affirm our findings that the collected spectrogram features are good discriminants between liquids and solids. The Random Forest Classifier resulted in a 86.6% F-measure compared to 77% and 78.1% F-measures of Bayesian Networks and kNN (with k=3 yielding the best results), respectively.

When testing the ability of the algorithm to distinguish between hot tea and water, our results show that the Random Forest Classifier resulted in a 90% F-measure, with a liquid precision and recall of 88.5% and 92.0% respectively. The solid precision and recall is also high at 91.7% and 88.0%. Table VII provides the confusion matrix between the Hot Tea and Water class label. The Random Forest Classifier resulted in a 90.0% F-measure compared to 69.0% and 84.0% F-measures of Bayesian Networks and kNN (with k=3 yielding the best results), respectively.

TABLE VI
CONFUSION MATRIX FOR EXPERIMENT 2 LIQUIDS VERSUS SOLIDS UNDER RANDOM FOREST CLASSIFIER

	Predicted Outcome		
Swallow Type	Liquid	Solid	Recall
Liquid	207	33	86.3%
Solid	31	209	87.1%
Precision	87.0%	86.4%	

TABLE VII
CONFUSION MATRIX FOR EXPERIMENT 2 HOT TEA WATER UNDER RANDOM FOREST CLASSIFIER

	Predicted Outcome		
Swallow Type	Hot Tea	Water	Recall
Hot Tea	92	8	92.0%
Water	12	88	88.0%
Precision	88.5%	91.7%	

TABLE VIII
CONFUSION MATRIX FOR EXPERIMENT 2 DISTINGUISHING SOLIDS UNDER RANDOM FOREST CLASSIFIER

	Predicted Outcome			
Swallow Type	Nuts	Chocolate	Patty	Recall
Nuts	36	10	4	72.0%
Chocolate	5	41	4	82.0%
Patty	2	5	43	86.0%
Precision	83.7%	73.2%	84.3%	

It's interesting to note the challenge of distinguishing between solids. While there are an infinite number of food types, people often maintain a regular regimen. Such a system can become customized based on each subjects diet. As seen in Table VIII the Random Forest Classifier consistently outperforms other well-known classifiers, even when distinguishing between solids, achieving an F-measure of 80%. The Bayesian Network and kNN classifier resulted in a 72.6% and 65.4% F-measure, respectively. Table VIII provides the confusion matrix for the Random Forest Classifier.

VII. CONCLUSION

In this paper we performed classification of swallows using statistical features collected from spectrograms generated from piezoelectric sensor signals. Our results show promise in using spectrogram analysis in combination with piezoelectric sensors as opposed to audio sensors. We have developed and tested a necklace prototype which has shown the ability to successfully distinguish between liquids and solids in two experiments using Random Forest Classifier with 100 trees resulting in an F-measure above 90%. We show a system and framework capable of distinguishing between hot and cold drinks with an F-measure of 90%. We also show potential for distinguishing between solid food types with an F-measure of about 80%. Our future work intends to expand classification to different types of foods, and test in more natural living environments.

REFERENCES

[1] U.S. Department of Agriculture and U.S. Department of Health and Human Services, *Dietary Guidelines for Americans*, 7th ed. Washington, DC, USA: U.S. Government Printing Office, Dec. 2010.

[2] M. Aboofazeli and Z. Moussavi, "Automated classification of swallowing and breadth sounds," in *Proc. 26th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (IEMBS)*, vol. 2, Sep. 2004, pp. 3816–3819.

[3] N. Alshurafa, J. Eastwood, S. Nyamathi, W. Xu, J. J. Liu, and M. Sarrafzadeh, "Battery optimization in smartphones for remote health monitoring systems to enhance user adherence," in *Proc. 7th Int. Conf. Pervasive Technol. Rel. Assistive Environ.*, 2014, Art. ID 8.

[4] N. Alshurafa *et al.*, "Anti-cheating: Detecting self-inflicted and impersonator cheaters for remote health monitoring systems with wearable sensors," in *Proc. 11th Int. Conf. Wearable Implant. Body Sensor Netw. (BSN)*, Jun. 2014, pp. 92–97.

[5] N. Alshurafa, H. Kalantarian, M. Pourhomayoun, S. Sarin, J. Liu, and M. Sarrafzadeh, "Non-invasive monitoring of eating behavior using spectrogram analysis in a wearable necklace," in *Proc. IEEE EMBS Healthcare Innov. Point Care Technol. (HIPT)*, Oct. 2014, pp. 71–74.

