Computers in Biology and Medicine 65 (2015) 1–9

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/cbm

Audio-based detection and evaluation of eating behavior using the smartwatch platform

Haik Kalantarian, Majid Sarrafzadeh

Wireless Health Institute, Department of Computer Science, University of California, Los Angeles, United States

ARTICLE INFO

Article history:
Received 2 May 2015
Accepted 16 July 2015

Keywords:
Signal processing
Wireless Health
Nutrition
Smartwatch
Machine learning
Pervasive computing

ABSTRACT

In recent years, smartwatches have emerged as a viable platform for a variety of medical and health-related applications. In addition to the benefits of a stable hardware platform, these devices have a significant advantage over other wrist-worn devices, in that user acceptance of watches is higher than other custom hardware solutions. In this paper, we describe signal-processing techniques for identification of chews and swallows using a smartwatch device's built-in microphone. Moreover, we conduct a survey to evaluate the potential of the smartwatch as a platform for monitoring nutrition. The focus of this paper is to analyze the overall applicability of a smartwatch-based system for food-intake monitoring. Evaluation results confirm the efficacy of our technique; classification was performed between apple and potato chip bites, water swallows, talking, and ambient noise, with an *F*-measure of 94.5% based on 250 collected samples.

1. Introduction

There is little doubt that obesity is associated with various negative health outcomes such as an increased risk for stroke, diabetes, various cancers, heart disease, and other conditions. In 2008, medical costs associated with obesity were estimated to exceed $147 billion, with over one-third of adults in the United States estimated to be obese [1]. The two major contributors to weight gain are an inactive lifestyle and poor diet. Though the former has been addressed by many wearable devices in recent years both in research and the consumer electronics field, few works exist on automatic detection of dietary habits in an inconspicuous form-factor [2–4]. Instead, characterization of an individual's eating habits is possible through manual record keeping such as food diaries, 24-h recalls, and food frequency questionnaires. However, these approaches suffer from low accuracy, high user burden, and low rates of long-term compliance. Wireless health-monitoring technologies have the potential to promote healthy behavior and address the ultimate goal of enabling better lifestyle choices.

In recent years, several electronic devices have been proposed for monitoring dietary habits. However, most works attempt to characterize eating from patterns in chewing and swallow counts, and very few proposed attempt to identify the nutritive properties of the foods themselves. Therefore, a fundamental question in the field of electronic food monitoring is the validity of chew and swallow counts as a heuristic for estimation of Caloric intake. A recent work by Fontana et al. [5] addresses this issue by comparing several different techniques for estimation of Caloric intake: weighed food records (gold standard), diet diaries, and electronic sensor-based measurements of chews and swallows. Though the study was conducted under constrained conditions, the results suggest that chew and swallow counts may be a promising alternative to manual self-reporting techniques.

While many audio-based nutrition monitors are novel from a perspective of algorithmic techniques, they generally propose custom hardware solutions or bulky non-standard equipment which are of limited use outside of clinical environments. The primary challenge of monitoring a subject's eating habits is creating a system that provides passive monitoring of behavior, presenting a low level of user burden and providing no compromises on comfort and appearance: even the most accurate techniques have very limited scope if they do not encourage repeated use from users.

Recently, smartwatches have emerged as a new platform that provide several promising applications such as wrist-worn activity monitoring, heart rate tracking, and even stress measurement. Watch usage is well established and has a high level of social acceptance, as confirmed not only by our personal studies but by their ubiquity in day-to-day life. Furthermore, the smartwatch platform provides many useful services that can collectively improve user adherence rates, rather than specialized devices

E-mail addresses: kalantarian@cs.ucla.edu (H. Kalantarian), majid@cs.ucla.edu (M. Sarrafzadeh).

http://dx.doi.org/10.1016/j.compbiomed.2015.07.013
0010-4825/

with just one application that may fail to sustain a user's interest. These devices contain a multitude of sensors including but not limited to: a microphone, camera, accelerometer, and gyroscope. Due to the ubiquity of watches, this technology can be used for various wireless health monitoring applications discretely, with low user burden. Furthermore, from a user-acceptance standpoint, these systems have a clear advantage over other proposed solutions based on custom hardware, which may require that these bulky and non-standard devices be worn in unconventional ways. Clearly, the multitude of sensors available on the smartwatch platform, wireless connectivity, as well as the comfort and social acceptance of the form-factor warrant further study into their potential applications in the medical and health-monitoring domain.

