



Published in final edited form as:

Physiol Meas. 2014 May ; 35(5): 739–751. doi:10.1088/0967-3334/35/5/739.

A novel approach for food intake detection using electroglottography

Muhammad Farooq, Juan M Fontana, and Edward Sazonov

Department of Electrical and Computer Engineering, University of Alabama, Tuscaloosa, AL 35487 USA

Edward Sazonov: esazonov@eng.ua.edu

Abstract

Many methods for monitoring diet and food intake rely on subjects self-reporting their daily intake. These methods are subjective, potentially inaccurate and need to be replaced by more accurate and objective methods. This paper presents a novel approach that uses an Electroglottograph (EGG) device for an objective and automatic detection of food intake. Thirty subjects participated in a 4-visit experiment involving the consumption of meals with self-selected content. Variations in the electrical impedance across the larynx caused by the passage of food during swallowing were captured by the EGG device. To compare performance of the proposed method with a well-established acoustical method, a throat microphone was used for monitoring swallowing sounds. Both signals were segmented into non-overlapping epochs of 30 s and processed to extract wavelet features. Subject-independent classifiers were trained using Artificial Neural Networks, to identify periods of food intake from the wavelet features. Results from leave-one-out cross-validation showed an average per-epoch classification accuracy of 90.1% for the EGG-based method and 83.1% for the acoustic-based method, demonstrating the feasibility of using an EGG for food intake detection.

Keywords

Eating disorders; Ingestive behavior; Dietary intake monitoring; Electroglottography sensor; Artificial Neural Network; Swallowing sound

1. Introduction

An accurate and objective monitoring of ingestive behavior is particularly important for research in populations suffering from eating disorders and obesity. An eating disorder is a medical condition that causes a serious disruption in a person's diet. People suffering from eating disorders have abnormal eating behaviors due to the consumption of either insufficient or excessive amounts of food. Common eating disorders include anorexia nervosa, bulimia nervosa and binge-eating disorders (Fairburn 2001). People suffering from anorexia nervosa are 18 times more likely to have nearly death than the general population (Steinhausen 2009). People with binge-eating disorders tend to develop severe medical

The authors have no potential conflicts of interest.

conditions such as cardiovascular disease and high blood pressure (Wilfley *et al* 2000). Obesity is a condition of having excess body fat and is considered to be one of the major contributors towards the decrease in life expectancy in the USA (Olshansky *et al* 2005). According to the World Health Organization (WHO), overweight and obesity are the 5th major cause of death worldwide with 2.8 million people dying each year (WHO 2012). The study of ingestive behavior is particularly important to identify and diagnose food intake patterns associated with eating disorders and obesity. However, an accurate dietary assessment has been difficult to achieve due to the reliance on self-reporting and the lack of tools for objective monitoring of eating in free living conditions.

Food frequency questionnaires, food records and random 24-hour dietary recalls are commonly used methods for dietary monitoring that require active participation of the subjects in reporting their daily intake (Livingstone and Black 2003, Thompson and Subar 2008). These methods are subjective and inaccurate mainly due to incorrect reporting of foods consumed, erroneous estimations of portion sizes and failure to report certain foods (Black *et al* 1991, Livingstone and Black 2003). A potential solution based on electronic devices was presented to overcome self-reporting problems. Some of the techniques developed were based on the use of a mobile phone equipped with a digital camera (Liu *et al* 2012, Martin *et al* 2009, Weiss *et al* 2010). Subjects took pictures of the meal before and after eating while a computer algorithm was developed to determine the volume of food consumed using those pictures. These techniques may improve the accuracy of food intake monitoring, but they still require an active participation of the subjects.

Automatic methods for recognition of food intake were developed based on the identification of important features related to a particular stage of the food consumption process: hand gestures, bites, chewing and/or swallowing (Dong *et al* 2012, Jia *et al* 2012, Lopez-Meyer *et al* 2010, Päßler *et al* 2012, Passler and Fischer 2011, Sazonov and Fontana 2012, Sazonov *et al* 2008, Sun *et al* 2010). In most of the proposed methods, minimal participation of the subjects is required, thus reducing the recording burden, however, accuracy of food intake detection is still far from desired. A possible reason is that many methods of food intake detection are based on acoustic signals (Sazonov *et al* 2010, Amft 2010, Päßler *et al* 2012,) that suffer from sensitivity from external noise, which can hamper the performance in realistic environments outside of quiet laboratories. For example, (Sazonov *et al* 2010) used recognition of swallowing sounds recorded at the throat level using a miniature microphone. Individual swallows related to food intake were detected with an accuracy of 84.7% using individual models, with the experimental conditions including simulated noises of urban environment. An attempt to use noise cancellation techniques to improve the accuracy of food intake detection (Päßler *et al* 2012) used sounds recorded by a microphone located in the outer ear canal and a reference microphone to cancel out external noise. This method was able to detect food intake with an accuracy of 83% and to classify among 8 different food items with an accuracy of 79%. The relatively low accuracy of acoustical methods suggests that a methodology tolerant to significant levels of external noise would be of great interest for practical applications of food intake monitoring.

This paper presents a novel approach for food intake detection based on Electroglottography. An Electroglottograph (EGG) device is impervious to external noise

and operates by measuring the transverse electrical impedance across the neck at the larynx level. An EGG signal is recorded by sending and receiving a high frequency signal through guard-ring electrodes placed at the larynx level. For that reason, EGG has been widely used for speech and swallowing analysis (Childers and Larar 1984, Hodgson *et al* 2003, Nozaki *et al* 1994, Schultz *et al* 1994). Excitation frequencies ranging from 300 kHz to 5 MHz are generally used, so that the current avoids the less conductive skin layer without the use of an additional conductive paste (Rothenberg and Mahshie, 1988). Typical use of guard-ring electrodes provides a reference for noise reduction without the need of an extra electrode. In speech studies, EGG is used to monitor changes in the electrical impedance across the larynx due to the changes in the contacts in the vocal folds which are separated by glottis. During phonation the impedance across the larynx increases as the vocal folds move apart and decreases when the vocal folds come closer. In swallowing studies, EGG is used to monitor submental muscle activity and laryngeal elevation as suggested by (Ding *et al* 2002). Similar action of laryngeal elevation and submental muscle activity takes place when a bolus of food passes through larynx during food swallowing, thus, EGG can potentially be used for food intake detection by measuring changes in the neck impedance. According to authors' knowledge, no one has reported the use of EGG for food intake detection. As a side note, Electroglottography should not be confused with Electro-gastrography (measurement of electrical and magnetic fields of stomach muscle) that has the same acronym (EGG) and was suggested but not studied for monitoring of food intake in (Amft and Troster 2009).

