Your Glasses Know Your Diet: Dietary Monitoring Using Electromyography Sensors

Qianyi Huang, *Student Member, IEEE*, Wei Wang, *Member, IEEE*, and Qian Zhang, *Fellow, IEEE*

***Abstract*—Dietary monitoring can provide valuable information for disease diagnosis, body weight control, and dietary habit management, and thus it is welcomed by patients, dieters, and nutritionists. While various techniques have been used for dietary monitoring in clinical trials and user studies, they are not ready for daily use. Existing solutions either require tedious manual recording or may impede normal daily activities. In this paper, a pair of diet-aware glasses is designed. The key idea here is that when people wear glasses, the temples of the glasses are in touch with the lower part of the temporalis muscle, one of the mastication muscles. By integrating an electromyography (EMG) sensor into glasses, the glasses can measure the muscle activity of the temporalis to detect intake-related events. This paper instantiates the idea by building a prototype equipped with an EMG sensor, a microcontroller, SD shield/card and a Bluetooth radio. When working together with a smartphone, the glasses can provide detailed information on intake schedule, the number of chewing cycles and broad food category. Extensive experiments are conducted on seven subjects and the results show appealing prospects for our diet-aware glasses. The prototype achieves 96% accuracy for counting the number of chewing cycles and up to 90.8% accuracy for classifying five types of food.**

***Index Terms*—Dietary monitoring, electromyography (EMG), glasses.**

I. Introduction

THE MODERN lifestyle, with its heavy workload and excessive traveling, is altering people's dietary habits. On one hand, people tend to skip meals or eat irregularly when tied up in meetings or hurrying to meet a deadline; on the other hand, the stressful life leads to emotional eating, as people turn to food not for hunger but for comfort and stress relief. In addition to disordered eating, a fast-paced life also shortens the mealtime. A survey by Malakoff Médéric, an insurance company, observed that the average time for the French workers to have lunch had shrunk from 90 min 20 years ago to 22 min nowadays [1]. People rush through meals without chewing the food thoroughly. Researchers discovered that eating quickly and irregularly is associated with being overweight, increasing the risk of developing cancer, cardiovascular diseases and diabetes [10].

Dietary monitoring can provide personal and detailed information on various aspects (e.g., intake schedule, food amount, and meal composition), which is of great help to prevent and intervene with unhealthy dietary habits. Long-term dietary monitoring can also provide valuable information for chronic disease diagnosis. Thus dietary monitoring technology is gaining popularity. Dietary logging is a built-in function in some commercial wearable devices (e.g., Fitbit and Jawbone Up) and some smartphone apps (e.g., Cholesterol Manager and VisualFoodLog). However, they require manual logging, such as taking a photo of the food or scanning a barcode. Manual logging requires a considerable amount of labor from the user, which makes it unpractical for long-term use. In the research community, scientists advocated the use of on-body sensors to automatically monitor diet. There is a number of works that use a microphone to record the sounds of crushing and grinding food (e.g., [2] and [16]), which can guarantee accurate and unobtrusive monitoring. However, for some soft foods (e.g., banana and cooked rice), the sounds of chewing are too weak for identification. Other solutions make use of surrounding objects and environment, such as the smart dining table [4] and the smart kitchen [6], where monitoring scope is limited to specific locations. These solutions do not fit well into the modern lifestyle, where people eat at various places throughout the day. To fill in this gap, in this paper, a wearable, unobtrusive and reliable dietary monitoring system is proposed, which does not require users' manual logging and has no invasion of privacy.

There is an interesting observation that, by integrating electromyography (EMG) sensors into the arms of eyeglasses, daily dietary can be monitored with high accuracy in a comfortable manner. Glasses are necessaries for many people. According to the statistics from the Vision Council of America, approximately 64% of adults wear eyeglasses [20]. Moreover, besides traditional eyeglasses, smart glasses have seen a consistent increase in popularity [17], [23]. As unobtrusive and comfortable devices, glasses are popular wearable gadgets which are suitable for daily dietary monitoring.

The physiological principle behind the observation is that when people wear glasses, the lower part of temporalis muscle is exactly the place where glasses touch the skull.

Manuscript received October 12, 2016; revised December 15, 2016; accepted December 30, 2016. Date of publication January 20, 2017; date of current version June 15, 2017. This work was supported in part by the 973 Project under Grant 2013CB329006, in part by the China NSFC under Grant 61502114, in part by the RGC under Contract CERG M-HKUST609/13, Contract 622613, Contract 16212714, Contract 16203215, and Contract ITS/143/14FP-A, in part by the Huawei-HKUST Joint Laboratory, and in part by the IoT WiFi Key Technologies.

Q. Huang and Q. Zhang are with the Department of Computer Science and Engineering, Hong Kong University of Science and Technology, Hong Kong (e-mail: qhuangaa@cse.ust.hk; qianzh@cse.ust.hk).

W. Wang is with the School of Electronic Information and Communications, Huazhong University of Science and Technology, Wuhan 430074, China, and also with the Fok Ying Tung Research Institute, Hong Kong University of Science and Technology, Hong Kong (e-mail: weiwangw@hust.edu.cn).

