Video Article

Design and Evaluation of Smart Glasses for Food Intake and Physical Activity Classification

Jungman Chung[1], Wonjoon Oh[2], Dongyoub Baek[3], Sunwoong Ryu[2], Won Gu Lee[4], Hyunwoo Bang[2]

[1]School of Mechanical and Aerospace Engineering, Seoul National University

[2]Envisible, Inc.

[3]Curiosis, Inc.

[4]Department of Mechanical Engineering, Kyung Hee University

Correspondence to: Won Gu Lee at termylee@khu.ac.kr, Hyunwoo Bang at bang@envisible.com

URL: https://www.jove.com/video/56633
DOI: doi:10.3791/56633

Keywords: Engineering, Issue 132, Wearable device, monitoring of ingestive behavior (MIB), food intake, physical activity, load cell, PCB fabrication, 3D printing, machine learning, support vector machine (SVM)

Date Published: 2/14/2018

Citation: Chung, J., Oh, W., Baek, D., Ryu, S., Lee, W.G., Bang, H. Design and Evaluation of Smart Glasses for Food Intake and Physical Activity Classification. *J. Vis. Exp.* (132), e56633, doi:10.3791/56633 (2018).

Abstract

This study presents a series of protocols of designing and manufacturing a glasses-type wearable device that detects the patterns of temporalis muscle activities during food intake and other physical activities. We fabricated a 3D-printed frame of the glasses and a load cell-integrated printed circuit board (PCB) module inserted in both hinges of the frame. The module was used to acquire the force signals, and transmit them wirelessly. These procedures provide the system with higher mobility, which can be evaluated in practical wearing conditions such as walking and waggling. A performance of the classification is also evaluated by distinguishing the patterns of food intake from those physical activities. A series of algorithms were used to preprocess the signals, generate feature vectors, and recognize the patterns of several featured activities (chewing and winking), and other physical activities (sedentary rest, talking, and walking). The results showed that the average F_1 score of the classification among the featured activities was 91.4%. We believe this approach can be potentially useful for automatic and objective monitoring of ingestive behaviors with higher accuracy as practical means to treat ingestive problems.

Video Link

The video component of this article can be found at https://www.jove.com/video/56633/

Introduction

Continuous and objective monitoring of food intake is essential for maintaining energy balance in the human body, as excessive energy accumulation may cause overweightness and obesity[1], which could result in various medical complications[2]. The main factors in the energy imbalance are known to be both excessive food intake and insufficient physical activity[3]. Various studies on the monitoring of daily energy expenditure have been introduced with automatic and objective measurement of physical activity patterns through wearable devices[4,5,6], even at the end-consumer level and medical stage[7]. Research on the monitoring of food intake, however, is still in the laboratory setting, since it is difficult to detect the food intake activity in a direct and objective manner. Here, we aim to present a device design and its evaluation for monitoring the food intake and physical activity patterns at a practical level in daily life.

There have been various indirect approaches to monitor the food intake through chewing and swallowing sounds[8,9,10], movement of the wrist[11,12,13], image analysis[14], and electromyogram (EMG)[15]. However, these approaches were difficult to apply to daily life applications, because of their inherent limitations: the methods using sound were vulnerable to be influenced by environmental sound; the methods using the movement of the wrist were difficult to distinguish from other physical activities when not consuming food; and the methods using the images and EMG signals are restricted by the boundary of movement and environment. These studies showed the capability of automated detection of the food intake using sensors, but still had a limitation of practical applicability to everyday life beyond laboratory settings.

In this study, we used the patterns of temporalis muscle activity as the automatic and objective monitoring of the food intake. In general, the temporalis muscle repeats the contraction and relaxation as a part of masticatory muscle during the food intake[16,17]; thus, the food intake activity can be monitored by detecting the periodic patterns of temporalis muscle activity. Recently, there have been several studies utilizing the temporalis muscle activity[18,19,20,21], which used the EMG or piezoelectric strain sensor and attaching them directly onto human skin. These approaches, however, were sensitive to the skin location of the EMG electrodes or strain sensors, and were easily detached from the skin due to the physical movement or perspiration. Therefore, we proposed a new and effective method using a pair of glasses that sense the temporalis muscle activity through two load cells inserted in both the hinges in our previous study[22]. This method showed great potential of detecting the food intake activity with a high accuracy without touching the skin. It was also un-obtrusive and non-intrusive, since we used a common glasses-type device.

Copyright © 2018 Journal of Visualized Experiments

In this study, we present a series of detailed protocols of how to implement the glasses-type device and how to use the patterns of temporalis muscle activity for monitoring the food intake and physical activity. The protocols include the process of hardware design and fabrication that consists of a 3D-printed frame of the glasses, a circuit module, and a data acquisition module, and include the software algorithms for data processing and analysis. We furthermore examined the classification among several featured activities (*e.g.*, chewing, walking, and winking) to demonstrate the potential as a practical system that can tell a minute difference between the food intake and other physical activity patterns.

Protocol

NOTE: All the procedures including the use of human subjects were accomplished by a non-invasive manner of simply wearing a pair of glasses. All the data were acquired by measuring the force signals from load cells inserted in the glasses that were not in direct contact with the skin. The data were wirelessly transmitted to the data recording module, which, in this case is a designated smartphone for the study. All the protocols were not related to *in vivo*/*in vitro* human studies. No drug and blood samples were used for the experiments. Informed consent was obtained from all subjects of the experiments.

1. Manufacturing of a Sensor-integrated Circuit Module

1. **Purchase electronic components for manufacturing the circuit module.**
    1. Purchase two ball-type load cells, each of which operates in a range between 0 N and 15 N, and produces an output of low differential voltage with maximum 120 mV span in a 3.3 V excitation.
    NOTE: These load cells are used to measure force signals on both the left and right sides of the glasses.
    2. Purchase two instrumentation amplifiers and two 15 kΩ gain-setting resistors.
    NOTE: The instrumentation amplifier and the gain-setting resistor are used to amplify the force signal of the load cell eight times, up to 960 mV.
    3. Purchase a micro controller unit (MCU) with wireless capability (*e.g.*, Wi-Fi connectivity), and a 10-bit analog-to-digital converter (ADC).
    NOTE: The MCU is used to read the force signals and transmit them to a data acquisition module wirelessly. Because one analog input pin is used for two analog force inputs, the use of a multiplexer is introduced in the next step 1.1.4.
    4. Purchase a two-channel analog multiplexer that handles the two input signals with one ADC pin on the MCU.
    5. Purchase a lithium-ion polymer (LiPo) battery with 3.7 V nominal voltage, 300 mAh nominal capacity, and 1 C discharge rate.
    NOTE: The battery capacity was chosen to supply enough current per hour more than 200 mAh and to operate the system reliably for about 1.5 h of an experiment.
    6. Purchase a 3.3 V voltage regulator for linear down-regulation of the 3.7 V battery voltage to the 3.3 V operating voltage of the system.
    7. Purchase five 12 kΩ surface-mounted devices (SMD) type resistors as pull-up resistors of the MCU. The resistor's footprint is 2.0 mm x 1.2 mm (size 2012).