The goal of this study was to evaluate the feasibility of using an EGG device and the related pattern-recognition methodology to detect periods of food intake by comparing the performance of the proposed method with the acoustical method in a controlled lab environment. The subsequent sections of the paper are organized as follows: Methods section provides a detailed description of data collection protocol and both signal processing and pattern-recognition algorithms. Results of the food intake detection by EGG and acoustical methods are presented in Section 3. Section 4 gives a detailed discussion of the results and compares the performance of the proposed system with other food intake monitoring systems found in the literature, which included different sensor modalities for detection of food consumption. Finally, Section 5 concludes the paper.

2. Methods

2.1 Data collection

Data from a group of 30 healthy subjects were collected for this study. Data from 5 subjects were later discarded due to equipment failure (i.e., electrode detachment) and operator's error during experiments (i.e., failure to center the video camera for subject observation). The remaining population consisted of 13 females and 12 males (average age of 29 ± 12 y, range: 19–58 y). The average body mass index (BMI) of the population (in kg/m^2) was 27.47 ± 5.45 (range: 20.5–41.7), which represented a range from normal weight to severely obese individuals. Subjects did not show any medical condition that affected their ability to eat. An Institutional Review Board approval for this study was obtained from Clarkson University, Potsdam, NY, and all subjects signed an informed-consent form before participation.

Each subject came to the laboratory in four separate visits, all of which occurred at the same time of day (breakfast, lunch or dinner). Each visit was divided into three parts:

- a. An initial resting period of 5 min in which subjects remained seated in relaxed position.
- b. An unlimited food intake period during which subjects were asked to consume a meal.
- c. A second resting period of 5 min identical to the first one.

Subjects self-selected two meals (different in content and size) from the menu of one of the food courts at Clarkson University. The first meals election was served in 3 of the visits and the second meal selection was served in the remaining visit. The consumption of the whole meal was not required, and there were no restrictions in the manner and order of food consumption. Contents of meals selected by subjects consisted of bananas, cereals, muffins, milk, coffee, toast, bacon, and different juices for breakfast. For lunch, subjects selected such foods as stir fry, chicken tenders, fries, chips, turkey sandwiches, pizza, burgers and salads with a wide variety of beverages. For dinner meals, subjects consumed items like corn, chef's salad, pizza, yogurt, spaghetti and meatballs, apples, cookies and brownies, along with soft drinks.

Subjects were allowed to talk, cough, clear throat, etc. and perform body movements while remaining seated (stretch, nod, turn head, etc.) at any time during the experiment, thus exposing the sensors to a variety of acoustical and motion artifacts not originating from food consumption.

A multimodal sensor system was used to monitor the subjects during the entire course of the experiment (Fontana *et al* 2011). A commercially available portable digital Laryngograph (EGG-D200 from Laryngograph, Ltd) was used to record the EGG signal $EGG(t)$. The excitation frequency was 3 Mhz and two gold-plated guard ring electrodes captured $EGG(t)$ within a 1 Hz to 10 kHz frequency range, which was then amplified in the Laryngograph unit. EGG electrodes were placed in contact with skin on both sides of the larynx using a collar. This is a standard location implemented in most of the swallowing studies (Childers and Larar 1984). A miniature throat microphone (IASUS NT) placed over the laryngopharynx was used to capture swallowing sounds. This microphone and location provided higher sensitivity to swallowing sounds and a lower sensitivity to noise when compared to other microphones and locations studied for food intake detection (Sazonov *et al* 2008). The microphone had a dynamic range of 46 +/- 3 dB with a frequency range of 10 Hz to 8 kHz. The acoustic signal ($MIC(t)$) was pre-amplified by a custom-designed amplifier. Figure 1 (left) shows the EGG electrodes and the throat microphone attached to a neoprene collar (paintball neck protector from JT Sports) for a comfortable wear and breathability. Figure 1 (right) show a subject wearing the collar equipped with the sensors. Both $EGG(t)$ and $MIC(t)$ were sampled at 44100 Hz with 16 bits of resolution using a USB-160HS-2AO data acquisition card (Measurement Computing) and stored on a computer. Figure 2 shows an example of $EGG(t)$ for a whole experiment. Approximately 60 hours of EGG and acoustic data were obtained from 25 subjects. Approximately 9 hours of data belonged to food intake (bites, chews and swallows).

A digital camera (PS3Eye camera, Sony) captured video of subjects time-synchronously with the sensor signal collection. The video stream and the sensor signals were then used to annotate periods of food intake by means of custom-designed LabVIEW software (Sazonov *et al* 2008). In the annotation process, a class label (T_i) was assigned to each sample of the sensor signals, where T_i {"no food intake," "food intake"}. For solid food items, a period of food intake involved a sequence of events including bite, chews and one or more swallows. For liquid items, a period of food intake involved a sip from the container and one or more swallows. Sensor signal epochs were annotated based on the presence of one of these events instead of the presence of individual swallows. A trained human rater reviewed the acquired video and sensor signals to identify when those events occurred in the experiment and to manually mark them in both $EGG(t)$ and $MIC(t)$. Inter-rater reliability of this annotation methodology was demonstrated in a previous study (Sazonov *et al* 2008). The remaining parts of the signals, including spontaneous swallows (saliva swallows), were marked as "no food intake". The annotated data was used as the gold standard for the development of automatic food intake detection algorithms. Figure 3 shows a segment of $EGG(t)$ during periods of resting and food intake along with an example of the food intake annotation. The resulting annotated data was used for training, validation and performance evaluation of the of the pattern-recognition algorithms.

2.2 Feature Extraction and Selection

Since both $EGG(t)$ and $MIC(t)$ carry similar spectral contents (from 1 Hz to 10 kHz for EGG and 10 Hz to 8 kHz for acoustical signal), they were processed with the same preprocessing and feature-extraction algorithms, thus eliminating any variability in the recognition results that could be attributed to the difference in the processing algorithms. First, signals were normalized with respect to their medians to account for variations in the signal amplitude among visits. Signals were then divided into non-overlapped time segments of fixed-time length referred to as epochs. Selection of the epoch length is important as it controls the time resolution of the decision stream. Using short epochs leads to a fine time resolution, which can help in the detection of short periods of food intake such as snacking. On the other hand, long epochs involve more data in the decision process, which may result in models with better performance but with poorer time resolution. Since the average frequency of spontaneous swallowing during waking is approximately 2 swallows per minute (Sazonov *et al* 2009), the food intake detection based on monitoring of swallowing pathway need to recognize an increase in swallowing rate (and corresponding changes in sensor signal) due to food consumption. Thus, the time resolution of such methodology is defined by the events with the lowest frequency (spontaneous swallowing) and an epoch length of 30 s was used as good trade-off between recognition accuracy and time resolution, capable of detecting small intake periods such as snacking. More information on epoch size selection and its relation to swallowing frequency may be found in (Sazonov *et al* 2008, Sazonov *et al* 2009 and Lopez-Meyer *et al* 2010).

