Wearable Eating Habit Sensing System Using Internal Body Sound*

Masaki SHUZO**, †, Shintaro KOMORI**, Tomoko TAKASHIMA**, †
Guillaume LOPEZ**, †, Seiji TATSUTA***, †, Shintaro YANAGIMOTO****, †
Shin’ichi WARISAWA**, †, Jean-Jacques DELAUNAY**, † and Ichiro YAMADA**, †

**School of Engineering, The University of Tokyo,
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan
E-mail: shuzo@lelab.t.u-tokyo.ac.jp,
***Olympus Corporation,
2-3 Kuboyama-cho, Hachioji-shi, Tokyo 192-8512, Japan
****The University of Tokyo Hospital,
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan
†Japan Science and Technology Agency, CREST

Abstract
Continuous monitoring of eating habits could be useful in preventing lifestyle diseases such as metabolic syndrome. Conventional methods consist of self-reporting and calculating mastication frequency based on the myoelectric potential of the masseter muscle. Both these methods are significant burdens for the user. We developed a non-invasive, wearable sensing system that can record eating habits over a long period of time in daily life. Our sensing system is composed of two bone conduction microphones placed in the ears that send internal body sound data to a portable IC recorder. Applying frequency spectrum analysis on the collected sound data, we could not only count the number of mastications during eating, but also accurately differentiate between eating, drinking, and speaking activities. This information can be used to evaluate the regularity of meals. Moreover, we were able to analyze sound features to classify the types of foods eaten by food texture.

Key words: Medical Engineering, Pattern Recognition, Wearable Sensing, Eating Habit, Chewing Sound, Mastication Count, Food Texture

1. Introduction

With the increasing number of patients with lifestyle diseases, the Japanese government has been promoting health preservation in everyday life, focusing on exercise, rest, meals, smoking, and alcohol consumption. It is possible to monitor calorie consumption or quality of sleep with wearable devices such as a digital camera or a wristwatch. However, although handwritten check sheets are available, no objective method has been clearly established for monitoring eating habits. Healthcare specialists such as dentists and nutritionists identify time of meals (regularity of meals), mastication count during meals, and types of foods eaten as medically meaningful factors in evaluating eating habits.

As previous work on mastication counting has shown, current devices that measure myoelectric potential from the masseter muscle can count bites (1), but wearing the device is
a significant burden for the user. Another technique using an infrared sensor can detect small changes in temporal muscle tension \(^{(2)}\), but we consider that this method is not applicable in the sense that it bothers users during meals due to ways of sensing and appearance. Recently, analysis of internal body sounds spectra has attracted attention as a way to differentiate between biting and speaking activity \(^{(3)}\) and to classify several types of food \(^{(4)}\) with less burden. However, activity discrimination performance is influenced too much by individual differences, and the classification method is limited to specific aliments.

Most of these just-begun works using internal body sounds had only one function of eating habit analysis. Also they were useful in the limited laboratory situation or didn’t have enough performance. So, at first, it is significant to develop a non-invasive sensing system focusing on sound information in daily life by synthesizing and improving those previous works. Next, it is also important to storage the sensing raw data considering a public database. The database system is expected to have some analysis functions related to eating habit. By using analysis results, users will comprehend their own activities, and the medical staffs will be able to advice exercise and dietary therapy. So, in this paper, our objectives are to develop a non-invasive sensing system as well as sound analysis methods that can accurately sensing eating habits in daily life.

2. Eating Habits Monitoring System

2.1 Overview

Practically, we suppose that sound data recorded over a few days by a wearable sensor would be sent to a data center and analysis results calculated in the data center would be presented to the user himself/herself or to his/her doctor. The analyses of meal times, mastication count during meals, and types of foods eaten are supposed to be useful in this paper. Then the user could make dietary modifications in accordance with regular and objective feedback about his/her eating habits. Figure 1 shows our proposed schematic of the eating habits monitoring system.

![Figure 1. Schematic of the eating habits monitoring system.](image)

2.2 Wearable sensor using a bone conduction microphone

We developed a prototype of a wearable eating habits sensor using a bone conduction microphone (Vibraudio EM20 from TEMCO corp.) for monitoring internal body sounds, and a condenser microphone (WM-E13U from Panasonic corp.) for monitoring environmental sound. This sensing element was placed inside the user’s ear so that we could record internal body sound information over a long period of time in daily life with little interference from external environmental noise. In this study, for the feasibility study of
eating habits analyses using internal body sound data, microphone output signals were recorded with an IC recorder (LS-10 from Olympus corp.) as a 16-bit signal sampled at 48 kHz (Fig. 2).

![Image](https://example.com/image1.png)

**Figure 2.** Overview of the eating habits sensing system.

### 2.3 Eating habit analyses using sound information

Using the prototype described above, we collected internal body sound information, such as mastication and conversation during meals, and analyzed these data to detect meal-related activity, count mastications, and classify the foods eaten into categories based on texture. Figure 3 shows a flow chart of our proposed analysis methods.

First, a sound segment was classified as originating from mastication or not. Next, for each segment classified as a mastication sound, we processed mastication counting and food type classification. The following section describes each operation in detail.