2. **Fabricate printed circuit boards (PCBs). This step is about drawing the circuit boards, and making the artwork (*i.e.*, the board layout, the .brd file) and the schematic (*i.e.*, the .sch file) for PCB fabrication. A basic understanding of the process of creating artwork and schematic files is required for development.**
    1. Draw a schematic of a left circuit containing the battery using an electronic design application as shown in **Figure 1A**. Save the result as both artwork (.brd) and schematic (.sch) files.
    2. Draw a schematic of a right circuit containing the MCU using an electronic design application as shown in **Figure 1B**. Save the result as both artwork (.brd) and schematic (.sch) files.
    3. Fabricate the circuit boards by placing an order with a PCB fabrication company.
    4. Solder every electronic component prepared in step 1.1 to the PCBs as shown in **Figure 2** and **Figure 3**.
    CAUTION: The instrumentation amplifier is very sensitive to the soldering temperature. Make sure that lead temperature does not exceed 300 °C for 10 s during soldering, otherwise it may cause permanent damage to the component.

2. 3D Printing of a Frame of the Glasses

1. Draw the 3D model of the head piece of the glasses using a 3D modeling tool as shown in **Figure 4A**. Export the result to the .stl file format.
2. Draw the 3D model of the left and right temples of the glasses using a 3D modeling tool as shown in **Figure 4B** and **Figure 4C**. Export the results to the .stl file format.
3. Print the head piece and temple parts using a 3D printer and a carbon fiber filament at 240 °C of a nozzle temperature and 80 °C of a bed temperature.
NOTE: The use of any commercial 3D printer and any types of filaments such as acrylonitrile butadiene styrene (ABS) and polylactide (PLA) can be permitted. The nozzle and bed temperatures may be adjusted according to the filament and printing conditions.
4. Heat the tips of the temples using a hot air blower of a 180 °C setting and bend them inward about 15 degrees to contact the epidermis of the temporalis muscle like conventional glasses.
NOTE: The degree of bending of the glasses temple does not need to be rigorous as the purpose of the curvature is to increase a form factor by helping the glasses fit on a subject's head when equipped. Be careful, however, as excessive bending will prevent the temples from touching the temporalis muscle, which makes it impossible to collect significant patterns.
5. Repeat the steps from step 2.1–2.4 to print two different sizes of the glasses frame to fit multiple head sizes as shown in **Figure 4**.

3. Assembly of All Parts of the Glasses

1. Insert the PCBs on both sides of the temples of the glasses using M2 bolts as shown in **Figure 5**.
2. Assemble the head piece and the temples by inserting the M2 bolts into the hinge joints.

3. Connect the left and right PCBs using the 3-pin connecting wires as shown in **Figure 5**.
4. Connect the battery to the left circuit and attach it with an adhesive tape to the left temple. The mounting side of the battery is not critical, since it may vary depending on the PCB design.
5. Cover the glasses with rubber tapes on the tip and the nose pad to add more friction with the human skin as shown in **Figure 5**.

4. Development of a Data Acquisition System

NOTE: The data acquisition system is composed of a data transmitting module and a data receiving module. The data transmitting module reads the time and the force signals of both sides, and then sends them to the data receiving module, which gathers the received data and writes them to .tsv files.

1. **Upload the data transmitting application to the MCU of the PCB module following the procedures in steps 4.1.1–4.1.3.**
    1. Run the "GlasSense_Server" project attached to the supplementary files using a computer.
    NOTE: This project was built with Arduino integrated development environment (IDE). It provides the ability to read the time and force signals with 200 samples/s, and transmit them to the data receiving module.
    2. Connect the PCB module to the computer via a universal serial bus (USB) connector.
    3. Press the "Upload" button on the Arduino IDE to flash the programming codes from step 4.1.1 into the MCU.

2. **Upload the data receiving application to a smartphone, which is used to receive the data wirelessly, following the procedures in steps 4.2.1–4.2.3.**
    1. Run the "GlasSense_Client" project attached to the supplementary files using a computer.
    NOTE: This project was built with C# programming language. It provides the ability to receive data and save the .tsv files, which contain a subject's information, such as name, sex, age, and body mass index (BMI).
    2. Connect the smartphone to the computer via a USB connector to build the data receiving application.
    3. Press the "File > Build & Run" button on the C# project to build the data receiving application to the smartphone.

5. Data Collection from a User Study

NOTE: This study collected six featured activity sets: sedentary rest (SR), sedentary chewing (SC), walking (W), chewing while walking (CW), sedentary talking (ST), and sedentary wink (SW).

1. Select a pair of glasses which have an appropriate size to the user to be tested. Fine-tune the tightness with the support bolts at both the hinges (**Figure 5**).
CAUTION: The force values must not exceed 15 N, since the force sensors used in this study may lose the fine linear characteristic beyond the operation range. The force values can be fine-tuned by loosening or tightening the support bolts.
2. **Record the activities of all subjects by pressing the "Record" button on the application built in step 4.2.3.**
    1. Record an activity during a 120-s block and generate a recording file of it.
        1. In the case of SR, sit the subject in a chair and have them use a smartphone or read a book. Allow movement of the head, but avoid movement of the whole body.
        2. In the cases of SC and CW, have the subjects eat two types of food texture (toasted bread and chewing jelly) in order to reflect different food properties. Serve the toasted bread in slices of 20 mm x 20 mm, which is a good size for eating.
        3. In the case of W, have the subjects walk at a speed of 4.5 km/h on a treadmill.
        4. In the case of ST, sit the subjects down and have them read a book out loud in a normal tone and speed.
        5. In the case of SW, inform the subjects to wink on the timing of a bell sound of 0.5 s long every 3 s.

    2. Generate a recording file in .tsv format from the data collected in step 5.2.1.
    NOTE: This file contains a sequence of the time when the data were received, a left force signal, a right force signal, and a label representing the current facial activity. Visualizations of temporal signals of all activities in a block of a user were depicted in **Figure 6**. The six featured activity sets (SR, SC, W, CW, ST, and SW) were labeled as 1, 2, 3, 4, 5, and 6, respectively. The labels were used to compare the predicted classes in section 8 of the protocol.
    3. Take a 60-s break after the recording block. Take off the glasses during the break, and re-wear them again when the recording block restarts.
    4. Repeat the block-and-break set of steps 5.2.1 and 5.2.2 four times for each activity.
    5. In the case of SW, have the subject wink repeatedly with the left eye during one block, and then wink repeatedly with the right eye during the next block.

3. Repeat steps 5.1–5.2 and collect the data from 10 subjects. In this study, we used five males and five females, the average age was 27.9 ± 4.3 (standard deviation; s.d.) years, which ranged at 19–33 years, and the average BMI was 21.6 ± 3.2 (s.d.) kg/m^2, which ranged at 17.9–27.4 kg/m^2.
NOTE: In this study, the subjects who did not have any medical conditions to chew food, wink, and walk were recruited, and this condition was used for inclusion criteria.