Digital Object Identifier 10.1109/JIOT.2017.2656151

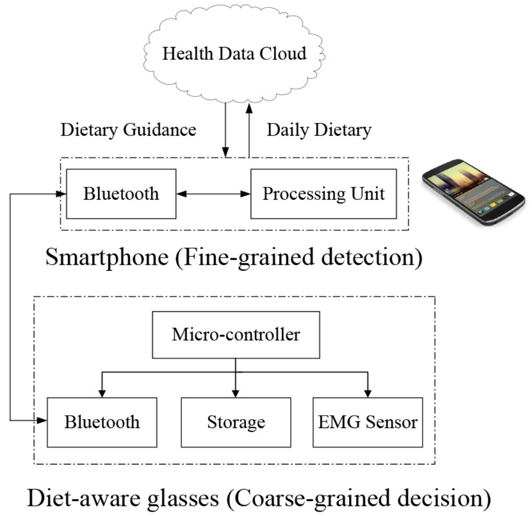

Fig. 1. Glasses measure temporalis EMG signals and perform coarse-grained decision. The glasses send relevant data samples to the smartphone for fine-grained detection. The smartphone will also upload the dietary information to the health data cloud. Combining users' exercise and sleep information, the cloud will feedback dietary guidance to the users.

The temporalis is one of the mastication muscles. Its major action is to elevate the lower jaw, supporting the chewing action [8]. We can feel the temporalis contracting when clenching the teeth, which indicates the potential of using EMG sensors to monitor mastication and thereby to monitor dietary related behavior. Furthermore, compared with existing solutions based on audio or video recording (e.g., [21] and [22]), the EMG-based solution will cause much less privacy breach.

Although the idea sounds straightforward, there are several challenges. First, existing techniques for surface EMG signal acquisition require the usage of adhesive electrodes, which are inconvenient for daily use. Moreover, the natural shifting of glasses due to head movements makes it even harder for signal acquisition. To cope with low-quality signals, we use adaptive thresholding for chewing spotting. Second, despite mastication, there are other activities involving jaw movements (e.g., talking and laughing). To distinguish mastication from other similar activities, we observe that the mastication of a bite of food usually involves a sequence of chews. This repetitive characteristic makes it different from other activities, e.g., talking. Last, as a wearable device, it has limited battery life and storage space. Transmitting or storing all sensor data would quickly drain the battery and use up the memory. Only data containing potential durations of food intake are of interest. Thus, a real-time algorithm is in demand to decide whether to keep or discard the data.

In this paper, a glass-based dietary monitoring system is proposed, which is wearable, unobtrusive and reliable. As shown in Fig. 1, the system contains three components: 1) the wearable glasses; 2) a central processor (e.g., smartphone); and 3) the health data cloud. The glasses provide the capability for EMG sensing, on-board storage and wireless communication with the central processor. The central processor is responsible for the majority of data processing. By analyzing the EMG signals, it can provide detailed information on intake schedule and the number of chewing cycles, which are of great value for dietary habits management. There is also an initial attempt to monitor food category. The central processor will also upload daily dietary information to the health data cloud. Combining the detailed dietary information with users' exercise and sleep information, the cloud can provide users with dietary guidance. Extensive experiments are conducted on seven volunteers with five different types of food. These seven volunteers are specifically selected to cover a wide BMI range, from 19 (normal weight) to 32 (obesity). Experiment results show that we can achieve 96% accuracy for detecting chewing cycles and up to 90.8% accuracy for classifying five types of known food. There is also a discussion that explores the feasibility of food classification in daily use.

II. Motivation and System Overview

A. Motivation for Dietary Monitoring

Of the three key healthy lifestyle activities (exercise, sleep, and eat), dietary monitoring is the most difficult to accomplish, but also very important. The motivation for dietary monitoring is to provide detailed and personal feedback on intake-related behavior, so as to prevent and intervene with unhealthy eating habits. When the dietary information is further combined with exercise and sleep log, it can provide the users with accurate and personalized dietary guidance. The functionality can also be integrated into remote healthcare systems, so as to help doctors monitor patients remotely.

In this paper, the focus is on the following dimension of dietary monitoring. The first one is intake schedule. People may skip meals when overwhelmed by work or eat excessively for comfort and stress relief. Intake schedule is about recording the time when intake events take place in daily routines. It also reflects the duration of intake events. The second is the number of chewing cycles per bite. The fast-paced life makes it a common scenario that people are in a hurry, eating on the run and swallowing the food without chewing thoroughly. Eating too fast is a health hazard, setting the body up for the development of obesity. Monitoring the number of chewing cycles can detect whether the user chews the food properly and send an alert when he/she is eating too fast. As chewing is the most fundamental activity during food intake, monitoring chewing events over daily routines can generally reflect intake schedule. Last but not the least, there is an initial attempt to monitor food category. Food category describes the composition of diet and reveals the nutrition information, which is an important factor in a healthy diet. Although the glasses can only identify food category among a small set of predefined foods, Section VI discusses how it can be used in daily life.

B. System Architecture and Implementation

There are three components in the dietary monitoring system, the diet-aware glasses, a central processor (e.g., smartphone), and the health data cloud, as shown in Fig. 1. The user is wearing the glasses when he/she goes about daily life. The central processor is responsible for the majority of data processing and provides feedback to the user. These two devices communicate via Bluetooth. The smartphone will

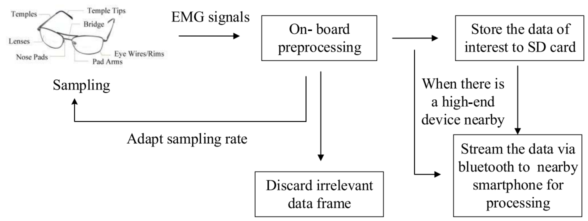

Fig. 2. Overview of system design.