This paper explores the idea of tracking eating habits using a custom Android application on the smartwatch platform. Though identifying eating-related gestures using wrist-worn devices is a viable application of the watch, the focus of our work is to explore the idea of using audio to detect eating behavior based on bites, rather than swallows as other works have done. A high-level system architecture is presented in Fig. 1. The first step is audio-based acquisition of eating-related sounds such as bites, acquired from the microphone integrated within the smartwatch. After data acquisition, the audio is processed using various classifiers to identify the sound and infer the associated activity.

In addition, we conducted two surveys in order to evaluate the potential of the smartwatch platform for nutrition monitoring. The surveys were conducted online, with 221 respondents in the first and 55 in the second. In the first survey, we asked subjects various questions about their general habits with respect to watches. For example, subjects were asked which hand they prefer to wear a watch, and whether they were willing to wear a watch on the opposite hand on which they were accustomed. In the second survey, respondents provided information about their opinion of various wearable form factors. Fifty-five subjects rated their receptiveness to smartwatches, necklace-based wearables, custom wrist-worn hardware, and smart glasses.

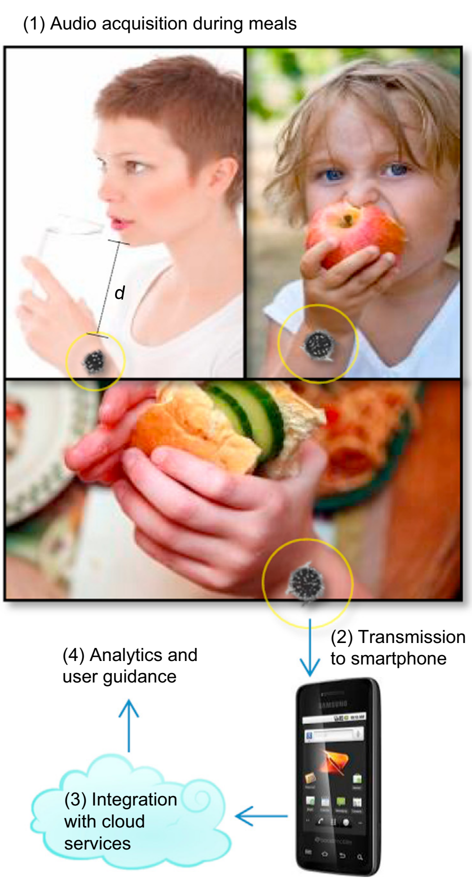

Fig. 1. A high level architecture of the proposed system is shown above. Many different forms of eating can be detected using a smartwatch, provided the appropriate hand is used and the watch is brought close enough to the mouth.

This paper is organized as follows. Section 2 provides an overview of related work, primarily in the scope of audio-based analysis of eating habits. Section 3 describes the hardware architecture of the system, based on around the Samsung Galaxy Gear smartwatch. Section 4 describes the algorithmic aspects of our work. Section 5 outlines the experimental procedure. Section 6 provides results and Section 7 provides concluding remarks.

2. Related work

The use of audio signals for analysis of swallows or eating behavior has been explored in several other works. For example, the work in [6] 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. In this work, acquired data is manually labelled to provide a benchmark for future classification. Analyzing wave shape in the time domain or feature extraction and machine learning [7] has resulted in an 86% swallow detection accuracy in an in-lab controlled environment. In [8], 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. In [9], the authors are able to achieve a food detection accuracy of 79% using hidden Markov models based on data acquired from microphones in the ear canal.

In the work by Amft et al. in [10], 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 [11], 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.

Though the signal processing aspects of this application are relatively well developed, our work differs from prior approaches in two significant ways. First, we propose the use of an audio spectrogram for representing the changes in the frequency distribution of the signal over time, which is subsequently subdivided into bins and used for feature extraction and selection. Therefore, the classifier can distinguish between different foods based on the frequency distribution of the signal over time. Secondly, prior works rely on cumbersome multi-sensor hardware approaches that have little utility out of laboratory environments, while our algorithm runs on off-the-shelf Samsung hardware. Lastly, we show how the extensive openSMILE feature extraction tool for analyzing human vocalizations can be applied to other domains for classification of sounds unrelated to speech patterns.

3. System architecture

Our proposed system does not require any custom hardware: the Android application runs on Samsung Galaxy Gear smartwatch running Android 4.2.1. This device, shown in Fig. 2, features an 800 MHz ARM-based processor, 512 MB of RAM, and a 320×320 pixel 1.6 inch display. The device also supports transfer of data using the Bluetooth LE protocol, and can be configured to access the Internet using Bluetooth tethering with compatible