6. Signal Preprocessing and Segmentation

NOTE: The left and right signals are calculated separately in the following procedures.

1. **Prepare a series of temporal frames of 2 s long.**
    1. Segment the 120 s recorded signals into a set of 2 s frames by hopping them at 1-s intervals using MATLAB as shown in **Figure 6**.

NOTE: The segmented frames of 2 s long were used to extract features in section 7. The 1 s hopping size was determined to divide the signals by the 3 s wink interval already mentioned in step 5.2.1.
2. Apply a low-pass filter (LPF) using a 5th order Butterworth filter with a cutoff frequency of 10 Hz for each frame.
3. Save the results of step 6.1.2 as the temporal frames for the next steps in step 7.1.

2. **Prepare a series of spectral frames.**
   1. Subtract the median from the original signals of each frame to remove the preload when wearing the glasses.
      NOTE: The preload value is not required for the following frequency analysis, since it does not include any information about chewing, walking, wink, *etc.* It could, however, contain significant information, which can vary from subject to subject, from every setting of the glasses, and even from the moment a subject wears the glasses.
   2. Apply a Hanning window to each frame to reduce a spectral leakage on frequency analysis.
   3. Produce and save a single-sided spectrum by applying a fast Fourier transform (FFT) to each frame.

3. Define a combination of a temporal and a spectral frame of the same time as a frame block (or simply a frame).

7. Generation of Feature Vectors

NOTE: A feature vector is generated per frame produced in section 6 of the protocol. The left and right frames are calculated separately and combined into a feature vector in the following procedures. All the procedures were implemented in MATLAB.

1. Extract statistical features from a temporal frame in step 6.1 of the protocol. A list of the total number of 54 features is given in **Table 1**.
2. Extract statistical features from a spectral frame in step 6.2 of the protocol. A list of the total number of 30 features is given in **Table 2**.
3. Generate an 84-dimensional feature vector by combining the temporal and spectral features above.
4. Label the generated feature vectors from the recordings in step 5.2 of the protocol.
5. Repeat the steps from steps 7.1–7.4 for all frame blocks and generate a series of feature vectors.

8. Classification of the Activities into Classes

NOTE: This step is to select the classifier model of a support vector machine (SVM)[23] by determining parameters that show the best accuracy from the given problem (*i.e.,* feature vectors). The SVM is a well-known supervised machine learning technique, which shows excellent performance in generalization and robustness using a maximum margin between the classes and a kernel function. We used a grid-search and a cross-validation method to define a penalty parameter C and a kernel parameter γ of the radial basis function (RBF) kernel. A minimum understanding of machine learning techniques and the SVM is required to perform the following procedures. Some referential materials[23,24,25] are recommended for better understanding of machine learning techniques and the SVM algorithm. All the procedures in this section were implemented using LibSVM[25] software package.

1. Define a grid of pairs of (C, γ) for the grid-search. Use exponentially growing sequences of C (2^{-10}, 2^{-5}, …, 2^{30}) and γ (2^{-30}, 2^{-25}, …, 2^{10}).
   NOTE: These sequences were determined heuristically.
2. Define a pair of (C, γ) (*e.g.*, (2^{-10}, 2^{-30})).
3. **For the defined grid in step 8.2, perform the 10-fold cross-validation scheme.**
   NOTE: This scheme divides the entire feature vectors into 10-part subsets, then test one subset from the classifier model trained by the other subsets, and repeat it over all the subsets, one by one. Therefore, every feature vectors can be tested sequentially.
   1. Divide the entire feature vectors into 10-part subsets.
   2. Define a testing set from a subset, and a training set from the remaining 9 subsets.
   3. Define a scale vector that scales all elements of the feature vectors to the range of [0, 1] for the training set.
      NOTE: The scale vector has same dimension with the feature vector. It consists of a set of multipliers that scales the same row (or column) of all feature vectors to the range of [0, 1]. For example, the first feature of a feature vector is linearly scaled to the range of [0, 1] for the all first features of the training feature vectors. Note that the scale vector is defined from the training set, because the testing set should be assumed to be unknown. This step increases the accuracy of the classification by making the features the equal range and avoiding numerical errors during the calculation.
   4. Scale each feature of the training set to the range of [0, 1] using the scale vector obtained in step 8.2.3.
   5. Scale each feature of the testing set to the range of [0, 1] using the scale vector obtained in step 8.2.3.
   6. Train the training set through the SVM with the defined pair of (C, γ) in step 8.2, and then build a classifier model.
   7. Test the testing set through the SVM with the defined pair of (C, γ) in step 8.2, and the classifier model obtained from the training procedure.
   8. Calculate a classification accuracy on the testing set. The accuracy was calculated from the percentage of feature vectors which are correctly classified.
   9. Repeat the steps 8.2.2–8.2.8 for all the subsets, and calculate the average accuracy of all subsets.

4. Repeat the steps 8.2–8.3.9 for all grid points of a pair of (C, γ).
5. Find the local maximum of the highest accuracy of the grid. All the procedures of section 8 are illustrated in **Figure 7**.
6. (Optional) If the step of the grid is considered coarse, repeat the steps 8.1–8.5 in a finer grid near the local maximum found in step 8.5, and find a new local maximum of the fine grid.
7. Compute the precision, recall, and F_1 score of each class of activities from the following equations:

$$\text{precision} = \frac{TP}{TP + \sum FP} \qquad \textbf{Equation 1}$$

$$\text{recall} = \frac{TP}{TP + \sum FN} \qquad \textbf{Equation 2}$$

$$F_1\ score = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}$$ **Equation 3**

where TP, FP, and FN represent true positives, false positives, and false negatives for each activity, respectively. The confusion matrix of all the activities is given in **Table 3**.

Representative Results

Through the procedures outlined in the protocol, we prepared two versions of the 3D printed frame by differentiating the length of the head piece, L_H (133 and 138 mm), and the temples, L_T (110 and 125 mm), as shown in **Figure 4**. Therefore, we can cover several wearing conditions, which can be varied from the subjects' head size, shape, *etc.* The subjects chose one of the frames to fit to their head for the user study. The vertical distance, L_h, between the hinge joint and the hole for the support bolt was set to 7.5 mm so that the amplified force would not exceed 15 N, which is the linear operating range of the load cell. Finally, the head piece should have a thickness, t_H, that can resist the bending moment transmitted from both support bolts when equipped. We chose the t_H to be 6 mm with a use of carbon fiber material from a heuristic approach. The contact points can be adjusted through the support bolts to fine-tune the tightness of the glasses as shown in **Figure 5**.

