

Article

Augmenting Ear Accessories for Facial Gesture Input Using Infrared Distance Sensor Array [†]

Kyosuke Futami ^{1,2,*} , Kohei Oyama ³ and Kazuya Murao ^{1,4} 

¹ Graduate School of Information Science and Engineering, Ritsumeikan University, 1-1-1 Nojihigashi, Kusatsu 525-8577, Shiga, Japan; murao@cs.ritsumei.ac.jp

² Digital Spirits Teck, Kusatsu 525-8577, Shiga, Japan

³ Nara Institute of Science and Technology, Ikoma 630-0192, Nara, Japan; oyama.kohei.ol8@is.naist.jp

⁴ Strategic Creation Research Promotion Project (PRESTO), Japan Science and Technology Agency (JST), 4-1-8 Honmachi, Kawaguchi 332-0012, Saitama, Japan

* Correspondence: futami@fc.ritsumei.ac.jp

[†] This paper is an extension version of the conference paper: Futami, K.; Oyama, K.; Murao, K. A Method to Recognize Facial Gesture Using Infrared Distance Sensor Array on Ear Accessories. In Proceedings of the 23rd International Conference on Information Integration and Web Intelligence, Linz, Austria, 29 November–1 December 2021; Association for Computing Machinery: New York, NY, USA, 2021; pp. 650–654.

Abstract: Simple hands-free input methods using ear accessories have been proposed to broaden the range of scenarios in which information devices can be operated without hands. Although many previous studies use canal-type earphones, few studies focused on the following two points: (1) A method applicable to ear accessories other than canal-type earphones. (2) A method enabling various ear accessories with different styles to have the same hands-free input function. To realize these two points, this study proposes a method to recognize the user's facial gesture using an infrared distance sensor attached to the ear accessory. The proposed method detects skin movement around the ear and face, which differs for each facial expression gesture. We created a prototype system for three ear accessories for the root of the ear, earlobe, and tragus. The evaluation results for nine gestures and 10 subjects showed that the F-value of each device was 0.95 or more, and the F-value of the pattern combining multiple devices was 0.99 or more, which showed the feasibility of the proposed method. Although many ear accessories could not interact with information devices, our findings enable various ear accessories with different styles to have eye-free and hands-free input ability based on facial gestures.

Keywords: ear accessories; infrared distance sensor; photo reflector; skin movement; facial gesture recognition; hands-free input interface; earable



Citation: Futami, K.; Oyama, K.; Murao, K. Augmenting Ear Accessories for Facial Gesture Input Using Infrared Distance Sensor Array. *Electronics* **2022**, *11*, 1480. <https://doi.org/10.3390/electronics11091480>

Academic Editors: Hsiang-Chieh Chen, Yi-Zeng Hsieh and Chih-Hsien Hsia

Received: 21 March 2022

Accepted: 27 April 2022

Published: 5 May 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In recent years, it has become important to provide hands-free input methods for information devices. This is because, as information devices are used in all situations, the number of scenarios in which it is difficult to use the hand for input operations has increased. Scenes in which the user's hand is used for purposes other than operating information devices [1] include scenes of holding luggage in the hand, walking, writing with a pen, playing a musical instrument, participating in sports and cooking. Another scenario is that using the hand-input method strains the hand and arm. For example, long-term aerial hand input may cause arm fatigue and pain [2,3]. Although many hand-based input methods (e.g., touch input, aerial hand gesture input and controller operation input) are still being used, a hands-free input method effectively supports such a scenario.

Many studies have proposed using ear accessories, such as earphones, to provide a hands-free input method. As the shape and wearing of ear accessories are socially acceptable, devices that use them as controllers for hands-free input are likely to be accepted

by society. Gesture input methods based on intentional face movements using sensors (e.g., barometers, microphones, infrared distance sensors and electrodes) mounted on canal-type earphones have been proposed in previous studies [4–10]. Such gesture input methods based on facial movements using ear accessories is effective as a simple input method in scenarios where the number of input contents is not large and can be used even in situations where it is difficult to use the hands-free input method using voice recognition (e.g., scenes with a lot of environmental voices and a high psychological load of speaking).