![Image](https://example.com/image2.png)

**Figure 3.** Proposed analysis flow of eating habits.

### 3. Experiments and Results

#### 3.1 Extraction of time of meals

For a feasibility study of discriminating internal body sounds by using the above mentioned wearable sensor, the used sound data which were produced by four types of actions: eating a hard food, eating a soft food, drinking water, and speaking were collected in a laboratory environment. The background noise was 40 dB. The participants were five males in their twenties. They were instructed to eat two types of food products normally with the mouth closed during chewing, to drink off a glass of water, and to read a book out loud during about 40 seconds.

At first, the internal body sounds were divided into one-second data segments and calculated to the power spectrum by the fast Fourier transformation (FFT) (Fig. 4). Then, the features related to the frequency were extracted from these data. Referring to previous work (5–8), the features used were the barycentric frequency, the maximum peak frequency,
the roll-off frequency, the coefficient of the linear predictive coding (LPC), the total power (integration value of the frequency section from 200 to 3000 Hz), and the ratios of divided section power of 1/3 octave band to the total power. Here, the frequency band of total power was limited where the bone conduction microphone used in this study has a flat gain of frequency characteristics (see Fig. 4). Subsequently, a three-second data set of these features was normalized. For a pattern recognition, we used the nearest neighbor algorithm by which a target data is classified with distances to some groups of training data. The Euclid distance was used in this algorithm. This pattern recognition was carried out by a calculation software (MATLAB R2008 from Mathworks inc.) on a personal computer.

Figure 5 shows a sample of recorded internal body sound data (top) and the corresponding discrimination result (bottom) for a participant. When we classified activities of a participant with his own database, we could differentiate the four identified activities. The average classification score for the five participants was more than 80% accuracy for each activity (Fig. 6 above).

However, when we classified activities of one participant with the other four participants’ databases to check the effect of individual differences, the average classification score dropped to about 70% (Fig. 6 below). Although the classification scores for eating hard and soft foods were still high, the score for drinking water was very low (about 50%). Differences are expected in how individuals eat and drink. Therefore, it is necessary to select features that are less affected by individual differences or to improve the classification algorithm.
3.2 Counting the number of mastications

We evaluated the number of mastications from the sound data identified as “eating”. We counted by applying a low-pass filter based on the person’s masticatory cycle and then detecting peaks. This analysis flowchart is shown in Fig. 7. To check the validity of this algorithm, we collected internal body sound data during mastication by above mentioned wearable sensor with a sampling rate of 48 kHz. A participant masticated a food sample 30 times with a chewing cycle of 0.4–1.6 seconds. Gum was used as the food sample because its texture does not change during mastication.

First, the cut-off frequency of the filter was set to a constant value of 1.5 Hz because a
typical chewing cycle is said to be 0.6–0.8 seconds. As an example, the extracted results of
the number of mastications for a piece of gum are shown in Fig. 8. Problematically, the
difference between the actual and the calculated numbers became larger in shorter or longer
chewing cycles.

Next, the cut-off frequency was set to a variable determined from the power spectrum
of sound data at that moment. Several tens of seconds of sound data which were defined as
“eating” in the meal time extraction process were calculated by FFT. We chose the first
frequency of the local maximum peak which appeared in the range of about 0.5–5 Hz. A
power spectrum of Fig. 7’s sample is shown in Fig. 9. As a result, we were able to achieve a
reliable count of mastications regardless of chewing rate.

Figure 7. Mastication counting flow.

Figure 8. Calculation of mastications number using a constant or a variable cut-off frequency in low-pass filter.

Figure 9. The first frequency of the local maximum peak.
3.3 Classification of food types

Concerning food type classification, we focused on the food texture, defined by mechanical properties such as hardness, cohesiveness (brittleness, chewiness, and gumminess), viscosity, springiness, and adhesiveness \(^9\). During the early stage of mastication, hardness and cohesiveness are likely to change. For the feasibility study of food type classification regarding texture, we carried out principal components analysis (PCA) on the mastication sounds.

We used food samples with three types of food textures: hard-crumbly (rice crackers and peanuts), high springiness (fish sausage and konjac food (*konnyaku*)), and soft (rice and bread rolls). Figure 10 shows the food textures of these samples \(^10\). A participant was instructed to masticate 3 grams of each sample with a mastication frequency of 1 Hz. Same as above mentioned two signal processes, we collected internal body sound data during mastication by the wearable sensor. Then we obtained the power spectrum by FFT, and also extracted the features shown in Fig. 4. Using these features, the PCA was carried out by MATLAB software. Figure 11 shows results of food type clustering regarding food texture. It was found that the first principal component (49% in contribution ratio) indicated the effect of sound from hard foods, so we were able to classify the hard-crumbly group and the others in the feature space. On the other hand, although each data point was organized in the feature space, we could not clearly classify the other two groups, possibly because the first defined food groups included other types of textures.
4. Conclusion

In this paper, we proposed a non-invasive wearable sensing system that can record sound data related to eating habits in daily life. Analyzing the internal body sound data from the bone conduction microphone, we extracted meal times, counted the number of mastications during eating, and classified the types of foods eaten based on their textures.

Now, we are building a sound information database system for accumulating and sharing a large amount of data collected with our wearable sensor. Also we are planning to demonstrate our proposed system in the filed of health check. From the feedback of medical users, this system will be improved in the future.

Acknowledgement

This research was supported by JST, CREST.

References