Table 3 shows the representative results of the classification for all the activity sets. The average F_1 score resulted in 80.5%. If considered as a single score, the performance may seem to be relatively degraded compared to the result of our previous study[22]. We, however, can extract significant information by comparing the outcomes between each activity. The SR was relatively well distinguished from the SC, CW, and SW, but not from the W and ST. Both chewing activities, SC and CW, were difficult to distinguish from each other. On the other hand, it can be observed that both chewing activities can be easily distinguished from the SR, W, ST, and SW, which represent the other physical activities. In the case of the SW, the wink activity turned out to be misclassified slightly throughout the other activities.

From the results of the **Table 3**, we can observe in-depth details of the classification. First, the two chewing activities, SC and CW, were clearly distinguished from the other activities. Among them, the distinction from the walking activity suggests a possibility that the food intake activity, which is the main purpose of this study, can be easily separable from the active physical activity, such as walking, using our system. As shown in **Figure 6**, it can be verified that the chewing and wink signals, activated from the temporalis muscle activity, were significantly different from those not activated by the temporalis muscle activity. On the other hand, the distinction between the two chewing activities showed relatively high misclassifications. They played a dominant role in lowering the both the precision and recall of the chewing activities.

In terms of chewing detection, the SR, W, and ST can be regarded as unintended noise in daily life. The wink activity, on the other hand, can be considered as meaningful measurement, because it is also activated from the temporalis muscle activity as well. Based on the above, the two chewing activities were bounded into a chewing activity (CH), and the other activities except for the wink were grouped into a physical activity (PA). **Table 4** shows the classification results on these activities: chewing (CH), physical activity (PA), and sedentary wink (SW). We can find more remarkable results from it. It predicts information about whether the system is robust for detecting food intake without being affected by other physical activities. Furthermore, it also indicates whether it is possible to distinguish food intake from other face activity such as wink. The results show that the chewing activity can be well distinguished from the other activities by a high F_1 score of 93.4%. In the case of wink, the recall (85.5%) was slightly lower than that of the other activities. This means that the quality of the collected data of wink was likely to be low, as the users had to wink at the exact time in 3 s intervals. In fact, it was observed that the users missed the wink or the glasses shifted down occasionally during the user study.

In order to obtain more meaningful results from the above, we grouped and re-defined the activities into new ones. The two chewing activities, SC and CW, were grouped into one activity, and defined as chewing. The SR, W, and ST, which had a large degree of misclassification among themselves, were also grouped into one activity, defined as physical activity. As a result, we obtained new representative results of the classification re-performed through the activities featured as chewing (CH), physical activity (PA), and sedentary wink (SW), as shown in **Table 4**. The results showed that a high prediction score with an average F_1 score of 91.4% of.

Copyright © 2018 Journal of Visualized Experiments

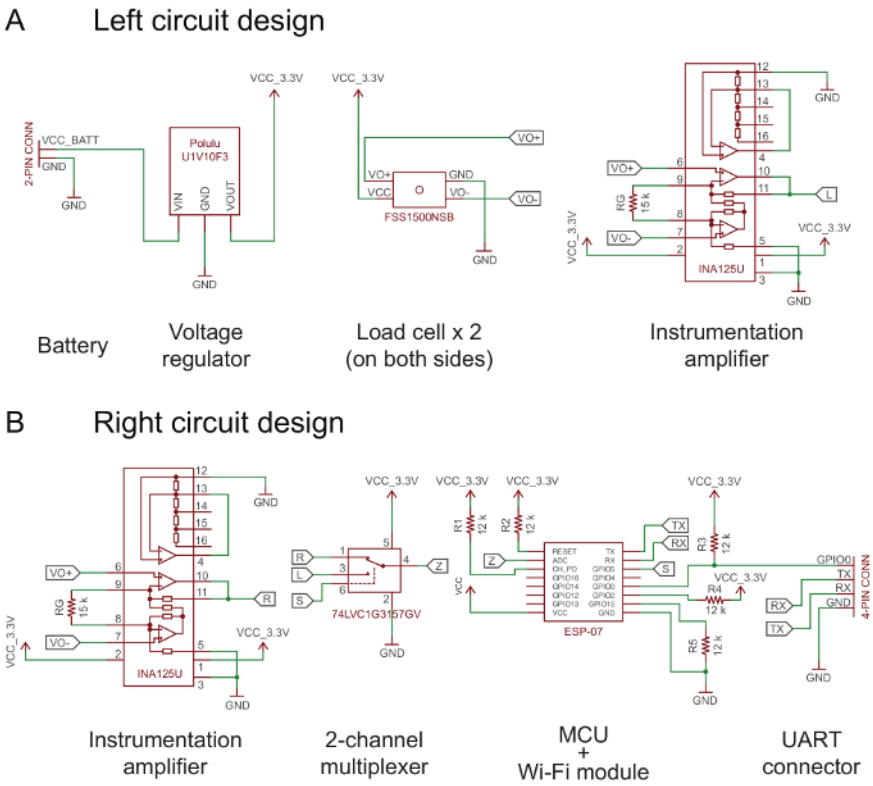

Figure 1: Schematic diagrams of both left and right circuits. (**A**) Schematic diagrams of the left circuit. It contains a battery to supply power to the both left and right circuit. A 3.3 V voltage regulator with bypass capacitor was provided to supply a stable operating voltage to the system. Load cells presented here were inserted into both sides of the circuit (**B**) Schematic diagrams of the right circuit. It contains a micro controller unity (MCU) with Wi-Fi capability. A two-channel multiplexer was provided to process two force signals from both sides with one analog-to-digital converter (ADC) of the MCU. A universal asynchronous receiver/transmitter (UART) connector was used to flash the MCU. Please click here to view a larger version of this figure.

