

A Wearable System for Detecting Eating Activities with Proximity Sensors in the Outer Ear

Abdelkareem Bedri, Apoorva Verlekar, Edison Thomaz, Valerie Avva, Thad Starner
{abedri6, ayp3, ethomaz, vavva3, thad}@gatech.edu
Georgia Institute of Technology

ABSTRACT

This paper presents an approach for automatically detecting eating activities by measuring deformations in the ear canal walls due to mastication activity. These deformations are measured with three infrared proximity sensors encapsulated in an off-the-shelf earpiece. To evaluate our method, we conducted a user study in a lab setting where 20 participants were asked to perform eating and non-eating activities. A user dependent analysis demonstrated that eating could be detected with 95.3% accuracy. This result indicates that proximity sensing offers an alternative to acoustic and inertial sensing in eating detection while providing benefits in terms of privacy and robustness to noise.

Author Keywords

Eating Detection; Activity Recognition; Dietary Intake; Proximity Sensors

ACM Classification Keywords

H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

INTRODUCTION

Obesity is a global health problem that is known to be a risk factor for a number of chronic diseases such as diabetes and coronary heart disease. Since 1980, the incidence of obesity has more than doubled, and in 2014, 11% of men and 15% of women aged 18 years and older were considered obese [2].

One of the most effective methods for combating obesity is through dietary self monitoring, which has been shown to correlate positively with weight change. However, long-term adherence to food journaling is poor. This trend is due to several factors, from the inherent difficulty in logging meals and snacks with current self-tracking tools to individuals simply forgetting to log consistently [4]. In this context, automatically identifying when eating activities are taking place is seen as an important mechanism by which the food journaling task can be facilitated.

One difficulty with wearable monitors is compliance. Users do not want to call attention to themselves by standing out in public. Embedding eating detection in a socially-acceptable

device, such as headphones or earbuds, might alleviate these concerns and lead to more sustained usage.

In this paper, we investigate the use of infrared proximity sensors mounted in a non-occluding earphone to monitor lower jaw movements associated with eating. The system, called the Outer Ear Interface (OEI), is shown in Figure 1. OEI measures the magnitude of deformations in the ear canal during speech and mastication activities.

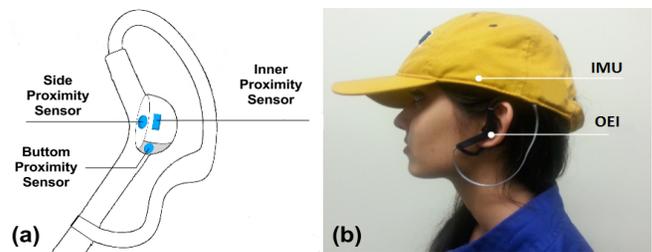


Figure 1. The current Outer Ear Interface system

RELATED WORK

Amft et al. showed that food intake detection and food type classification can be done acoustically for sufficiently “crunchy” foods, with accuracies of 99% for recognition of intake and between 80% to 100% for estimating food type [3]. Fontana et al. demonstrated a wearable sensor system for food intake detection that combines data from a jaw motion sensor, a hand gesture sensor and an accelerometer. Average accuracy of 89.8% was achieved [5]. In the more general case of detecting facial gestures, Kazuhiro et al. developed a device that utilizes optical sensors placed in the outer ear. [1].

Proximity sensing represents a promising direction for detecting eating and possibly other facial activities since it requires less processing power, is robust to noise, and imposes no privacy concerns when compared with auditory sensing.

SYSTEM DESCRIPTION

The OEI interface exploits proximity sensing to identify mandible location by measuring the volume of deformation it causes in the ear canal wall. Due to the heterogeneity of the outer ear phenotype among individuals, it is difficult to recover the mandible location and motion with one proximity sensor. As shown in Figure 1-a, OEI utilize a set of three infrared proximity sensors placed orthogonally with respect to each other to allow for a wider coverage of the ear canal. As depicted in Figure 1-b the hardware for eating detection contains an OEI unit and an Inertial Measuring Unit (IMU) inside the hat for monitoring body movement. Data from the proximity sensors and the 3D gyroscope are sampled at 100 Hz and are saved in an SD card for off-line analysis.

SYSTEM EVALUATION

We conducted a user study in our laboratory to examine how well our system distinguished between eating and non-eating activities. We recruited a total of 23 participants (10 females and 13 males) with an average age of 24 years. Due to data collection errors, we had to discard data from 3 participants.

