Obesity Medicine 1 (2016) 6–14

Contents lists available at ScienceDirect

Obesity Medicine

journal homepage: http://www.journals.elsevier.com/obesity-medicine

Original research

A comparison of piezoelectric-based inertial sensing and audio-based detection of swallows

Haik Kalantarian [a,*], Bobak Mortazavi [c], Nabil Alshurafa [b], Costas Sideris [a], Tuan Le [a], Majid Sarrafzadeh [a]

[a] Department of Computer Science, University of California, Los Angeles, United States
[b] Department of Preventative Medicine, Northwestern University, United States
[c] School of Medicine, Yale University, United States

ARTICLE INFO

Article history:
Received 21 December 2015
Accepted 12 January 2016

Keywords:
Piezoelectric sensor
Wireless health
Nutrition
Necklace

ABSTRACT

Background: Prior research has shown a correlation between poor dietary habits and countless negative health outcomes such as heart disease, diabetes, and certain cancers. Automatic monitoring of food intake in an unobtrusive, wearable form-factor can encourage healthy dietary choices by enabling individuals to regulate their eating habits.
Methods: This paper presents an objective comparison of two of the most promising methods for digital dietary intake monitoring: piezoelectric swallow sensing by means of a smart necklace which monitors vibrations in the neck, and audio-based detection using a throat microphone.
Results: Data was collected from twenty subjects with ages ranging from 22 to 40 as they consumed a variety of foods using both devices. In Experiment I, we distinguished sandwich, chips, and water. In Experiment II, we distinguished nuts, chocolate, and a meat patty. F-Measures for the audio based approach were 91.3% and 88.5% for the first and second experiments, respectively. In the piezo-based approach, F-measures were 75.3% and 79.4%.
Conclusion: The accuracy of the audio-based approach was significantly higher for classifying between different foods. However, this accuracy comes at the expense of computational overhead increased power dissipation due to the higher sample rates required to process audio signals compared to inertial sensor data.

1. Introduction

Healthy eating can reduce the risk of heart disease, stroke, diabetes, and several cancers. In 2008, medical costs associated with obesity were estimated at $147 billion, and the Centers for Disease Control (CDC) believes that the best areas for treatment and prevention are monitoring behavior and environment settings (Centers for disease control, 2014). Wireless technologies and health-related wearable devices have the potential to enable healthier lifestyle choices. These devices and systems are designed to encourage behavior modifications needed to reduce the risk of obesity and obesity-related diseases (Dorman et al., 2010).

Studies have shown that the number of swallows recorded during a day strongly correlate with weight gain on the following day (Stellar and Shrager, 1985). This provides motivation for the analysis of food intake patterns based on volume. Though many wearable devices have been designed for monitoring activity (Freedson et al., 1998; 2011; Patel et al., 2012), automatically and accurately inferring eating durations and patterns in a non-intrusive manner has been for the most part an unaddressed challenge.

Prior works have attempted to characterize eating habits through various means. Though many methods have been proposed, two of the more promising techniques include inertial-systems using piezoelectric sensors, as well as audio-based detection using throat microphones. In piezoelectric-based techniques, piezoelectric sensors, which produce a voltage in response to mechanical stress, can be used to detect movement in the skin on the lower-neck associated with swallowing. This approach differs from

* Corresponding author.
E-mail addresses: kalantarian@cs.ucla.edu (H. Kalantarian), bobak.mortazavi@yale.edu (B. Mortazavi), nabil@northwestern.edu (N. Alshurafa), costas@cs.ucla.edu (C. Sideris), tuanle@cs.ucla.edu (T. Le), majid@cs.ucla.edu (M. Sarrafzadeh).

http://dx.doi.org/10.1016/j.obmed.2016.01.003
2451-8476/

microphones based on piezoelectric technology: our system does not detect sound waves, instead assessing motion in the skin that results from swallows and chewing. Alternatively, audio-based techniques typically place a small microphone near the jaw or neck, and record eating noises such as chewing and swallowing. These sounds can be disambiguated from other background noises using classifiers and other signal-processing techniques. These approaches differ significantly from a perspective of comfort, practicality, convenience, power usage, and detection accuracy.

The primary novelties of our work are the description of a system in which a piezoelectric sensor is placed in the lower part of the neck for detecting swallow motions, and a comparison of this technique with audio-based monitoring using datasets derived from the same experiments. This provides a much more objective comparison of these two technologies than otherwise possible by comparing results from separate papers using different datasets and methodologies. Furthermore, we provide an evaluation of the power overhead of these techniques as a function of sample rate, computational overhead, and Bluetooth connection interval.

This paper is organized as follows. Section 2 presents related work in dietary monitoring technologies. In Section 3, we describe the hardware architecture of the two schemes, followed by algorithms in Section 4. In Section 5, we describe the experimental procedure. In Section 6, we describe experimental results. In Section 7, we describe our methods for monitoring the power and energy overhead of these techniques, which is followed by a presentation of results in Section 8. Finally, limitations and future work are described in Section 9, followed by concluding remarks in Section 10.