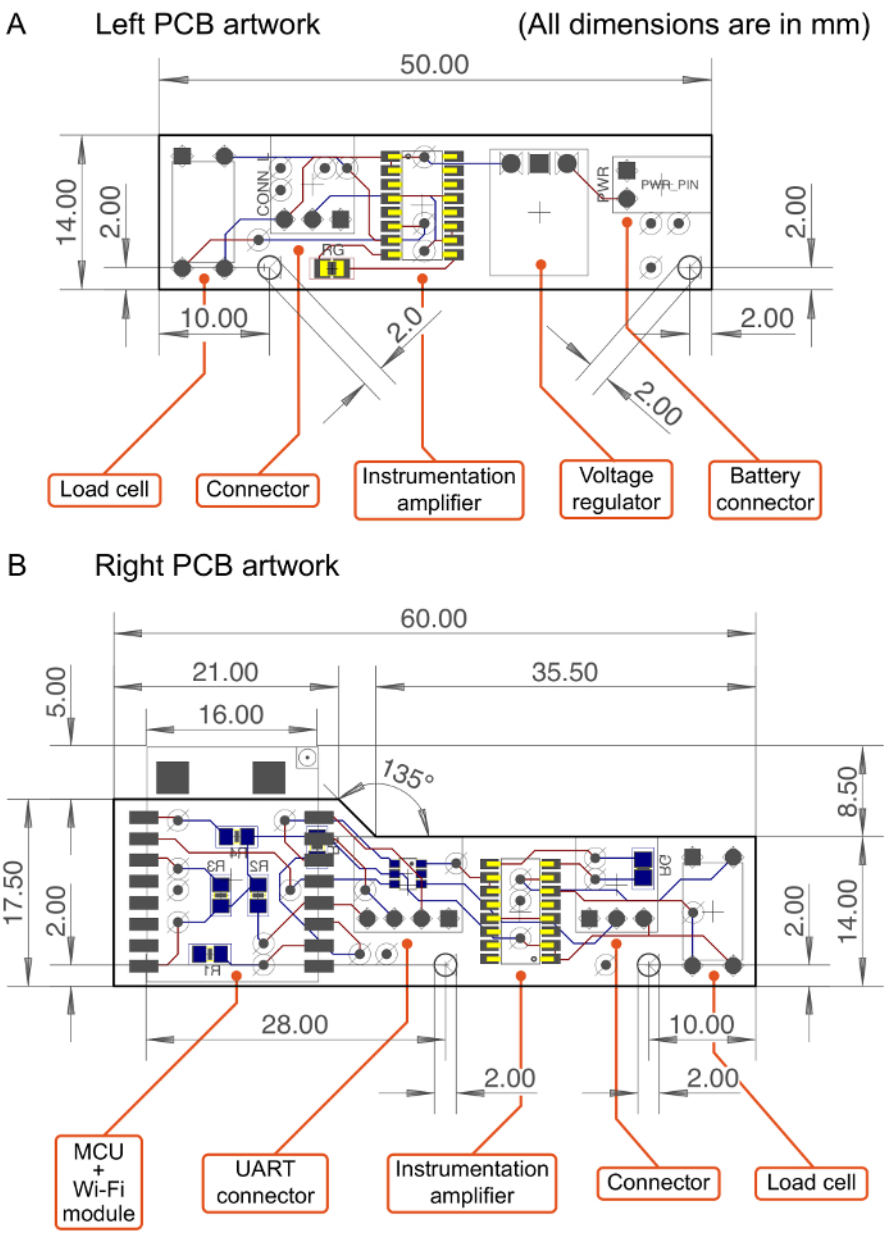

Figure 2: PCB artworks of both left and right circuits. (**A**) An artwork of the left PCB. All electronic components are displayed as actual measurements in mm. (**B**) An artwork of the right PCB. Please click here to view a larger version of this figure.

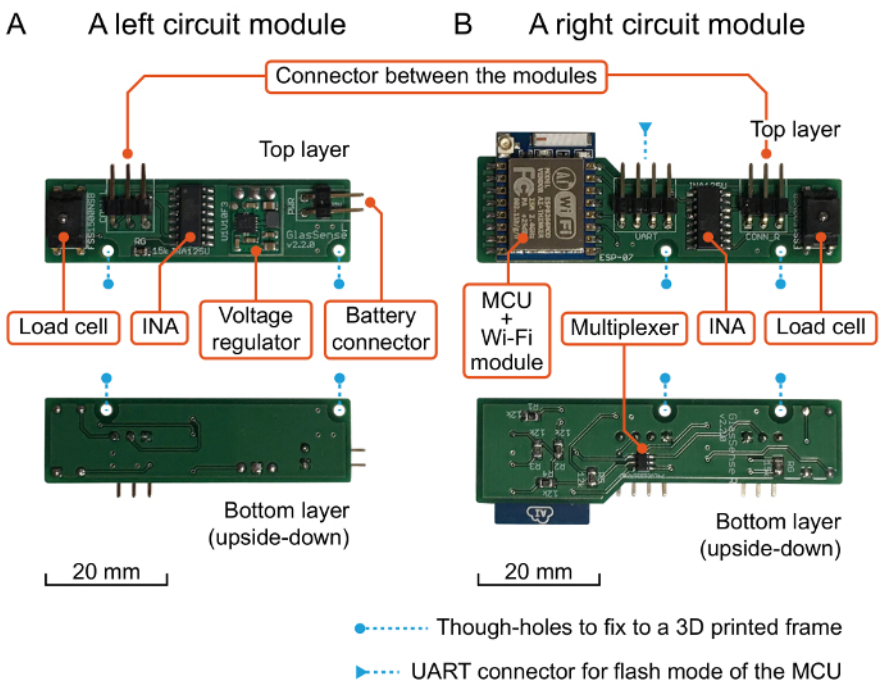

Figure 3: Representative results of PCBs soldered with all components. (**A**) The left circuit module. The load cell was integrated into the board. It contains a 2-pin connector for battery and a 3-pin connector to connect to the right board. (**B**) The right circuit module. The load cell was also integrated into the board. It contains a 4-pin connector for flashing mode of the MCU, and a 3-pin connector to connect to the left circuit. Please click here to view a larger version of this figure.

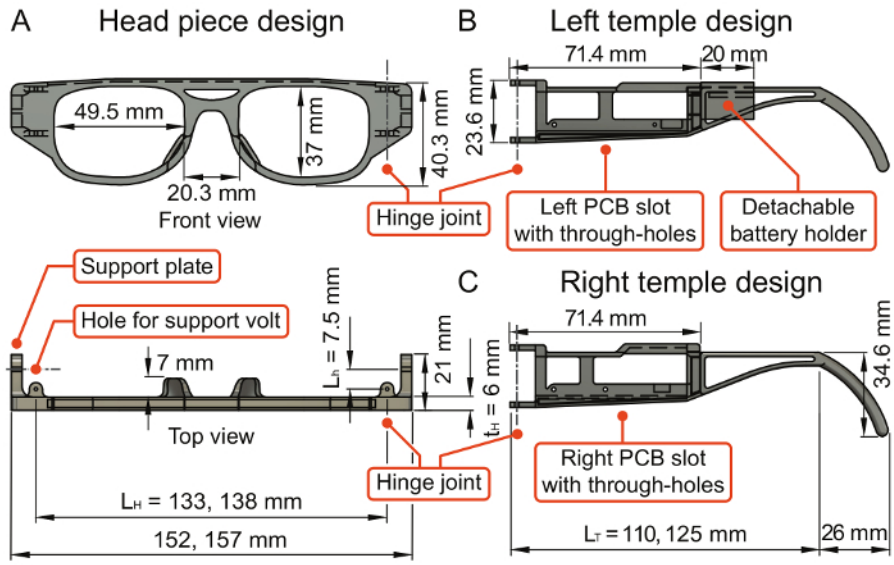

Figure 4: A 3D model design of the frame of the glasses. (**A**) The design of the head piece. The upper figure shows a front view, and the lower figure shows a top view of the head piece. The length of the head piece, L_H, is a design parameter to cover various head size of subjects. We 3D printed two versions of the head piece by differentiating it. The thickness of the head piece, t_H, was defined by heuristic. The distance between a hinge joint and a hole for a support bolt, L_h, was set from the mechanical amplification factor. (**B**) The design of the temples. The upper figure shows the left temple, and the lower figure shows the right temple. The PCBs in **Figure 3** were inserted into slots and a battery was mounted to a battery holder. Please click here to view a larger version of this figure.