The division of the signal into epochs resulted in some data samples labeled as food intake and some data samples labeled as no food intake within the same epoch. To tackle this issue, the 50% determination rule was implemented to assign a class label T_i {'+1', '-1'} to each 30s epoch. An epoch was labeled as "food intake" ($T_i = +1$) if at least 50% of the data

samples within the epoch belonged to food intake; otherwise the epoch was labeled as “no food intake” ($T_i = -1$), which means that if in an epoch, there was food intake of 15 s or more, the epoch was labeled as food intake. This ensured the detection of short intake periods. For feature computation, each epoch $x(n)$, (where $n = 1, 2, \dots, N$ is the total number of samples within the epoch), was decomposed using a discrete wavelet transform.

The Wavelet Transform (WT) is an alternate signal analysis technique to the Short Time Fourier Transform (STFT) and provides better time and frequency resolution compared to STFT. The WT provides high frequency resolution and low time resolution at low frequencies and vice versa at high frequencies. The Discrete Wavelet Transform (DWT) is an implementation of the WT that gives a computationally efficient and compact representation of the signal in time and frequency (Heil and Walnut 1989). The computational complexity of DWT is of the order of $O(N)$ for an N -length data sequence compared to the Fast Fourier Transform (FFT) which has a computational complexity of $O(N \log_2(N))$. The mother wavelet and the number of decomposition levels are important parameters of the DWT, which are selected based on the application. In this study, the Coiflet mother wavelet (coif5) was chosen for feature extraction as it is able to pick up details that are missed by the simpler wavelets and may be of importance for food intake detection. For each i -th epoch of 30 s, the DWT decomposition was performed at 4 levels using Wave-kit wavelet toolbox (Ojanen 1998), which resulted in five frequency sub-bands: $D1$ to $D4$ and $A4$ (detail and approximation coefficients respectively). The detail and approximation coefficients provided a compact representation of the signal's energy distribution in both time and frequency domain and were used as the features that represented the $EGG(t)$ and $MIC(t)$ signals. The feature vector f_i for each epoch of the signal was formed by computing various metrics over the set of the wavelet coefficients for each sub-band. As a result, each sub-band was represented by 10 features (Table 1), which were combined together to form f_i , consisting of 50 features (5 sub-bands with 10 features each). For each subject, features from all four visits were combined to form a combined feature vector which was used to train the food intake classification models.

2.3 Artificial Neural Network (ANN)

Artificial neural network (ANN) is a supervised machine learning technique that has shown excellent results for a number of pattern recognition and classification problems (Hudson and Cohen, 2000). ANN is robust, flexible and has the ability to create complex decision boundaries and handle noisy data. In this study, subject-independent models (group models) based on the ANN were trained using the wavelet features extracted from $EGG(t)$ (EGG model) and from $MIC(t)$ (MIC model). These models had the goal of detecting and separating food intake from other factors that resulted in changes of EGG and acoustical signals (such as intrinsic speech, head movements, bodily sounds, etc.). Subject-independent models were developed because they do not require individual calibration and ensure the applicability of the proposed technique to a wider population.

In this study, a three-layer (input layer, hidden layer and output layer) feed forward neural network trained by the back-propagation algorithm was used. In the input layer 50 predictors were used (one for each feature) whereas the hidden layer consisted of 10 neurons. The

output layer consisted of only one neuron, which indicated the final output T_i (“food intake” or “no food intake”) corresponding to the input feature vector f_i . The hyperbolic tangent sigmoid was the transfer function used for the hidden and output layers. Training, validation and testing of the model was done using the Neural Network toolbox available in Matlab R2011b (The Mathworks Inc).

A leave one out cross-validation procedure was used to evaluate the performance of the ANN model. Data from 24 subjects was divided into non-overlapping training (80% of the data) and validation (20% of the data) sets. Testing of the model was performed with data from the subject that was left out (25th subject). This procedure was repeated 25 times, so that each subject could be used as the test subject. Since the initial weights and bias were randomly generated, the leave-one-out cross validation procedure was performed 10 times to get generalizable results. The average of the accuracies for 10 iterations was taken to get final accuracy value for each subject. Finally, the overall result was obtained by averaging the results across all the subjects. Per-epoch classification accuracy was the metric used to evaluate the performance of the classification model. It was defined as the average between Precision and Recall, to account for a high number of true negatives that are typical in the monitoring of food intake over long periods of time.

$$Accuracy = (Precision + Recall) / 2 \quad (1)$$

$$Precision = \frac{T_+}{T_+ + F_+}; Recall = \frac{T_+}{T_+ + F_-} \quad (2)$$

where T_+ was the number of food intake epochs correctly classified by the model as food intake, F_+ was the number of no food intake epochs incorrectly classified as food intake epochs, and F_- as the number of food intake epochs incorrectly classified as no food intake epochs. The choice of accuracy metrics is defined by the fact that food intake constitutes only 2–3% of duration of daily activities (Fontana *et al* 2013), and a large number of true negatives (periods/epochs of no food intake) are typical for experiments in free living. Thus, precision and recall are best suited used for quantification of food intake detection accuracy as these metrics do not account for true negatives.

A statistical comparison (using t-test) was performed between *EGG* and *MIC* models for their ability to differentiate between epochs of food intake from epochs of no food intake. Both *EGG* and *MIC* models were evaluated for male and female subjects separately to determine if models were able to achieve similar accuracies for both genders and one tailed t-test (with significance level $p = 0.05$) was used to determine statistical significant difference between model performance. Additionally, the study's population was divided into 3 groups according to subjects' BMI (normal, overweight and obese) and a statistical analysis was performed to find significant differences among model performances for different level of adiposity (one way ANOVA with significance level $p = 0.05$).

Finally, to demonstrate the effect of background noise on the *EGG* and the acoustic signals, a small experiment was performed where a volunteer was asked to drink water with and without the presence of external acoustic noise originating from a laptop playing a song.

3. Results

For food intake recognition, the results of the leave-one-out cross validation procedure showed average accuracies of 90.1% (SD \pm 8.50%) and 83.1% (SD \pm 10.8%) for *EGG* and *MIC* models, respectively. The box-plot in figure 4 shows the accuracy distributions for each methodology. The statistical analysis showed significant differences between model performances (p value < 0.001).