Fig. 3. Prototype.

upload users' dietary information to the health data cloud, on which stores users' health-related data, including their exercise and sleep log. By combining all the available and related information, the cloud can provide users with accurate and personalized dietary guidance.

There is an EMG sensor, a microcontroller, a Bluetooth radio, an SD breakout shield and an SD card mounted on the glasses. EMG signals are continuously sampled at a certain frequency. As shown in Fig. 2, the glasses have two modes, idle mode and working mode. By default, the glasses are in idle mode. The EMG sensor samples at a low rate (i.e., 25 Hz). Based on the sensor data, the glasses decide whether the user is probably eating or not. The decision is made by comparing the signal variance to an adaptive threshold. Irreverent data frames are discarded. When detecting that the user is likely eating something, the glasses would switch to the working mode. The sensor would sample at a high frequency (i.e., 50 Hz) and transmit the signals to the smartphone for more fine-grained data processing. If the smartphone is not nearby, it can temporarily store sensor data to the on-board SD card. The detailed algorithm is presented in Section III-A.

The prototype is shown in Fig. 3. For all the components to be mounted on the temple of glasses, the selected hardware are thin and lightweight. The EMG sensor adopted is the Muscle Sensor from Advancer Technologies. The two data channels of the cable adhere to the places where the arms of the glasses are in touch with the skull. The feedback channel adheres to the temple tip on the left side. Bluno nano is an off-the-shelf product made by DFRobot. It integrates Bluetooth (version 4.0) with Arduino UNO development board and comes in a size of a gum. It has a 16 MHz ATmega328 microcontroller with 32 kB program space and 2 kB SRAM for variables. The ADC converts an input voltage between 0 and 5 volts into an integer value in 0–1023. On-board storage is provided by an SD card breakout shield and a 16 GB SD card.

III. Chewing Detection

Chewing is the fundamental activity during food intake. Chewing event detection can provide fine-grained information on the number of chewing cycles and it can also generally reflect intake schedule. The algorithm for chewing detection mainly contains two parts: 1) on-board real-time coarse-grained decision and 2) the real-time chewing detection on the central processor.

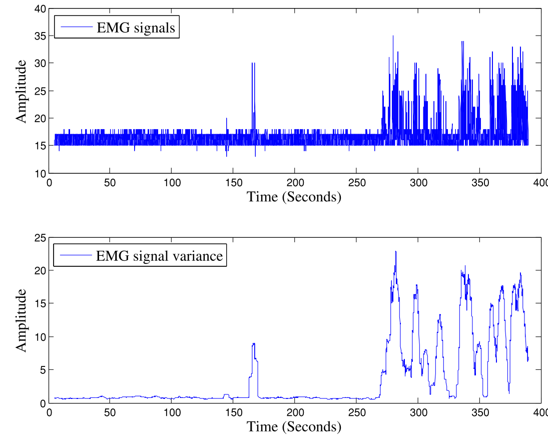

Fig. 4. Data samples and data variance.

A. On-Board Real-Time Decision

The goal of the on-board algorithm is to find possible intervals containing eating periods.

To study the distinction of EMG signals between eating and noneating periods, we show a sequence of 400-s EMG signal samples in Fig. 4 containing both eating and noneating periods. For the first 150 s, the subject is talking; from 150 to 270 s, the subject is walking and looking around; the subject starts eating at 270 s. Fig. 4 (top) shows the raw data while the bottom one shows the data variation. Let $R = \{r_1, r_2, \ldots, r_n\}$ denote the time series of raw samples. The data variance Var is defined as

$$\mathrm{Var} = 1/n \times \sum_{i=1}^{n} (r_i - \bar{r})^2$$

where $\bar{r}$ is the mean of R.

For the first 150 s, the Var is about 1. After the subject starts eating at 270 s, Var increases to 5 and above. It indicates that Var may serve as criteria for differentiating eating and noneating periods. The signal fluctuation at around 165 s results from the subject turning her head to the left. The fine-grained detection algorithm in Section III-B is designed to eliminate such false positives.

A set of feasibility studies is conducted to validate above observation.

1) Feasibility Study: We invite three subjects to participate in the feasibility studies, including one female and two males. They are all college students. Continuous EMG signals are collected from three subjects while they are performing various daily activities, such as eating different types of food (e.g., crackers, potato chips, and corn), talking, looking around, and working on computers. The whole process is captured by a camera. The overall study lasts for about 112 min. The data are manually labeled for each second.

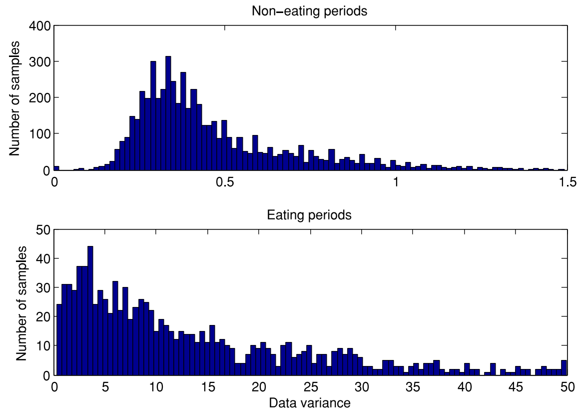

Fig. 5. Histograms of data variance.