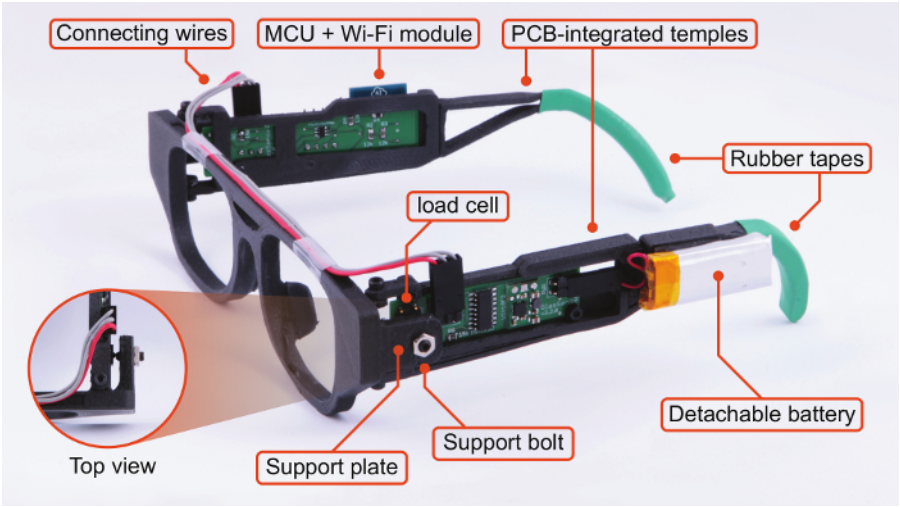

Figure 5: A representative result of thePCB-integrated glasses. The PCBs were inserted into the slots with bolts. The nose pads and the tips of the temples were covered by rubber tapes to add friction with skin. When the glasses are equipped, the load cells are pressed by support bolts on both sides. The tightness of the glasses can be fine-tuned by loosening or tightening the support bolts. Please click here to view a larger version of this figure.

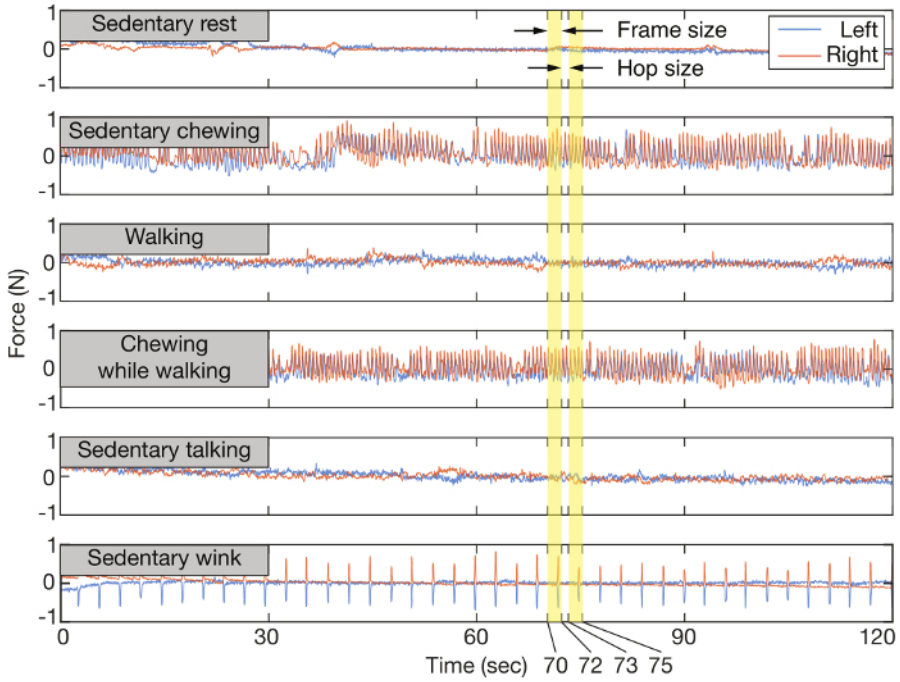

Figure 6: Temporal signals in a recording block of a user for all activities. The y-axis represents the measured force, which was subtracted by its median of the recording block for a visualization purpose. The maximum amplitudes of the chewing activities are larger than the other activities. Left and right signals of wink activity are inverted. The figure shows an example of the left wink. A 2 s frame was used to define a feature vector by hopping the signals at 1 s interval. Please click here to view a larger version of this figure.

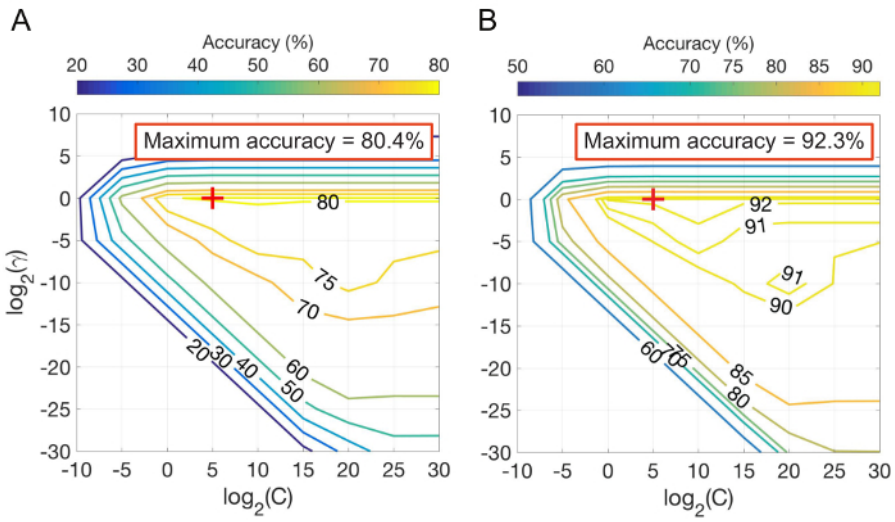

Figure 7: Representative results of finding the local maximum accuracy through various pairs of (C, γ). (**A**) A contour plot of cross-validated accuracies of all activities defined in **Table 3**. Each axis increases exponentially and the range was heuristically selected. The local maximum accuracy of 80.4% occurred at (C, γ) = (2^5, 2^0). (**B**) A contour plot of cross-validated accuracies of re-defined activities in **Table 4**. The maximum accuracy of 92.3% occurred at (C, γ) = (2^5, 2^0), and was much accurate than the result of (**A**). Please click here to view a larger version of this figure.