Results of evaluating the models for both genders indicated that *EGG* model achieved average accuracies of 89.7% (SD \pm 8.35%) and 90.3% (SD \pm 8.97%) for female and male subjects respectively. A statistical analysis showed no significant differences between models performances for males and females (p > 0.05). Similarly, *MIC* model achieved average accuracies of 85.2% (SD \pm 10.3%) for male and 81.1% (SD \pm 11.3%) for female subjects without significant differences between performances (p > 0.05).

Results of model performances for different levels of adiposity are presented in figure 5. The *EGG* models achieved average accuracies of 91.8% (SD \pm 9.00%), 84.17% (SD \pm 9.92%) and 89.8% (SD \pm 5.84%) for normal weight, overweight and obese subjects respectively. The statistical analysis showed no significant differences among model performances for all BMI groups: (p > 0.05 in all cases). Similarly, *MIC* models achieved average accuracies of 84.8% (SD \pm 8.49%), 70.1% (SD \pm 15.4%) and 86.9% (SD \pm 7.28%) for normal weight, over-weight and obese subjects respectively. There were no significance differences in the performance of both models for corresponding BMI groups.

Results of testing *EGG* and microphone in the presence of background noise are shown in figure 6. Figure 6 (left) shows the power spectra of the *EGG* and figure 6 (right) shows acoustic signals with and without the presence of external noise in the experiment. The acoustic signal was highly affected by the presence of noise whereas in the *EGG* signal the effect of noise was insignificant.

4. Discussion

This study introduced a new sensor modality (*EGG*) for food intake detection and compared it with an acoustic based approach. In the experiment, a wide variety of food items consumed without restriction ensured that the food detection models did not over-fit to a specific food type or intake of solid or liquid foods. Sensor signals from both sensors were processed in exactly the same manner to eliminate any potential bias due to processing. Wavelet features extracted from sensor signals were used in conjunction with ANN classifiers to derive subject-independent food intake detection models that eliminated the need for individual calibration as they account for inter-subject variability.

Results suggest that the *EGG* models achieved a statistically-significant higher average accuracy of food intake detection than *MIC* models (90% vs. 83%, respectively). This showed that *EGG* is potentially superior for this application than the acoustic based approach with proposed signal processing and pattern-recognition techniques on this data set. The average recall of 91.8% and an average precision of 88.4% of *EGG*-based models

suggest good overall performance of the method with just a few false negatives and false positives.

An advantage of the proposed *EGG* models is their ability to detect food intake equally well for both genders and for people with different levels of adiposity with no statistically significant differences among model's performances for different genders and different BMI groups (normal, overweight and obese). This capability is critical for studying food intake in obese populations where adipose tissue may negatively impact accuracy of some sensing modalities (Sazonov *et al* 2008).

The proposed *EGG* based methodology is also non-invasive and not sensitive to acoustic noise by the virtue of the physical principles used in *EGG* measurement. Figure 6 illustrates practical effects of background noise on the sensor signals with *EGG* clearly not being affected to the same extent as the acoustical signal. This figure reinforced the fact the *EGG* is immune to acoustic noise compared to acoustic signal. In most cases, acoustic based methods require a preprocessing step for noise removal to improve the results. A previous study based on the acoustic classification of sounds for food intake detection achieved an accuracy of 83% (Päßler *et al* 2012) using a reference microphone to remove background noise. In this work, we demonstrated that the *EGG* sensor can achieve a high food intake detection rate without requiring any noise cancellation procedure. However, the practical benefit of *EGG*'s insensitivity to background noise (in terms of corresponding improvement in food intake recognition accuracy in free living environment) will need further investigation by testing the method in noisy environments that are typical in everyday living.

The current study has established feasibility of using *EGG* for food intake detection, but free-living tests will be required to evaluate performance of the device over extended periods of time under realistic conditions of daily living. Impedance measurement techniques such as *EGG* are sensitive to motion artifacts that may negatively impact accuracy of food intake detection, but are plentiful in daily life. Similarly, acoustical sensors are prone to contamination from body motion, external noise and speech. A future study should evaluate how the proposed technique performs in free living, how motion artifacts of daily life impact the sensor signals, what electrodes and methods of attachment perform the best, what motion artifact cancellation techniques are appropriate, if and how the *EGG* sensor can be made into a convenient and socially acceptable device and thus more thoroughly estimate suitability of *EGG*-based food intake detection for long-term monitoring.

Swallowing is a complex process that involves coordinated muscle contractions, laryngeal movement and passage of the food bolus. This study used changes in *EGG* signal due to swallowing activity to detect food intake without evaluating relative contributions in impedance change from these processes. At this moment it is not clear that the impedance variations were caused by the passage of food bolus, the swallowing motions and muscle activity or a combination of both. Exact quantification of the contribution of these phenomena to the resulting signal is beyond the scope of this manuscript, but should be considered in the future as food impedance may carry information on food composition. Bolus impedance may potentially carry information needed to differentiate between nutritive

and non-nutritive swallowing, and different food items consumed and estimate their energy density. However this hypothesis needs to be carefully tested in further studies. Another possibility to characterize ingested foods is the addition of a camera that can be automatically triggered to take a picture when the EGG methodology detect food intake. Image processing techniques would be required to estimate portion size and the type of food.

Overall, the proposed EGG-based approach showed good results for food intake detection by outperforming an acoustic-based method. While further investigation is needed to evaluate performance of EGG in free living and its potential to differentiate between different foods types, these results suggest that EGG use may potentially be a promising foundation for development of a wearable sensor system for detection of food intake under free living condition.

5. Conclusion

This paper evaluated the feasibility of using an Electroglottograph device for automatic and objective monitoring of ingestive behavior. Signals acquired by an Electroglottograph and a microphone at the larynx level were segmented into non-overlapping epochs of 30 s and a set of 50 features was computed based on DWT decomposition. These features were used to train subject-independent food intake detection models using an artificial neural network as a classifier. A leave-one-out cross validation scheme was used for training, validation and testing of the models, resulting in average food intake detection accuracy of 90.1% for EGG-based models and 83.1% for acoustical models. The difference between the average accuracies was statistically significant. The EGG models performed equally well for both genders and for people with different levels of adiposity. As a non-invasive and insensitive to background noise, the proposed EGG-based methodology justifies further investigation as a potential option for automatic and objective monitoring of ingestive behavior under free living conditions.

Acknowledgments

This work was partially supported by Grant Number R21DK085462 from the National Institute of Diabetes and Digestive and Kidney Diseases. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Diabetes and Digestive and Kidney Diseases or the National Institutes of Health.