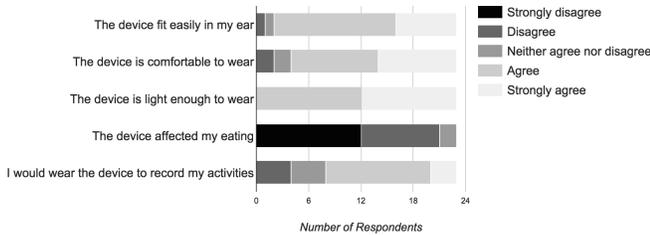


Figure 2. Laboratory study user survey results

Participants were instructed to begin the study by reading aloud from a book for 5 minutes, followed by another 5 minutes of silent book reading or internet browsing. Next, participants were asked to eat a variety of food items. They were asked to behave as they normally would during eating and were given the freedom to talk or carry other activities while they ate. The set of food items offered was controlled and included an apple, a banana and M&Ms. Eating typically lasted between 10 and 15 minutes, and participants were also asked to drink a glass of water. Afterwards, participants took a 5-minute walk outside the laboratory in a noisy city environment accompanied by one of the experimenters. Except for the walking activity, all activities were video-recorded; the videos were used for annotation and ground truth estimation.

After the study we asked the participants to fill in a survey in which they answered questions regarding the comfort and fit of the device (Figure 2).

ANALYSIS AND RESULTS

For eating activity detection we trained a set of hidden Markov models using five features extracted from the three proximity sensors and the 3D gyroscope data. Using empirical methods we selected a 10-state, left-right topology with no skip states to model “talking” and “silent,” and an 8-state, left-right topology with additional transitions between (4 → 1) and (7 → 5) to model “eating” and “walking.”

To extract features from the proximity sensors we applied an averaging window filter with a size of 5 samples to smooth the data. We then took the first derivative of the values to eliminate the offset effect on the data. To recover jaw motion from the three proximity sensors, we applied principal component analysis on the preprocessed data and used the first principle component as the first feature. We then applied a Short Time Fourier Transform on a sliding window of 50 samples on the first feature and used the average energy of the band between 1.2 to 4.6 Hz as our second feature. We found this feature helpful in distinguishing between mastication and talking activities. To help classifying walking, we included the x, y, and z gyroscope raw data as features.

To evaluate the system in the lab environment, we ran user dependent tests. We extracted frames from labeled data using

		Prediction			
		Eating	Silent	Talking	Walking
Truth	Eating	7692	97	179	260
	Silent	133	3551	180	51
	Talking	347	87	3767	114
	Walking	60	27	22	4413

Table 1. Confusion matrix for user dependent testing indicating number of frames tested and prediction results

15-second sliding windows with 50% overlap. We performed 10-fold (33%-66%) cross validation on each user. The average accuracy obtained was 95% (precision 93%, recall 96%).

DISCUSSION

From the results obtained on the frame level testing we believe there is high potential in using the current setup to achieve a low false positive rate per hour, making the system more convenient for food logging. Results from the user survey indicate that users found the system light-weight and comfortable. The confusion matrix in Table 1, shows that the current system is able to classify all classes with above 95% accuracy. This result indicates the capability of the current setup in classifying speech, silence and walking activities. Since our approach only monitors jaw motion, it might impose limitations on its capability to classify food types. However, our approach might prove complementary with the food classifying systems developed by Amft et al.

CONCLUSION AND FUTURE WORK

This work shows the feasibility of proximity sensing in the ear canal for eating detection. We believe this work could be an important stepping stone towards building systems for nutrition monitoring by prompting individuals to log their eating activity every time it is detected.

In the future we would like to run event level analysis on the lab dataset and test the real-time viability of the apparatus in the wild to evaluate the ecological validity of the system.

REFERENCES

1. Earable, commercial product. <http://www.earable.jp/>. Accessed: 2015-July-7.
2. *Global status report on noncommunicable diseases*. WHO, 2014. <http://www.who.int/nmh/publications/ncd-status-report-2014/en/>.
3. Amft, O., Stäger, M., Lukowicz, P., and Tröster, G. Analysis of chewing sounds for dietary monitoring. In *UbiComp 2005: Ubiquitous Computing*, Springer (2005), 56–72.
4. Cordeiro, F., Epstein, D. A., Thomaz, E., Bales, E., Jagannathan, A. K., Abowd, G. D., and Fogarty, J. Barriers and negative nudges: Exploring challenges in food journaling. *Proceedings of the ACM Conference on Human Factors in Computing Systems* (2015), 1159–1162.
5. Fontana, J. M., Farooq, M., and Sazonov, E. Automatic Ingestion Monitor: A Novel Wearable Device for Monitoring of Ingestive Behavior. *IEEE Transactions on Biomedical Engineering* 61, 6 (2014), 1772–1779.