Algorithm 1 On-Board Real-Time Decision Algorithm

Initialization:
Variance threshold $V \leftarrow T$; Sampling frequency $f \leftarrow F_i$; $Flag \leftarrow Flase$.
while True **do**
Wait until sensor data for current second available and calculate Var.
if $Var > V$ **then**
if $Flag == False$ **then**
$f \leftarrow F_w$; $Flag \leftarrow True$;
end if
$V \leftarrow max(V - \Delta, L)$.
Call data processing procedure.
else
if $Flag == True$ **then**
$f \leftarrow F_i$; $Flag \leftarrow Flase$; $V \leftarrow T$.
end if
Discard the data.
end if
end while

The whole process contains 1118 s of eating periods and 5645 s of noneating periods. Fig. 5 shows the distributions of data variance for eating periods and noneating periods. Fig. 5 indicates that the data variance for almost 97% of noneating periods and only 5% of eating periods are less than 1.5. Thus, data variance can serve as an effective rule for roughly deciding whether the user is probably eating or not.

Data variance is a simple criteria which can be calculated efficiently on the microcontroller. A threshold is set to separate eating periods and noneating periods, which is a commonly used method in EMG signal analysis [18]. It is natural that people chew a bite multiple times before swallowing, and thus if the user is chewing in the previous second, he/she is likely chewing in the current second. To reduce false-negative, the algorithm will adjust the threshold accordingly. The detailed pseudo-code is shown in Algorithm 1. Here, T is the initial variance threshold and Δ is the threshold adjustment step, while L is the lower bound for the threshold; F_i, F_w are the sampling frequency for idle mode and working mode, respectively; $Flag$ is a boolean variable to store the decision results for the previous second.

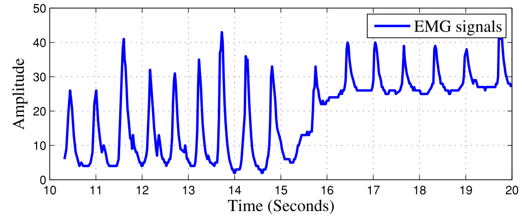

Fig. 6. Data samples.

The data processing procedure is automatically switched between two modes. When the smartphone is nearby, the glasses will continuously stream the sensor data to the smartphone for real-time processing; otherwise, data are temporarily stored to the on-board SD card and synchronized with the smartphone when possible.

B. Fine-Grained Chewing Detection

After raw data are synchronized with the central processor, the fine-grained chewing detection algorithm is executed on the central processor.

Although there are a lot of factors affecting EMG signals, there is a simple pattern. As the EMG signals are rectified and amplified by the Muscle Sensor, there is a pulse in the signals every time the user chews. Fig. 6 shows a 10 s' signals while the subject is having blueberries. Although the baseline drifts over this short time span, clearly, there is one pulse corresponding to each chew. Furthermore, chews occur repeatedly in sequence, which are quite different from other activities. The algorithm applies some data processing techniques to filter out high-frequency noise and adds some constraints to reduce false positives. The techniques are illustrated step by step as follows.

1) Preprocessing: In daily life, mastication is a low-frequency activity which happens at most two to three times in 1 s. Thus, the algorithm uses the Savitzky–Golay algorithm to smooth the EMG recordings of potential intake intervals, so as to filter out high-frequency noise. The filter width is set to be 5.

2) Pulse Identification: There are ups and downs in the smoothed signals, some of which correspond to chewing, while others may be noise or motion artifact. The following constraints are used to remove false-positives.

1) $P > P_{\min}$.
2) $D_{\min} < D < D_{\max}$.
3) $C > C_{\min}$.

Here P, D, and C denote peak-to-baseline amplitude, duration of a chewing cycle and the number of cycles in a chewing sequence, respectively. The explanations for each constraint are as follows.

The first and most straightforward constraint is the peak-to-baseline amplitude. A low threshold will result in false-positives and a high threshold will bring adverse effect. Due to the inherent instability of EMG signals, the signal amplitudes vary significantly among different users and on different days. To find a robust threshold, analysis is conducted among the 1600 chews from three subjects in the feasibility study. Fig. 7 plots the ratio of peak-to-baseline amplitude to data variance.

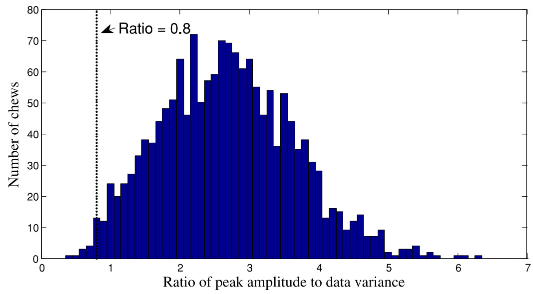

Fig. 7. Distribution of peak-to-baseline amplitude to data variance ratio.

Nearly, all the ratios are larger than 0.8. Thus, in evaluation, we set, that is,

$$P_{\min} = 0.8 \times \text{Var}.$$

The physical meaning behind this adaptive threshold is that, when the signals have larger variance, possibly the background noise level is higher, thus the peak-to-baseline amplitude should be larger so as to be recognized as chewing, and vice versa.