References

- Amft O. A wearable ear pad sensor for chewing monitoring. *IEEE Sensors*. 2010:222–227.
- Amft O, Troster G. On-Body Sensing Solutions for Automatic Dietary Monitoring. *IEEE Pervasive Comput*. 2009; 8:62–70.
- Black AE, Goldberg GR, Jebb SA, Livingstone MB, Cole TJ, Prentice AM. Critical evaluation of energy intake data using fundamental principles of energy physiology: 2. Evaluating the results of published surveys. *Eur J Clin Nutr*. 1991; 45:583–599. [PubMed: 1810720]
- Childers DG, Larar JN. Electroglottography for Laryngeal Function Assessment and Speech Analysis. *IEEE Trans Biomed Eng*. 1984; 31:807–817. [PubMed: 6526406]
- Ding R, Larson CR, Logemann JA, Rademaker AW. Surface electromyographic and electroglottographic studies in normal subjects under two swallow conditions: normal and during the Mendelsohn maneuver. *Dysphagia*. 2002; 17:1–12. [PubMed: 11820381]

- Dong Y, Hoover A, Scisco J, Muth E. A new method for measuring meal intake in humans via automated wrist motion tracking. *Appl Psychophysiol Biofeedback*. 2012; 37:205–215. [PubMed: 22488204]
- Fairburn, CG. *Eating Disorders* eLS. John Wiley & Sons, Ltd; 2001.
- Fontana, JM.; Farooq, M.; Sazonov, E. Estimation of feature importance for food intake detection based on Random Forests classification; Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); 2013. p. 6756-6759.
- Fontana JM, Lopez-Meyer P, Sazonov ES. Design of a instrumentation module for monitoring ingestive behavior in laboratory studies. *EMBC*. 2011:1884–1887.
- Heil CE, Walnut DF. Continuous and Discrete Wavelet Transforms. *Siam Rev*. 1989; 31:628–666.
- Hodgson M, Linforth RS, Taylor AJ. Simultaneous real-time measurements of mastication, swallowing, nasal airflow, and aroma release. *J Agric Food Chem*. 2003; 51:5052–5057. [PubMed: 12903969]
- Hudson DL, Cohen ME. *Neural networks and artificial intelligence for biomedical engineering*. (IEEE Press). 2000
- Jia, W.; Yue, Y.; Fernstrom, JD.; Zhang, Z.; Yang, Y.; Sun, M. 3D localization of circular feature in 2D image and application to food volume estimation; Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); 2012. p. 4545-4548.
- Liu, J.; Johns, E.; Atallah, L.; Pettitt, C.; Lo, B.; Frost, G.; Yang, G-Z. An Intelligent Food-Intake Monitoring System Using Wearable Sensors; 2012 Ninth International Conference on Wearable and Implantable Body Sensor Networks (BSN); 2012. p. 154-160.
- Livingstone MBE, Black AE. Markers of the validity of reported energy intake. *J. Nutr*. 2003; 133(Suppl 3):895S–920S. [PubMed: 12612176]
- Lopez-Meyer P, Makeyev O, Schuckers S, Melanson E, Neuman M, Sazonov E. Detection of food intake from swallowing sequences by supervised and unsupervised methods. *Ann Biomed Eng*. 2010; 38:2766–2774. [PubMed: 20352335]
- Makeyev, O.; Sazonov, E.; Schuckers, S.; Lopez-Meyer, P.; Baidyk, T.; Melanson, E.; Neuman, M. Recognition of swallowing sounds using time-frequency decomposition and limited receptive area neural classifier; The Twenty-eighth SGAI International Conference on Innovative Techniques and Applications of Artificial Intelligence; 2008. p. 33-46.
- Martin CK, Han H, Coulon SM, Allen HR, Champagne CM, Anton SD. A novel method to remotely measure food intake of free-living individuals in real time: the remote food photography method. *Br J Nutr*. 2009; 101:446–456. [PubMed: 18616837]
- Ojanen H. Wavekit: a Wavelet Toolbox for Matlab. 1998 available at <http://www.math.rutgers.edu/~ojanen/wavekit/>.
- Olshansky SJ, Passaro DJ, Hershov RC, Layden J, Carnes BA, Brody J, Hayflick L, Butler RN, Allison DB, Ludwig DS. A potential decline in life expectancy in the United States in the 21st century. *N. Engl. J. Med*. 2005; 352:1138–1145. [PubMed: 15784668]
- Päßler S, Wolff M, Fischer W-J. Food intake monitoring: an acoustical approach to automated food intake activity detection and classification of consumed food. *Physiol Meas*. 2012; 33:1073–1093. [PubMed: 22621915]
- Passler, S.; Fischer, W-J. Food Intake Activity Detection Using a Wearable Microphone System; 2011 7th International Conference on Intelligent Environments (IE); 2011. p. 298-301.
- Rothenberg M, Mahshie JJ. Monitoring vocal fold abduction through vocal fold contact area. *J Speech Hear Res*. 1988; 31:338–351. [PubMed: 3172751]
- Sazonov E, Fontana J. A Sensor System for Automatic Detection of Food Intake Through Non-Invasive Monitoring of Chewing. *IEEE Sens J*. 2012; 12:1340–1348. [PubMed: 22675270]
- Sazonov ES, Makeyev O, Schuckers S, Lopez-Meyer P, Melanson EL, Neuman MR. Automatic detection of swallowing events by acoustical means for applications of monitoring of ingestive behavior. *IEEE Trans Biomed Eng*. 2010; 57:626–633. [PubMed: 19789095]
- Sazonov ES, Schuckers SAC, Lopez-Meyer P, Makeyev O, Melanson EL, Neuman MR, Hill JO. Toward Objective Monitoring of Ingestive Behavior in Free-living Population. *Obesity*. 2009; 17:1971–1975. [PubMed: 19444225]