[6] N. Alshurafa *et al.*, "Designing a robust activity recognition framework for health and exergaming using wearable sensors," *IEEE J. Biomed. Health Informat.*, vol. 18, no. 5, pp. 1636–1646, Sep. 2014.

[7] N. Alshurafa *et al.*, "Robust human intensity-varying activity recognition using stochastic approximation in wearable sensors," in *Proc. IEEE Int. Conf. Body Sensor Netw. (BSN)*, May 2013, pp. 1–6.

[8] O. Amft and G. Tröster, "Methods for detection and classification of normal swallowing from muscle activation and sound," in *Proc. 1st Int. Conf. Pervasive Comput. Technol. Healthcare (ICST)*, Nov. 2006, pp. 1–10.

[9] O. Amft, H. Junker, and G. Troster, "Detection of eating and drinking arm gestures using inertial body-worn sensors," in *Proc. 9th IEEE Int. Symp. Wearable Comput.*, Oct. 2005, pp. 160–163.

[10] Y. Dong, A. Hoover, and E. Muth, "A device for detecting and counting bites of food taken by a person during eating," in *Proc. IEEE Int. Conf. Bioinformat. Biomed. (BIBM)*, Nov. 2009, pp. 265–268.

[11] F. G. D'Ottaviano, T. A. Linhares Filho, H. M. T. de Andrade, P. C. L. Alves, and M. S. G. Rocha, "Fiberoptic endoscopy evaluation of swallowing in patients with amyotrophic lateral sclerosis," *Brazilian J. Otorhinolaryngol.*, vol. 79, no. 3, pp. 349–353, May/Jun. 2013.

[12] P. Dugan, M. Pourhomayoun, Y. Shiu, R. Paradis, A. Rice, and C. Clark, "Using high performance computing to explore large complex bioacoustic soundscapes: Case study for right whale acoustics," *Proc. Comput. Sci.*, vol. 20, pp. 156–162, 2013.

[13] C. Ertekin, N. Yüceyar, and I. Aydoğdu, "Clinical and electrophysiological evaluation of dysphagia in myasthenia gravis," *J. Neurol., Neurosurgery, Psychiatry*, vol. 65, no. 6, pp. 848–856, 1998.

[14] J. Flanagan, *Speech Analysis Synthesis and Perception* (Kommunikation und Kybernetik in Einzeldarstellungen). New York, NY, USA: Springer, 1972.

[15] F. Fraternali, M. Rofouei, N. Alshurafa, H. Ghasemzadeh, L. Benini, and M. Sarrafzadeh, "Opportunistic hierarchical classification for power optimization in wearable movement monitoring systems," in *Proc. 7th IEEE Int. Symp. Ind. Embedded Syst. (SIES)*, Jun. 2012, pp. 102–111.

[16] World Cancer Research Fund / American Institute for Cancer Research, *Policy and Action for Cancer Prevention. Food, Nutrition and Physical Activity: A Global Perspective*. Washington, DC, USA: AICR, 2007.

[17] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *J. Mach. Learn. Res.*, vol. 3, pp. 1157–1182, Mar. 2003.

[18] M. A. Hall, "Correlation-based feature subset selection for machine learning," Ph.D. dissertation, Dept. Comput. Sci., Univ. Waikato, Hamilton, New Zealand, 1998.

[19] M. Jain, G. R. Howe, and T. Rohan, "Dietary assessment in epidemiology: Comparison of a food frequency and a diet history questionnaire with a 7-day food record," *Amer. J. Epidemiol.*, vol. 143, no. 9, pp. 953–960, 1996.

[20] H. Kalantarian, N. Alshurafa, T. Le, and M. Sarrafzadeh, "Monitoring eating habits using a piezoelectric sensor-based necklace," *Comput. Biol. Med.*, vol. 58, pp. 46–55, Mar. 2015.

[21] H. Kalantarian, N. Alshurafa, T. Le, and M. Sarrafzadeh, "Non-invasive detection of medication adherence using a digital smart necklace," in *Proc. IEEE PerCom, Smart Environ. Workshop*, Mar. 2015.