The second constraint is the duration of a chewing cycle. According to daily experience, the duration of the chewing cycle lasts from several hundred milliseconds to several seconds. Signal fluctuations with larger durations are likely due to other body movements, such as head movement. According to a previous study [15], the average cycle duration is 0.67 s for carrot and peanut, 0.77 s for cheese, and 0.82 s for cake. Thus, to be conservative, the thresholds are set as

$$D_{\min} = 0.2 \text{ s}, D_{\max} = 2 \text{ s}.$$

Another distinction between chewing and other activities is that chews are more likely to occur repeatedly in sequence. Even for a small bite, people have to chew multiple times before swallowing. One or two isolated pulses are likely noise or motion artifact. Here, the threshold is set as follows:

$$C_{\min} = 2.$$

As all the above algorithms leverage only the historical data and the data on the fly, it can be implemented as a real-time algorithm. When the sensing data is continuously transmitting to the central processor, the smartphone can provide feedback to the user in real time. When it detects that the user is not chewing the food thoroughly, it may remind the user to slow down and enjoy the food.

IV. Food Classification

Meal composition is an important dimension of dietary monitoring, which is of great value for chronic disease prevention and diagnosis. The goal of this paper is to recognize the broad food category. We discuss how to apply this functionality into practical use in Section VI.

A decision tree-based classifier is used for classification, for it is simple and easy to interpret. The following four sets of features are extracted from each single chew.

1) *P (Peak-to-Baseline Amplitude):* Existing study [11] shows that the electrical activity level of mastication muscles increases with the hard level of food. Peak-to-baseline amplitudes are used to capture the activity level of temporal muscle.
2) *D (Chewing Cycle Duration):* A previous study [12] indicates that the duration of jaw-closing muscle activity is closely related to the food fracturability and food adhesiveness.
3) *C (The Number of Chewing Cycles in the Sequence):* Similar to duration, the number of chewing cycles is also related to food textures. A previous study [12] points out that increase in both food fracturability and adhesiveness would decrease the number of chewing cycles.
4) T_{25}, T_{50}, T_{75}: T_p values are proposed by Miyaoka *et al.* [13] to quantify muscles' activity pattern. The duration of muscle contraction is standardized to 1.0. T_p corresponds to the time on the standardized time scale when the cumulative EMG reaches $p\%$ percent of its maximum value. Miyaoka *et al.* [12] have shown that T_p values of masseter muscle are effective in discriminating food with different textures. Similar to [12], T_{25}, T_{50}, and T_{75} are used as features.

For a chewing sequence, chews in the sequence may be classified with different labels. Since food content is unlikely to change within a bite, similar to [2], a majority vote is performed over a chewing sequence to decide on its food content. Evaluation results show that this step can improve the overall classification accuracy.

V. Performance Evaluation

This section describes the experimental settings and shows the evaluation results. Performance of the prototype is evaluated from four aspects. First is the accuracy of counting the number of chewing cycles. Then is the ability of the prototype to distinguish intake events from other daily activities. After that are the results of food classification. Finally is the analysis of the system energy consumption. We note that data reported in this section are independent of the data used for empirical algorithm design in Section III.

A. Experimental Settings

An iPhone 5s (16 GB) is used as the central processor, running on 1.3 GHz dual-core Apple A7 (64-bit ARMv8) processor with 1 GB of RAM. During experiments, sensor data are continuously streaming from the glasses to the smartphone via Bluetooth. Data are analyzed offline. All data are collected from a normal laboratory environment without any electromagnetic shielding unless otherwise specified. There is no requirement for skin preparation before the experiments.

Seven volunteers, one female and six males, are invited to participate in the experiments. They are all college students and staff aging from 20 to 30 years old. They are specifically selected to cover both slim people and obese people (BMI ranges from 19 to 32). While slim people wear glasses loosely, stout people wear them tightly. We want to see whether such a factor would affect experimental results. As it is infeasible to cover thousands of food types in experiments, we select these

TABLE I
CHEWING EVENT DETECTION

Subject No.		1	2	3	4	Rate
Apple	Truth	297	303	N/A	258	
	Result	293	320	N/A	276	
	FP	1	17	N/A	18	0.04
	FN	5	0	N/A	0	0.01
Banana	Truth	124	82	148	90	
	Result	119	80	146	84	
	FP	1	0	0	2	0.01
	FN	6	2	2	8	0.04
Blue berries	Truth	208	264	208	166	
	Result	231	265	213	155	
	FP	23	1	5	0	0.03
	FN	0	0	0	11	0.01
Bread	Truth	624	307	777	494	
	Result	657	312	791	510	
	FP	33	5	14	16	0.03
	FN	0	0	0	0	0
Crackers	Truth	179	262	237	277	
	Result	170	279	239	288	
	FP	1	17	2	11	0.03
	FN	10	0	0	0	0.01
Average FP Rate		0.04	0.03	0.02	0.05	0.03
Average FN Rate		0.01	0	0	0.02	0.01

five types of food with different food textures and shapes, i.e., apples, bananas, bread, blueberries, and crackers. Apples are wet and crispy whereas crackers are dry and crispy; bananas and bread are both soft texture foods, but bananas are wet whereas bread is dry. Blueberries are watery and small, which can be eaten in one gulp. The whole process is recorded by a camera. The chewing ground truth is acquired by manual labeling based on video footage.