- Sazonov E, Schuckers S, Lopez-Meyer P, Makeyev O, Sazonova N, Melanson EL, Neuman M. Non-invasive monitoring of chewing and swallowing for objective quantification of ingestive behavior. *Physiol Meas.* 2008; 29:525–541. [PubMed: 18427161]
- Schultz JL, Perlman AL, VanDaele DJ. Laryngeal movement, oropharyngeal pressure, and submental muscle contraction during swallowing. *Arch Phys Med Rehabil.* 1994; 75:183–188. [PubMed: 8311675]
- Steinhausen H-C. Outcome of Eating Disorders Child. *Adolesc Psychiatr Clin N Am.* 2009; 18:225–242.
- Sun M, Fernstrom JD, Jia W, Hackworth SA, Yao N, Li Y, Li C, Fernstrom MH, Sclabassi RJ. A Wearable Electronic System for Objective Dietary. Assessment *J. Am. Diet. Assoc.* 2010; 110:45–47.
- Thompson, FE.; Subar, AF. *Nutrition in the Prevention and Treatment of Disease.* San Diego, CA: Academic Press; 2008. Dietary assessment methodology.
- Weiss R, Stumbo PJ, Divakaran A. Automatic food documentation and volume computation using digital imaging and electronic transmission. *J Am Diet Assoc.* 2010; 110:42–44. [PubMed: 20102824]
- Wilfley DE, Schwartz MB, Spurrell EB, Fairburn CG. Using the eating disorder examination to identify the specific psychopathology of binge eating disorder. *Int J Eat Disord.* 2000; 27:259–69. [PubMed: 10694711]
- World Health Organization. Obesity and overweight. 2012. available at <http://www.who.int/mediacentre/factsheets/fs311/en>

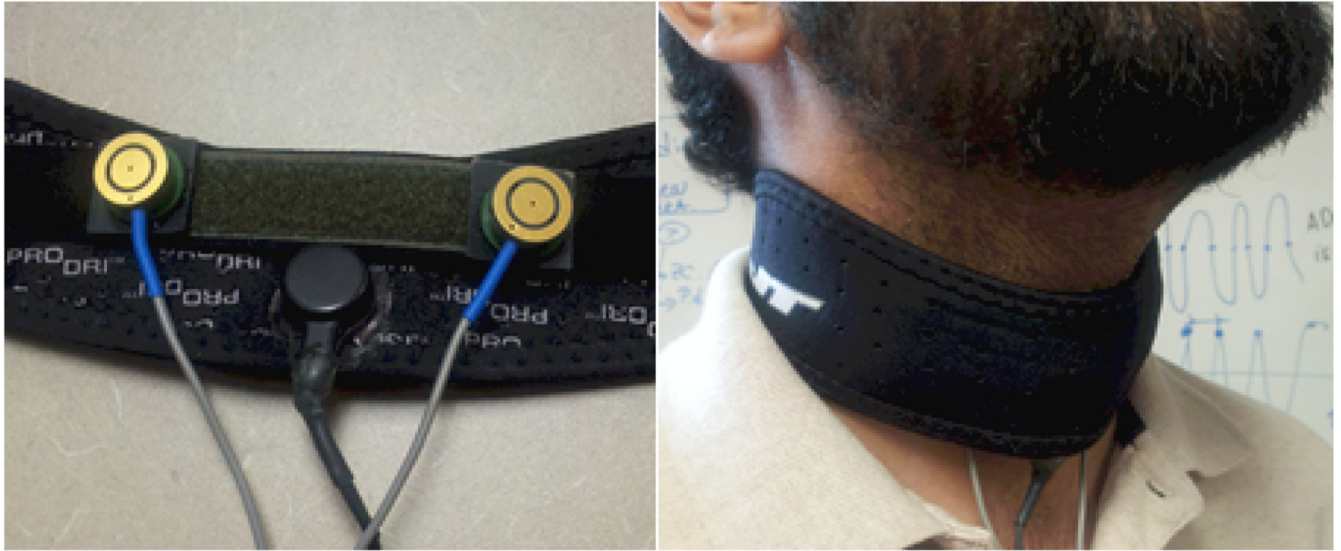


Figure 1.
Left: EGG and swallowing sensors attached to a neoprene collar. Right: Collar was fastened to the neck of the subject using flexible Velcro.

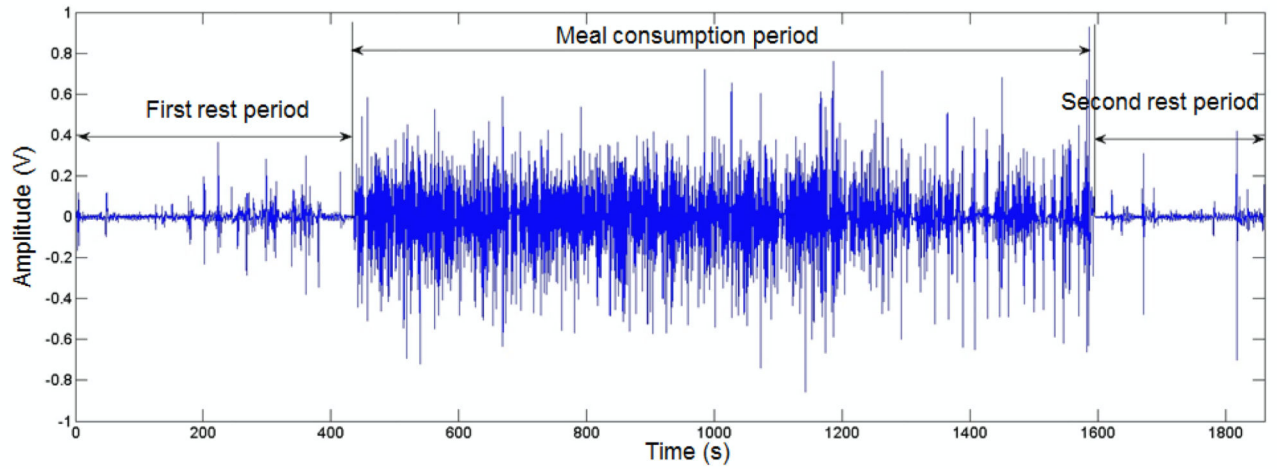


Figure 2.

EGG signal for a complete experiment of about 30 min in duration. Note that marked meal consumption period does not present a contiguous intake of food, but a complex microstructure with pauses in food intake, conversations with the research personnel and motion artifacts.