No.	Feature description	No.	Feature description
1	Standard deviation L	28	Skenwness R
2	Standard deviation R	29	Kurtosis L
3	Coefficient of variation L	30	Kurtosis R
4	Coefficient of variation R	31	Autocorrelation function coefficients L
5	Zero crossing rate L	32	Autocorrelation function coefficients R
6	Zero crossing rate R	33	Signal energy L
7	20th percentile L	34	Signal energy R
8	20th percentile R	35	Log signal energy L
9	50th percentile L	36	Log signal energy R
10	50th percentile R	37	Entropy of energy L
11	80th percentile L	38	Entropy of energy R
12	80th percentile R	39	Peak-to-peak amplitude L
13	Interquartile range L	40	Peak-to-peak amplitude R
14	Interquartile range R	41	The number of peaks L
15	Square sum of 20th percentile L	42	The number of peaks R
16	Square sum of 20th percentile R	43	Mean of time between peaks L
17	Square sum of 50th percentile L	44	Mean of time between peaks R
18	Square sum of 50th percentile R	45	Std. of time between peaks L
19	Square sum of 80th percentile L	46	Std. of time between peaks R
20	Square sum of 80th percentile R	47	Prediction ratio L
21	1st bin of binned distribution L	48	Prediction ratio R
22	1st bin of binned distribution R	49	Harmonic ratio L
23	2nd bin of binned distribution L	50	Harmonic ratio R
24	2nd bin of binned distribution R	51	Fundamental frequency L
25	3rd bin of binned distribution L	52	Fundamental frequency R
26	3rd bin of binned distribution R	53	Correlation coefficient of L and R
27	Skenwness L	54	Sigmal magnitude area of L and R

Table 1: Extracted statistical features of a temporal frame. A total of 54 features were extracted. The left and right signals were calculated separately except for the correlation features, 53 and 54.

Copyright © 2018 Journal of Visualized Experiments

No.	Feature description	No.	Feature description
1	Spectral energy L	16	Spectral spread R
2	Spectral energy R	17	Spectral entropy L
3	Spectral zone 1 of energy L	18	Spectral entropy R
4	Spectral zone 1 of energy R	19	Spectral entropy of energy L
5	Spectral zone 2 of energy L	20	Spectral entropy of energy R
6	Spectral zone 2 of energy R	21	Spectral flux L
7	Spectral zone 3 of energy L	22	Spectral flux R
8	Spectral zone 3 of energy R	23	Spectral rolloff L
9	Spectral zone 4 of energy L	24	Spectral rolloff R
10	Spectral zone 4 of energy R	25	Maximum spectral crest L
11	Spectral zone 5 of energy L	26	Maximum spectral crest R
12	Spectral zone 5 of energy R	27	Spectral skewness L
13	Spectral centroid L	28	Spectral skewness R
14	Spectral centroid R	29	Spectral kurtosis L
15	Spectral spread L	30	Spectral kurtosis R

Table 2: Extracted statistical features of a spectral frame. A total of 30 features were extracted. The left and right signals were calculated separately. From the features in **Table 1** and **Table 2**, a feature vector consists of a total of 84 features.

Predicted activity	Actual activity						Total	Precision
	[a]SR	[b]SC	[c]W	[d]CW	[e]ST	[f]SW		
SR	1222	18	79	6	168	75	1568	77.9%
SC	10	1268	17	159	46	15	1515	83.7%
W	55	19	1212	32	144	20	1482	81.8%
CW	3	158	34	1327	28	12	1562	85.0%
ST	192	75	185	19	1117	55	1643	68.0%
SW	78	22	33	17	57	1383	1590	87.0%
Total	1560	1560	1560	1560	1560	1560	9360	
Recall	78.3%	81.3%	77.7%	85.1%	71.6%	88.7%		80.4%
F_1 score	78.1%	82.5%	79.7%	85.0%	69.7%	87.8%		
Average F_1 score		80.5%						

Table 3: Confusion matrix of all the activities when (C, γ) = (2^5, 2^0) in Figure 7A. This matrix shows all the prediction results for all activities: [a]SR: sedentary rest, [b]SC: sedentary chewing, [c]W: walking, [d]CW: chewing while walking, [e]ST: sedentary talking, [f]SW: sedentary wink.

Predicted activity	Actual activity			Total	Precision
	[a]C	[b]PA	[c]SW		
C	2898	162	26	3086	93.9%
PA	201	4404	200	4805	91.7%
SW	21	114	1334	1469	90.8%
Total	3120	4680	1560	9360	
Recall	92.9%	94.1%	85.5%		92.3%
F_1 score	93.4%	92.9%	88.1%		
Average F_1 score		91.4%			

Table 4: Confusion matrix of all the re-defined activities when (C, γ) = (2^5, 2^0) in Figure 7B. This matrix shows all the prediction results for all re-defined activities: [a]CH: chewing, [b]PA: physical activity, [c]SW: sedentary wink.

Copyright © 2018 Journal of Visualized Experiments

Discussion

In this study, we first proposed the design and manufacturing process of glasses that sense the patterns of food intake and physical activities. As this study mainly focused on the data analysis to distinguish the food intake from the other physical activities (such as walking and winking), the sensor and data acquisition system required the implementation of mobility recording. Thus, the system included the sensors, the MCU with wireless communication capability, and the battery. The proposed protocol provided a novel and practical way to measure the patterns of temporalis muscle activity due to the food intake and wink in a non-contact manner: the tools and methodologies to easily detect the food intake in daily life without any cumbersome equipment were described.

There are important considerations for the procedure of manufacturing the glasses. The temple parts should be designed to integrate the PCB modules fabricated in step 1.2 as shown in **Figure 4B** and **Figure 4C**. The load cell should be placed so that it is pressed by a support bolt at a support plate of the head piece when equipped as illustrated in the top view of the hinge part in **Figure 5**. In step 2.4, the degree of bending of the glasses temple does not need to be rigorous, as the purpose of the curvature is to increase a form factor to better fit the glasses on a subject's head. Be careful, however, as excessive bending will prevent the temples from touching the temporalis muscle, which would make it impossible to collect significant patterns.

To obtain reliable data reflecting the different head sizes and shapes of subjects, two versions of the glasses were provided by varying the length of the head piece and the temples. In addition, by utilizing the support volts to fine-tune the wear-ability, we could adjust the tightness of the glasses. Thus, the data collected through the various glasses, subjects, and wearing-conditions can reflect intra- and inter-individual variability and different form factors.

In the user study, the subject took off the glasses during the break, and wore them again when the recording block restarted. This action prevented the data from overfitting to a specific wearing condition because it changed the wearing conditions (*e.g.*, left-and-right balance, preload on the load cells, contact area with the skin, *etc.*) every time the subject re-wore the glasses.

According to an earlier study of chewing frequency, the chewing activity mainly ranged from 0.94 Hz (5^{th} percentile) to 2.17 Hz (95^{th} percentile)[26]. Thus, we set the frame size to 2 s so that a frame contains multiple chewing activities. This frame size is also suitable for containing the one or more walking cycles, which generally range from 1.4 Hz to 2.5 Hz[27]. We conducted the walking activity at a speed of 4.5 km/h on a treadmill because the normal walking speed varies from 3.3 km/h to 6.5 km/h[27,28]. The hop size in **Figure 6** was determined from the recorded wink data where subjects were informed to wink at 3-s intervals. We also filtered the data with the cutoff frequency of 10 Hz, because we found from our previous study that signals over 10 Hz had no significant information on chewing detection[22].