[22] H. Kalantarian, N. Alshurafa, M. Pourhomayoun, S. Sarin, T. Le, and M. Sarrafzadeh, "Spectrogram-based audio classification of nutrition intake," in *Proc. IEEE EMBS Healthcare Innov. Point Care Technol. (HIPT)*, Oct. 2014, pp. 161–164.

[23] H. Kalantarian, N. Alshurafa, M. Pourhomayoun, and M. Sarrafzadeh, "Power optimization for wearable devices," in *Proc. IEEE PerCom, WristSense*, 2015.

[24] H. Kalantarian, N. Alshurafa, and M. Sarrafzadeh, "A wearable nutrition monitoring system," in *Proc. IEEE 11th Int. Conf. Wearable Implant. Body Sensor Netw.*, Jun. 2014, pp. 75–80.

[25] T. Klap and Z. Shinar, "Using piezoelectric sensor for continuous-contact-free monitoring of heart and respiration rates in real-life hospital settings," in *Proc. Comput. Cardiol. Conf. (CinC)*, Sep. 2013, pp. 671–674.

[26] A. K. Limdi, M. J. McCutcheon, E. Taub, W. E. Whitehead, and E. W. Cook, III, "Design of a microcontroller-based device for deglutition detection and biofeedback," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc., Images 21st Century*, vol. 5, Nov. 1989, pp. 1393–1394.

[27] O. Makeyev *et al.*, "Recognition of swallowing sounds using time-frequency decomposition and limited receptive area neural classifier," in *Applications and Innovations in Intelligent Systems XVI*, T. Allen, R. Ellis, and M. Petridis, Eds. London, U.K.: Springer-Verlag, 2009, pp. 33–46.

[28] S. Mallat, *A Wavelet Tour of Signal Processing: The Sparse Way*, 3rd ed. New York, NY, USA: Academic, 2008.

[29] S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Process.*, vol. 41, no. 12, pp. 3397–3415, Dec. 1993.

[30] R. D. Mattes and W. W. Campbell, "Effects of food form and timing of ingestion on appetite and energy intake in lean young adults and in young adults with obesity," *J. Amer. Dietetic Assoc.*, vol. 109, no. 3, pp. 430–437, 2009.

[31] D. K. Mellinger and C. W. Clark, "Recognizing transient low-frequency whale sounds by spectrogram correlation," *J. Acoust. Soc. Amer.*, vol. 107, no. 6, p. 3518, 2000.

[32] B. J. Mortazavi, M. Pourhomayoun, G. Alsheikh, N. Alshurafa, S. I. Lee, and M. Sarrafzadeh, "Determining the single best axis for exercise repetition recognition and counting on SmartWatches," in *Proc. 11th Int. Conf. Wearable Implant. Body Sensor Netw. (BSN)*, Jun. 2014, pp. 33–38.

[33] M. Nagae and K. Suzuki, "A neck mounted interface for sensing the swallowing activity based on swallowing sound," in *Proc. Annu. Int. Conf. Eng. Med. Biol. Soc. (EMBC)*, Aug./Sep. 2011, pp. 5224–5227.

[34] A. Savitzky and M. J. E. Golay, "Smoothing and differentiation of data by simplified least squares procedures," *Anal. Chem.*, vol. 36, no. 8, pp. 1627–1639, 1964.

[35] E. S. Sazonov and J. M. Fontana, "A sensor system for automatic detection of food intake through non-invasive monitoring of chewing," *IEEE Sensors J.*, vol. 12, no. 5, pp. 1340–1348, May 2012.

[36] D. Sussillo, A. Kundaje, and D. Anastassiou, "Spectrogram analysis of genomes," *EURASIP J. Adv. Signal Process.*, vol. 2004, pp. 29–42, Jan. 2004.

[37] N. Yao, R. J. Sclabassi, Q. Liu, and M. Sun, "A video-based algorithm for food intake estimation in the study of obesity," in *Proc. IEEE 33rd Annu. Northeast Bioeng. Conf. (NEBC)*, Mar. 2007, pp. 298–299.

Nabil Alshurafa, photograph and biography not available at the time of publication.

Haik Kalantarian, photograph and biography not available at the time of publication.

Mohammad Pourhomayoun, photograph and biography not available at the time of publication.

Jason J. Liu, photograph and biography not available at the time of publication.

Shruti Sarin, photograph and biography not available at the time of publication.

Behnam Shahbazi, photograph and biography not available at the time of publication.

Majid Sarrafzadeh, photograph and biography not available at the time of publication.