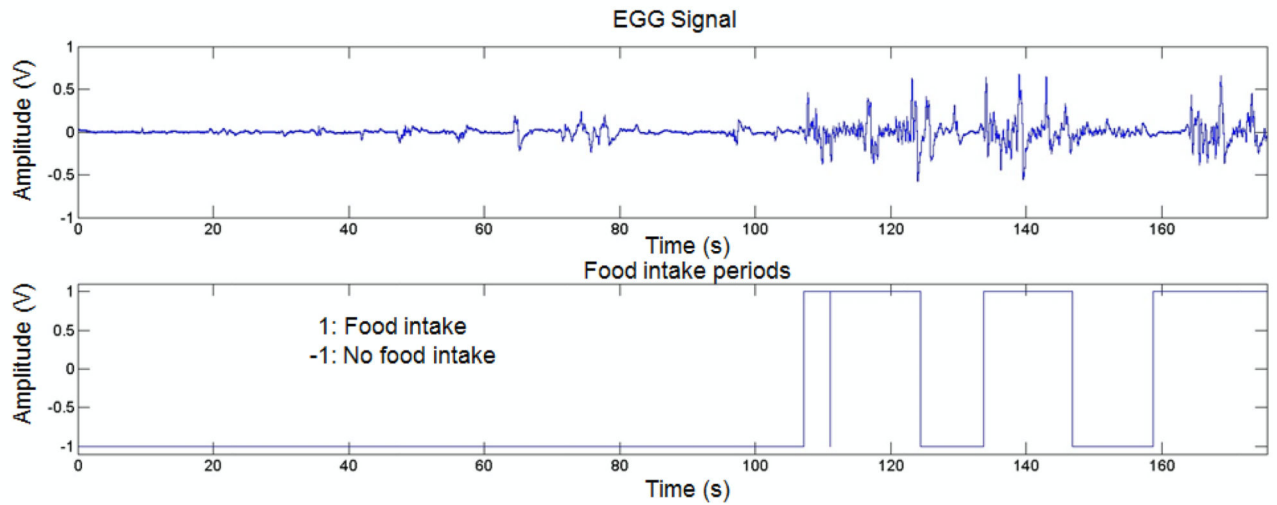


Figure 3.

A close-up of EGG signal and corresponding annotation of food intake. Note the complex structure of food intake marked on bottom graph and presence of artifacts on the EGG signal in the periods of “no food intake”.

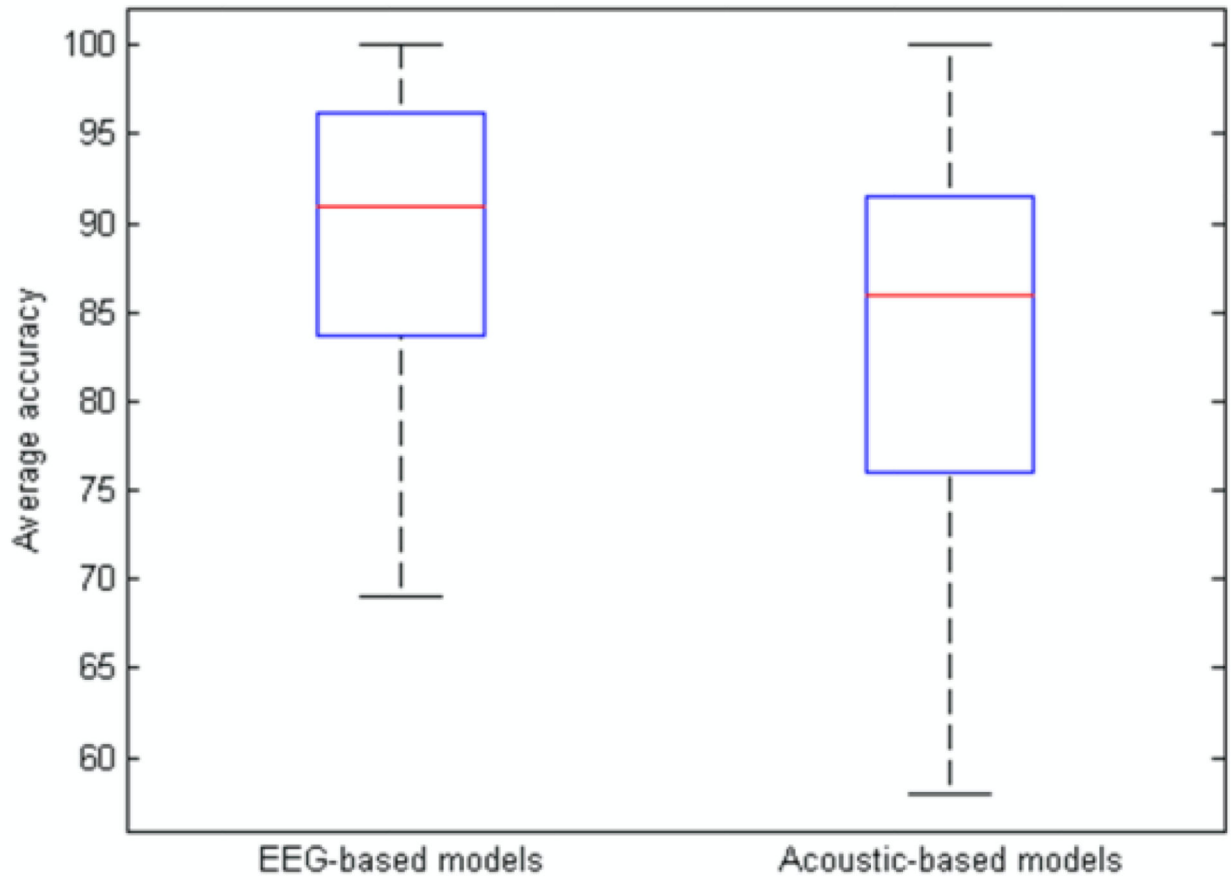


Figure 4. Distribution of the food intake detection accuracies obtained from both EEG-based and acoustic-based models.