B. Chewing Event Detection

This section shows the results for chewing event detection. Subjects 1–4 (BMI ranges from 19 to 32) participate in this part of evaluation. Subjects are seated during the process. They are instructed to chew the food as naturally as possible and at their preferred speeds. They may talk and drink between bites, look around and have some unconscious body movements. The experiment for each subject lasts about 40 min. The results are shown in Table I. Here FP is short for false positive and FN is short for false negative.

The overall false positive rate is 0.03 and false negative rate is 0.01. Note that for the majority of foods (except for banana) and test subjects, false positive rates are higher than false negative rates. The possible underlying reason is that people may chew unconsciously after swallowing a bite, to crush some small food debris. These tiny actions are difficult to identify from videos but the jaw movements are recorded by the EMG sensor. Thus these unconscious bites are not counted into *Truth*, but the algorithm identifies them. When eating bananas, which are very soft, subjects bite gently and the jaws move in a mild manner. Hence, the glasses miss a number of bites. The overall accuracy is 96%. The performance is stable regardless of the slimness or obesity of the subjects. From above results, it can be inferred that it is unlikely for the glasses to miss intake events that last for many minutes, such as big meals.

TABLE II
DAILY ACTIVITY MONITORING

Subject No.	Total Time	Proportion	No. of Detected Chews
1	105 min 27 sec	0.02	2
2	22 min 21 sec	0.04	0
3	14 min 09 sec	0.03	0
4	20 min 47 sec	0.01	0
5	14 min 0 sec	0.04	4
6	13 min 29 sec	0	0
7	13 min 30 sec	0.03	5

Thus, it can also generally reflect the user's intake schedule and duration of eating periods.

C. Daily Activity Monitoring

This set of experiments is conducted to evaluate the performance in differentiating between chewing and other activities. We consider three categories of head/body activities.

1) *Head Activity:* Talk, laugh, yawn, sneeze, cough, look around, turn the head to left and right, head lean against right/left shoulder, and head nodding/shaking.
2) *Deskwork:* Subjects are seated and they read books or work on computers.
3) *Body Activity:* Walk around.

The subjects were actively told to perform these confounding behaviors. Subject 1 participates in a two-hour experiment to cover different daily activities. The experiment is conducted in the living room of her apartment. The familiar environment makes it easier for her to behave naturally. She covered nearly all the activities listed above. The remaining six subjects participate in a 15-min experiment to verify the robustness of the algorithm across different users. Subjects 2–4 mainly covered head activity and deskwork. Subjects 5–7 mainly covered head activity and body activity. Table II gives the results.

Any chew detected here are all false positives. For subjects 2–4 and 6, there is no false positive. Overall, the false positive rate is very low. Although some activities (e.g., speak and laugh) also involve jaw movement, different from eating, they do not have a similar pattern that occurs repeatedly. Table II also shows the proportions of transmitted data in the whole experiments. From the results, it indicates that the on-board decision algorithm can significantly reduce the data volume to be transmitted or stored. Thus, it can save a large amount of on-board storage and energy for wireless communication.

D. Food Classification

This section presents the results of food classification among a small set of predefined foods. Section VI shows the feasibility for more wide use. Five-thousand three-hundred sixty-six single chews are collected in total. $1/5$ of the samples are randomly selected as the training set and the remaining $4/5$ samples as the test set. We use the J48 decision tree classifier from Weka. Then, a majority vote is performed over each chewing sequence. We use the time intervals between chewing sequences as indicators for segmentation, during which the user would sallow the food, maneuver the remaining food to the mouth, and take another bite. The time interval between

TABLE III
CONFUSION MATRIX FOR CHEWING SEQUENCES

a	b	c	d	e	← Classified as	Accuracy
37	6	4	1	2	a = Apple	74.0%
2	18	3	0	3	b = Banana	69.2%
2	3	69	0	2	c = Blue berries	90.8%
1	2	5	48	3	d = Bread	81.4%
4	2	3	3	29	e = Crackers	70.7%

two chewing sequences is usually larger than the gap between two consecutive chews. The final results are shown in Table III.

The accuracy for blueberries is the highest among all types of foods, for it requires the least amount of chewing cycles. When the product finally comes in a form-factor of glasses, we can further combine vision-based solutions to improve classification accuracy.

E. Energy Consumption

The energy consumption is mainly attributed to the EMG sensor, on-board processing and control by the microcontroller, wireless communication by the Bluetooth chip and file read/write operations by the SD breakout shield.

As the Bluetooth chip is integrated into Bluno nano and the SD shield is powered by the Bluno nano, we do a two-way breakdown of power measurement. The power consumption for the Muscle Sensor and Bluno nano is 33.4 and 138.3 mW, respectively. The prototype directly adopts the off-the-shelf products and does not have any hardware revision for the consideration of energy-saving. By adopting MSP430 series, the ultralow power processor from Texas Instruments, the power consumption for on-board microcontroller can be reduced to 0.3 mW. Furthermore, the latest Bluetooth Low Energy technology brings down the communication power consumption to 15 mW [7]. By integrating these latest ultralow power components, the power consumption of the diet-aware glasses can be greatly reduced.

VI. DISCUSSION

This section discusses some remaining issues and possible future directions.

The goal of this paper is to record the broad food category. Previous evaluation shows the classification results for five types of food. The accuracy ranges from 69.2% to 90.8%. The food classification capability is restrained by the limitation of EMG. EMG only records the muscle contraction information, which may be used to reveal the food texture, e.g., hard/soft, crispy/soft. It is challenging to recognize the exact food type, such as distinguishing noodles and rice.