Because the system has two load cells on both sides, it is possible to distinguish the left and right events of the chewing and wink, as proved in our previous study[22]. However, unlike the previous study, the aim of this study was to demonstrate that the system could effectively separate food intake from the physical activities. If the data are sufficiently accumulated through the user study, then further research on the left and right classification can be conducted, utilizing the correlation features included in the feature vector. On the other hand, it is difficult to distinguish between the sedentary activity and walking within the system. Further modifications to the system can provide detailed classification of the food intake, like eating while sitting and eating on the move, with a high accuracy. This can be implemented through a sensor fusion technique by adding an inertial measurement unit (IMU) to the system[18]. If so, the system can track the energy expenditure and the energy intake simultaneously. We believe that our approach provides practical and potential ways for detection of food intake and physical activities.

Estimation of energy intake is a crucial goal of research on dietary monitoring, and for example, can be analyzed by classifying the type of food, and then converting it into calories from predefined caloric information. A recent study suggested a method of classifying food types using food images and deep learning algorithms[14]. However, it is difficult to separate the food types with the force sensors used in this study; the addition of an image sensor to the front of the device could recognize the food types through image processing and machine learning techniques, and thus classify the food types. Through this sensor fusion technique with the force and image sensors, the potential of this study is application toward general dietary monitoring applications.

Disclosures

The authors have nothing to disclose.

Acknowledgements

This work was supported by Envisible, Inc. This study was also supported by a grant of the Korean Health Technology R&D Project, Ministry of Health & Welfare, Republic of Korea (HI15C1027). This research was also supported by the National Research Foundation of Korea (NRF-2016R1A1A1A05005348).

References

1. Sharma, A.M., Padwal, R. Obesity is a sign-over-eating is a symptom: an aetiological framework for the assessment and management of obesity. *Obes Rev.* **11** (5), 362-370 (2010).
2. Pi-Sunyer, F.X. *et al.* Clinical guidelines on the identification, evaluation, and treatment of overweight and obesity in adults. *Am J Clin Nutr.* **68** (4), 899-917 (1998).
3. McCrory, P., Strauss, B., Wahlqvist, M.L. Energy balance, food intake and obesity. *Exer Obes London.* (1994).

Copyright © 2018 Journal of Visualized Experiments

4. Albinali, F., Intille, S., Haskell, W., Rosenberger, M. In: *Proceedings of the 12th ACM international conference on Ubiquitous computing.* ACM, 311-320 (2010).
5. Bonomi, A., Westerterp, K. Advances in physical activity monitoring and lifestyle interventions in obesity: a review. *Int J Obes.* **36** (2), 167-177 (2012).
6. Jung, S., Lee, J., Hyeon, T., Lee, M., Kim, D.H. Fabric-Based Integrated Energy Devices for Wearable Activity Monitors. *Adv Mater.* **26** (36), 6329-6334 (2014).
7. Fulk, G.D., Sazonov, E. Using sensors to measure activity in people with stroke. *Top Stroke Rehabil.* **18** (6), 746-757 (2011).
8. Makeyev, O., Lopez-Meyer, P., Schuckers, S., Besio, W., Sazonov, E. Automatic food intake detection based on swallowing sounds. *Biomed Signal Process Control.* **7** (6), 649-656 (2012).
9. Päßler, S., Fischer, W. Food intake activity detection using an artificial neural network. *Biomed Tech (Berl).* (2012).
10. Passler, S., Fischer, W.-J. Food intake monitoring: Automated chew event detection in chewing sounds. *IEEE J Biomed Health Inform.* **18** (1), 278-289 (2014).
11. Kadomura, A. *et al.* In: *CHI'13 Extended Abstracts on Human Factors in Computing Systems.* ACM, 1551-1556 (2013).
12. Fontana, J.M., Farooq, M., Sazonov, E. Automatic ingestion monitor: A novel wearable device for monitoring of ingestive behavior. *IEEE Trans Biomed Eng.* **61** (6), 1772-1779 (2014).
13. Shen, Y., Salley, J., Muth, E., Hoover, A. Assessing the Accuracy of a Wrist Motion Tracking Method for Counting Bites across Demographic and Food Variables. *IEEE J Biomed Health Inform.* (2016).
14. Farooq, M., Sazonov, E. In: *International Conference on Bioinformatics and Biomedical Engineering.* Springer, 464-472 (2017).
15. Grigoriadis, A., Johansson, R.S., Trulsson, M. Temporal profile and amplitude of human masseter muscle activity is adapted to food properties during individual chewing cycles. *J Oral Rehab.* **41** (5), 367-373 (2014).
16. Strini, P.J.S.A., Strini, P.J.S.A., de Souza Barbosa, T., Gavião, M.B.D. Assessment of thickness and function of masticatory and cervical muscles in adults with and without temporomandibular disorders. *Arch Oral Biol.* **58** (9), 1100-1108 (2013).
17. Standring, S. *Gray's anatomy: the anatomical basis of clinical practice.* Elsevier Health Sciences (2015).
18. Farooq, M., Sazonov, E. A novel wearable device for food intake and physical activity recognition. *Sensors.* **16** (7), 1067 (2016).
19. Zhang, R., Amft, O. In: *Proceedings of the 2016 ACM International Symposium on Wearable Computers.* ACM, 50-52 (2016).
20. Farooq, M., Sazonov, E. Segmentation and Characterization of Chewing Bouts by Monitoring Temporalis Muscle Using Smart Glasses with Piezoelectric Sensor. *IEEE J Biomed Health Inform.* (2016).
21. Huang, Q., Wang, W., Zhang, Q. Your Glasses Know Your Diet: Dietary Monitoring using Electromyography Sensors. *IEEE Internet of Things Journal.* (2017).
22. Chung, J. *et al.* A glasses-type wearable device for monitoring the patterns of food intake and facial activity. *Scientific Reports.* **7**, 41690 (2017).
23. Cristianini, N., Shawe-Taylor, J. *An introduction to support Vector Machines: and other kernel-based learning methods.* Cambridge University Press Cambridge (2000).
24. Giannakopoulos, T., Pikrakis, A. *Introduction to Audio Analysis: A MATLAB Approach.* Academic Press (2014).
25. Chang, C.-C., Lin, C.-J. LIBSVM: a library for support vector machines. *ACM Trans Intell Syst Technol.* **2** (3), 27 (2011).
26. Po, J. *et al.* Time-frequency analysis of chewing activity in the natural environment. *J Dent Res.* **90** (10), 1206-1210 (2011).
27. Ji, T. Frequency and velocity of people walking. *Struct Eng.* **84** (3), 36-40 (2005).
28. Knoblauch, R., Pietrucha, M., Nitzburg, M. Field studies of pedestrian walking speed and start-up time. *Transp Res Red.* (1538), 27-38 (1996).

Copyright © 2018 Journal of Visualized Experiments