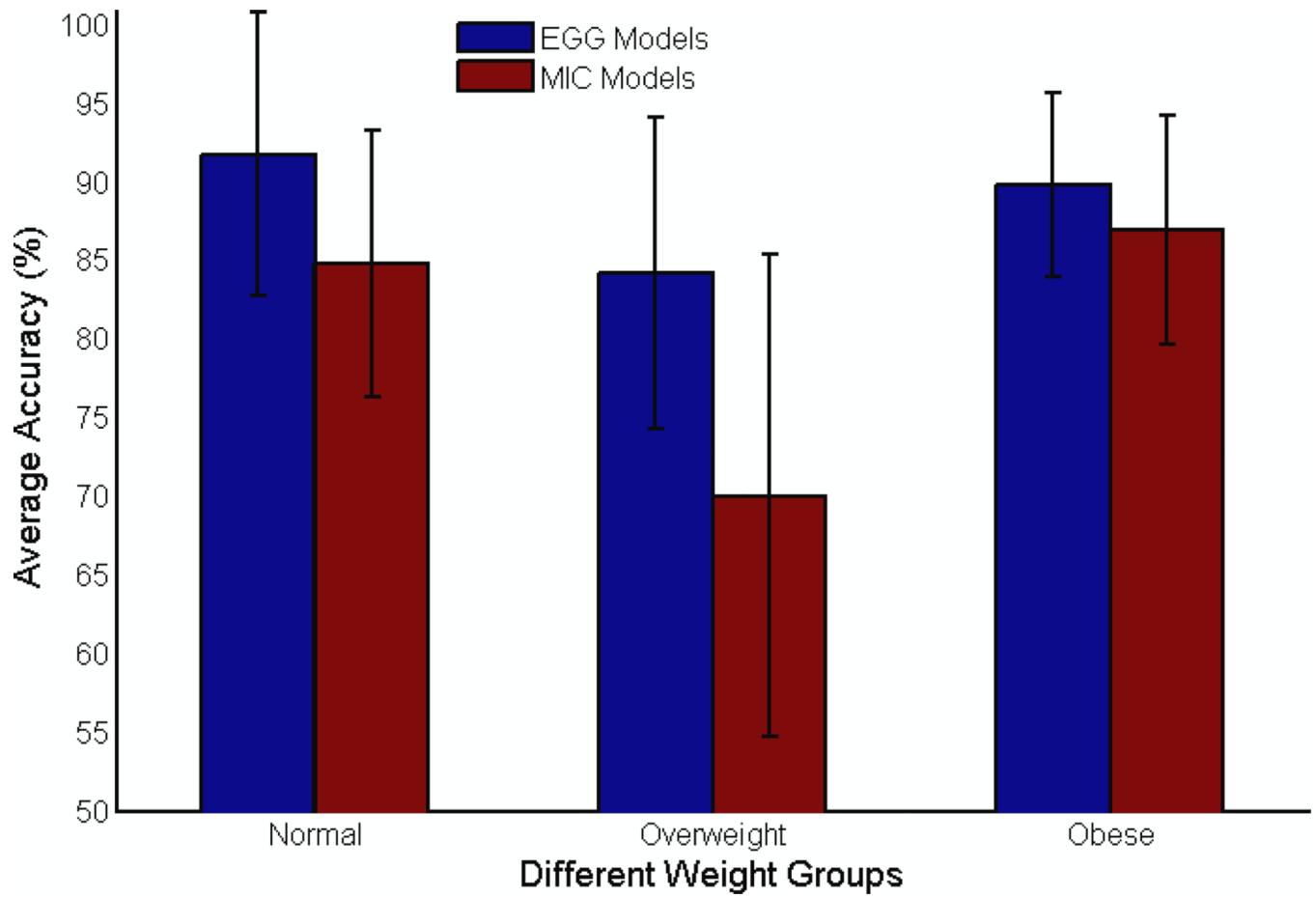


Figure 5.
Model performance for different levels of adiposity.

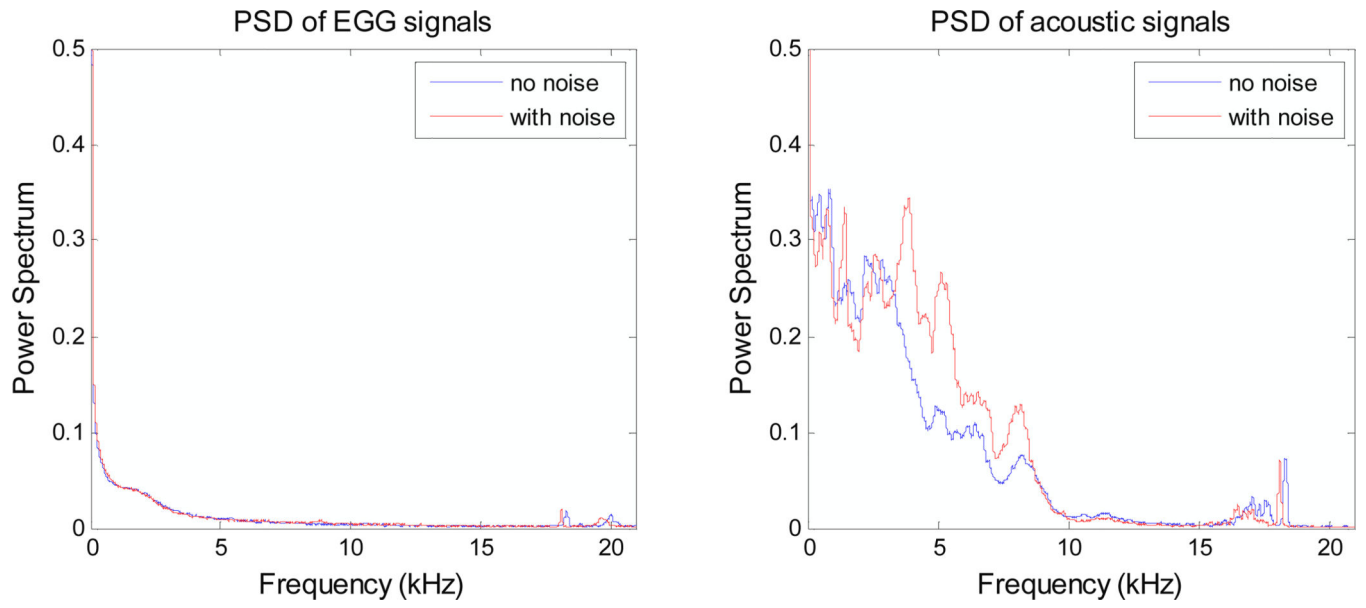


Figure 6. Left: Power spectral density (PSD) of EGG signal with and without noise Right: PSD of acoustic signal for the same experiment.

Table 1Features computed for each sub-band^a.

Feat #	Feature	Description
1	RMS value of $C(n)$	$RMS = \sqrt{\frac{1}{N} \sum_{n=1}^N C(n)^2}$
2	Entropy of $C(n)$	$H(C(n)) = (-\sum p_c \log_2(p_c))$; where p_c is the histogram of $C(n)$
3	Mean of Absolute Value (MAV)	$MAV = \frac{1}{N} \sum_{n=1}^N C(n) $
4	Max of Abs	$MaxAbs = \max(C(n))$
5	Ratio Max of Abs to RMS	$R_1 = MaxAbs/RMS$
6	Ratio RMS to MAV	$R_2 = RMS/MAV$
7	Standard deviation of $C(n)$	$\sigma_{X(n)} = \sqrt{\frac{1}{N} \sum_{n=1}^N (C(n) - \overline{C(n)})^2}$
8	Energy of $C(n)$	$E = \sum_{n=1}^N C(n)^2$
9	Power of $C(n)$	$P = \frac{1}{N} \sum_{n=1}^N C(n)^2$
10	Skew-ness $C(n)$	$\gamma = \frac{\sum_{n=1}^N (C(n) - \overline{C(n)})^3}{(N-1)\sigma^3}$ where σ is the standard deviation of $C(n)$

^a $C(n) = C(1), C(2), \dots, C(N)$ are the coefficients in each sub-band after DWT decomposition of each epoch.