Although it cannot recognize the exact food type, it is feasible for more wide use with some domain knowledge. For instance, combining electronic shopping list or credit card payment information, it can narrow down possible food categories. Context information, such as dining at western restaurants or Chinese restaurants, is also helpful for narrowing down the potential types. It may further combine video-based image processing techniques to improve food classification accuracy. When identifying foods from images, the knowledge of food textures can help to classify foods that have similar colors and shapes, such as apples (crispy) and tomatoes (soft).

In addition to monitoring food category, the initial efforts to distinguish basic food textures can benefit a number of applications. For example, some patients are prone to indigestion. When they are having some hard-texture foods, the glasses may remind them to chew the food thoroughly before swallowing.

Another limitation of our diet-aware glasses is that it cannot distinguish healthy food and unhealthy food. Doctors are more concerned about whether user's diet is healthy or not. However, the glass cannot distinguish foods with similar texture, such as chips and carrot slices. One possible solution is to invoke the camera periodically or at the first bite (the glasses can detect that the food texture is different from the last bite, i.e., the user has switched to another type of food).

VII. RELATED WORK

We review existing works in dietary monitoring.

One major line of research effort is devoted to wearable sensors. During meals, there are some specific intake gestures, such as fine-cutting and maneuvering food pieces to the mouth [3]. Leveraging inertial sensors, intake gestures can be recognized. During mastication, sounds of crushing and grinding food propagate through bones and tissue, which can be picked up by a microphone at the outer ear canal. There are a number of works belonging to this category (e.g., [2], [14], and [19]). Chewing sounds provide useful information on the eating microstructure and food category. However, for some soft foods (e.g., cooked rice), the chewing sounds are too weak for identification [2]. Cheng *et al.* [5] used capacitive sensors to detect swallowing events, which can be integrated into collars. Cheng *et al.* [5] were able to differentiate between swallowing 5 and 15 ml of water, shedding light on food amount estimation. Li *et al.* [9] proposed a prototype for embedding a small motion sensor inside artificial teeth for classification of oral activities, such as speaking, chewing and drinking. Muscle cells generate some electrical signals during normal functioning. EMG is a widely used technique to measure muscle activity. Previous studies used EMG to analyze masseter and temporalis muscle activity during mastication strokes under clinical settings (e.g., [11], [12], and [15]). However, they require the subjects to wear adhesive electrodes. Different from these solutions, the diet-aware glasses, leveraging EMG sensors on the arms of glasses, provide an unobtrusive, comfortable monitoring and have no invasion of privacy. As the glasses record the muscle activity of temporal muscle, it can detect food intake events, as long as there are jaw movement and teeth clenching. Thus, it is able to detect soft-texture foods (e.g., banana). Although there is a similar work [24], it fails to consider the limited capacity of wearable devices.

Another major line of research seeks to leverage ambient sensors. Wu and Yang [21] recognized fast food in videos by comparing images to the reference images in a database, which contains 101 types of fast foods collected by the authors.

Chang *et al.* [4] proposed a smart dining table with two layers of sensor surfaces: 1) a RFID surface and 2) a weighing surface. Such ambient-based solutions are location-dependent (i.e., are constrained by specific environments), implying some limitations for monitoring dietary behavior in modern lifestyle, where people eat at various locations throughout the day. Thus, compared with ambient-based solutions, the glasses-based dietary monitoring system can provide long-term, ubiquitous monitoring.

VIII. Conclusion

In this paper, the idea of diet-aware glasses is proposed and implemented in a hardware prototype platform, which can provide long-term dietary monitoring in an unobtrusive and comfortable manner. The glasses measure the muscle activity of the temporalis and the microcontroller on-board performs a real-time coarse-grained intake detection. When working together with a smartphone, it can provide fine-grained information on intake schedule, the number of chewing cycles, as well as broad food category. The performance of the prototype is evaluated from different aspects. Experiment results from seven subjects show that it can achieve 96% accuracy for counting the number of chewing cycles and up to 90.8% accuracy for classifying five types of food. It can also distinguish intake from other daily activities. It is expected that the system opens a new direction for dietary monitoring, which is promising for wide use in daily life.

References

[1] Y. Alléno, *Bien Déjeuner Dans ma "Boîte."* Paris, France: Malakoff Médéric, 2011.

[2] O. Amft, M. Stäger, P. Lukowicz, and G. Tröster, "Analysis of chewing sounds for dietary monitoring," in *Proc. UbiComp*, Tokyo, Japan, 2005, pp. 56–72.

[3] O. Amft and G. Tröster, "On-body sensing solutions for automatic dietary monitoring," *IEEE Pervasive Comput.*, vol. 8, no. 2, pp. 62–70, Apr./Jun. 2009.

[4] K.-H. Chang *et al.*, "The diet-aware dining table: Observing dietary behaviors over a tabletop surface," in *Proc. Pervasive*, Dublin, Ireland, 2006, pp. 366–382.

[5] J. Cheng, O. Amft, and P. Lukowicz, "Active capacitive sensing: Exploring a new wearable sensing modality for activity recognition," in *Proc. Pervasive*, Helsinki, Finland, 2010, pp. 319–336.