2. Related works

Many works have employed microphones for detecting food intake. For example, the work in Sazonov et al. (2008) uses acoustic data acquired from a small microphone placed near the bottom of the throat. Their system is coupled with a strain gauge placed near the ear. Other works suggest the use of throat microphones as a means of acquiring audio signals from throat and extracting swallowing sounds, for evaluation of dysphagia symptoms in seniors (Nagae and Suzuki, 2011; Tsujimura et al., 2010). Analyzing wave shape in the time domain or feature extraction and machine learning (Tsujimura et al., 2010) has resulted in an 86% swallow detection accuracy in an in-lab controlled environment. Similarly, the work featured in Nagae and Suzuki (2011) by Nagae et al. attempts to distinguish between swallowing, coughing, and vocalization using wavelet-transform analysis of audio data. However, identifying the volume or characteristic of food intake is not the focus of their work.

In Rahman et al. (2014), Rahman et al. present BodyBeat: a robust system for detecting human sounds. A similar work is presented by Yatani et al. in Yatani and Truong (2012). Our work differs from theirs for a number of reasons. First, we do not propose a custom hardware solution, instead employing a simple off-the-shelf throat microphone that connects directly to a mobile phone. Secondly, we emphasize classification between different foods, comparing the properties of celery, chocolate, nuts, water, chips, and sandwiches. Furthermore, we perform real-time experiments to measure the power overhead of frequency domain audio analysis, and Bluetooth 4.0 LE transmission of audio signals. Lastly, we directly compare this approach to the inertial-sensing approach on the basis of classification accuracy, and computational overhead.

In the work by Amft et al. in Amft et al. (2009), authors analyze bite weight and classify food acoustically from an earpad-mounted sensor. However, sound-based chewing recognition accuracy was low, with a precision of 60%–70%. In Amft (2010), the authors present a similar earpad-based sensor design to monitor chewing sounds. Food grouping analysis revealed three significant clusters of food: wet and loud, dry and loud, soft and quiet. An overall recognition accuracy of over 86.6% was achieved. Some studies have reached accuracy rates of 91.7% in an in-lab controlled environment using neural networks with false positives of 9.5% (Aboofazeli and Moussavi, 2004). A more recent study using support vector machines have been able to reach swallow detection accuracies of up to 84.7% in an in-lab setting (Sazonov et al., 2010). These devices are mounted very high in the upper trachea, near the laryngopharynx. In Passler and Fischer (2011), Pler, et al. proposed a system geared towards patients living in ambient assisted living conditions and used miniature electret microphones which were integrated into a hearing aid case, and placed in the ear canal. Our prior work described in Kalantarian et al. (2014a) also provided a foundation for spectrogram-based analysis of audio signals. A similar approach for analyzing bioacoustic signals using spectrograms was also presented by Pourhomayoun et al. in Pourhomayoun et al.

A “smart tablecloth” was presented in 2015 by Bo Zhou et al. in Zhou et al. (2015). The system detects eating behavior on solid surfaces (such as tables), based on changes in the pressure distribution of these tables during the eating process. The tablecloth was a matrix of pressure sensors based on a carbon polymer sheet, which changes its electrical resistance in response to electrical force. At the corners of the tablecloth, force-sensitive resistors (FSRs) are installed with the primary purpose of determining weight, rather than spatial density. Features extracted from the FSRs, as well as the pressure-sensitive tablecloth, are analyzed using classifiers such as decision trees to distinguish between various eating-related actions such as stirring, scooping, and cutting. Based on the ratio of different actions performed, the authors were able to distinguish between four different meal types with high accuracy. Furthermore, changes in the average pressure values from the data stream were associated with a decrease in the remaining amount of food on the table, which was used to estimate food weight with an error of approximately 16.62%.

The E-Button was presented in 2014 by Professor Mingui Sun at the University of Pittsburgh (Sun et al., 2014). In this work, Sun et al. propose a chest-mounted button with an embedded camera that among other applications, can be applied to the domain of dietary monitoring. The button is attached to a shirt using a pin or pair of disk magnets, and contains an ARM Cortex processor, two wide-angle cameras, a UV sensor for distinguishing between indoor and outdoor environments, inertial sensors, proximity sensors, a barometer, and a GPS. The acquired data is transmitted to a smartphone using Bluetooth or WiFi. The E-Button operates by taking photos at a preset rate, thereby recording the entire eating process. Using image processing techniques, the utensils (such as a plate or bowl) are detected. Subsequently, the food items are identified based on color, texture, and other heuristics. Using this information and additional DSP techniques, volume figures can be calculated for each food, which is converted to a Calorie count using a public domain database that equates volume and food type to Calories. Evaluation of 100 foods was conducted, and the error was approximately 30% for 85% of the foods, which were regularly shaped. However, irregularly shaped food was not detected with high accuracy.

Several prior works have attempted to detect swallow disorders using piezoelectric sensors. The work by Toyosato et al. in Toyosato et al. (2007) used a Piezoelectric Pulse Transducer to detect food bolus passage through the esophagus. In Ertekin et al. (1996), Ertekin et al. used piezoelectric sensors to evaluate dysphagia symptoms in a study with thirty normal subjects and 66 dysphagia patients. The authors concluded that piezoelectric sensors can be