[6] P.-Y. P. Chi, J.-H. Chen, H.-H. Chu, and J.-L. Lo, "Enabling calorie-aware cooking in a smart kitchen," in *Proc. Persuasive*, Oulu, Finland, 2008, pp. 116–127.

[7] *Bluetooth Smart*, Dialog Semicond., London, U.K., 2015. [Online]. Available: http://www.dialog-semiconductor.com/products/bluetooth-smart

[8] H.-U. Hecker and A. Steveling, *Color Atlas of Acupuncture: Body Points-Ear Points-Trigger Points*. New York, NY, USA: Thieme, 2011.

[9] C.-Y. Li, Y.-C. Chen, W.-J. Chen, P. Huang, and H.-H. Chu, "Sensor-embedded teeth for oral activity recognition," in *Proc. ISWC*, Sydney, NSW, Australia, 2013, pp. 41–44.

[10] K. Maruyama *et al.*, "The joint impact on being overweight of self reported behaviours of eating quickly and eating until full: Cross sectional survey," *Brit. Med. J.*, vol. 337, no. 7678, pp. 1091–1093, 2008.

[11] L. Mioche, P. Bourdiol, J.-F. Martin, and Y. Noël, "Variations in human masseter and temporalis muscle activity related to food texture during free and side-imposed mastication," *Arch. Oral Biol.*, vol. 44, no. 12, pp. 1005–1012, 1999.

[12] Y. Miyaoka *et al.*, "Analysis of masseter activity patterns using T_p values during chewing of foods with different shapes and textural properties," *J. Texture Stud.*, vol. 44, no. 3, pp. 196–204, 2013.

[13] Y. Miyaoka, I. Ashida, D. Inagaki, and S.-Y. Kawakami, "Differentiation of activity patterns in the suprahyoid muscles during swallowing of foods with five taste qualities," *J. Sens. Stud.*, vol. 20, no. 6, pp. 473–483, 2005.

[14] S. Päßler, M. Wolff, and W.-J. Fischer, "Food intake monitoring: An acoustical approach to automated food intake activity detection and classification of consumed food," *Physiol. Meas.*, vol. 33, no. 6, pp. 1073–1093, 2012.

[15] L. J. Pereira, M. B. D. Gavião, L. Engelen, and A. V. D. Bilt, "Mastication and swallowing: Influence of fluid addition to foods," *J. Appl. Oral Sci.*, vol. 15, no. 1, pp. 55–60, 2007.

[16] T. Rahman *et al.*, "BodyBeat: A mobile system for sensing non-speech body sounds," in *Proc. MobiSys*, Bretton Woods, NH, USA, 2014, pp. 2–13.

[17] S. Rallapalli, A. Ganesan, K. Chintalapudi, V. N. Padmanabhan, and L. Qiu, "Enabling physical analytics in retail stores using smart glasses," in *Proc. MobiCom*, Maui, HI, USA, 2014, pp. 115–126.

[18] M. B. I. Reaz, M. S. Hussain, and F. Mohd-Yasin, "Techniques of EMG signal analysis: Detection, processing, classification and applications," *Biol. Procedures Online*, vol. 8, no. 1, pp. 11–35, 2006.

[19] M. Shuzo *et al.*, "Wearable eating habit sensing system using internal body sound," *J. Adv. Mech. Design Syst. Manuf.*, vol. 4, no. 1, pp. 158–166, 2010.

[20] *VisionWatch Canada Market Overview*, Vis. Council, Alexandria, VA, USA, Jun. 2015.

[21] W. Wu and J. Yang, "Fast food recognition from videos of eating for calorie estimation," in *Proc. ICME*, New York, NY, USA, 2009, pp. 1210–1213.

[22] K. Yatani and K. N. Truong, "BodyScope: A wearable acoustic sensor for activity recognition," in *Proc. UbiComp*, Pittsburgh, PA, USA, 2012, pp. 341–350.

[23] L. Zhang *et al.*, "It starts with iGaze: Visual attention driven networking with smart glasses," in *Proc. MobiCom*, Maui, HI, USA, 2014, pp. 91–102.

[24] R. Zhang, S. Bernhart, and O. Amft, "Diet eyeglasses: Recognising food chewing using EMG and smart eyeglasses," in *Proc. BSN*, San Francisco, CA, USA, 2016, pp. 7–12.

Qianyi Huang (S'15) received the B.S. degree in computer science from Shanghai Jiao Tong University, Shanghai, China. She is currently pursuing the Ph.D. degree at the Department of Computer Science and Engineering, Hong Kong University of Science and Technology, Hong Kong.

Her current research interests include mobile computing, wearable applications, and digital health.

Wei Wang (S'10–M'16) received the bachelor's degree in electronics and information engineering from the Huazhong University of Science and Technology, Wuhan, China, in 2010, and the Ph.D. degree from the Department of Computer Science and Engineering, Hong Kong University of Science and Technology, Hong Kong.

He is currently a Professor with the School of Electronic Information and Communications, Huazhong University of Science and Technology.

Qian Zhang (M'00–SM'04–F'12) received the B.S., M.S., and Ph.D. degrees from Wuhan University, Wuhan, China, in 1994, 1996, and 1999, respectively, all in computer science.

She joined the Hong Kong University of Science and Technology, Hong Kong, in 2005, where she is a Full Professor with the Department of Computer Science and Engineering. She was a Research Manager with the Wireless and Networking Group, Microsoft Research Asia, Beijing, China, from 1999 to 2005.