Although there have been many studies on hands-free input methods using canal-type earphones, few have focused on the following two points. (1) Hands-free input methods that can be applied to ear accessories other than canal-type earphones for presenting sound information and (2) Hands-free input methods that can apply similar input methods to different ear accessories. The explanation is as follows. Canal-type earphones are not used except when presenting sound information. Such earphones that close the ear hole throughout daily life are often inconvenient from multiple perspectives (e.g., inhibition of environmental perception, poor hygiene of ear hole and the possibility of hearing loss). On the other hand, ear accessories come in various styles and shapes that can be worn. Ear accessories, such as fashionable earrings and earphones, come in various styles that are attached to different parts of the ear, such as the earlobe, tragus, helix and the root of the ear. Other than ear accessories, other accessories, such as glasses-type devices, also have parts that contact the ears. If the hands-free input method can be performed using such various ear accessories, the operation scene of the information device can be expanded, which is convenient. Furthermore, it would be convenient if a hands-free input method could apply similar input methods to different ear accessories and allow the user to select which ear accessories to use and perform similar input methods from any ear accessory.

In this study, we focus on the above two points and propose a method recognizing the gesture of the user's facial expression by using the infrared distance sensor attached to the ear accessory. When users move their face (e.g., the cheeks or chin), the skin inside and around their ear canal physically moves as does the distance between their ears and face. The proposed method detects skin movement by measuring the difference in the distance between the skin and the infrared distance sensor attached to the ear accessory. Then, based on the skin movement, the gestures of the facial expressions are recognized.

In Evaluation 1, we evaluated 10 subjects and nine types of gestures to verify the feasibility of the proposed method. In Evaluation 2, we verified whether the recognition accuracy of the proposed method can be improved by reducing the number of gestures. In Evaluation 3, we verified whether the user needs to prepare gesture learning data for each individual when using the proposed method.

We published the concept of the proposed method in the work-in-progress paper [11]. In this paper, we improve Evaluation 1 and add Evaluation 2 and Evaluation 3. Although Evaluation 1 was contained in the previous paper, this paper improved Evaluation 1 by increasing the number of subjects from 5 to 10 and adding the analysis of the recognition accuracy for each gesture. We also add a section of related research on skin-movement sensing methods using infrared distance sensors and hands-free input methods.

2. Related Research

2.1. Gesture Recognition Method Using Ear Accessories

The gesture recognition methods using ear accessories often use canal-type earphones. Manabe et al. [9] proposed an earphone that recognizes gaze input gestures using an EOG sensor. Amesaka et al. [8] proposed a method that recognizes the gestures of facial expressions by using active acoustic sensing and the changes of shape of the ear canal. Hand-input methods using the canal type earphone include EarTouch [12], which recognizes gestures by using an infrared distance sensor attached to the earphone and the changes in the shape of the ear that users pull with their hands.

CanalSense [4] is a method that recognizes the movements of the face by using a barometer attached to earphones. EarFieldSensing [6] is a method that recognizes gestures

of the movements of the face by using electrodes (EMG) attached to earphones and the changes in the electric field of the ear canal. Taniguchi et al. [7] proposed a method that recognizes the movement of the tongue by using an infrared distance sensor attached to the tip of a canal-type earphone and the movement of the bottom of the ear canal.

Bedri et al. [5] proposed a method that recognizes the movement of the jaw by using an infrared distance sensor attached to the tip of a canal-type earphone. An input method that uses canal-type earphones to acquire sound using a small muscle in the middle ear called the tensor tympani muscle has been proposed [10]. There is an ear-worn device with which users use their ears as a hand-held input controller [13]. In addition, there is the ready-made product (DashPro) that recognizes head gestures using earphones with an acceleration sensor is implemented in a canal-type earphone.

These previous studies have shown that the face and head gesture recognition method using ear accessories is effective as a hands-free input method. Although there are many studies on hands-free input methods using canal type earphones, canal-type earphones are not used except when presenting sound information and such earphones that close the ear hole throughout daily life are often inconvenient for particular people and situations. In this study, we propose a method that can apply the hands-free input function to ear accessories other than canal type earphones. In addition, we propose a method that can be applied to different types of ear accessory devices. If the hands-free input method can be performed using various ear accessories, the operation scene of the information device can be expanded, which is convenient.

2.2. Skin Movement Sensing Method Using an Infrared Distance Sensor

Many methods have been proposed, similar to ours, to recognize skin movement in each part of the body using an infrared distance sensor attached to the wearing. Many studies have been conducted that use infrared distance sensors in conjunction with eyeglasses. For recognizing the direction and movement of the eyeball, Futami et al. [14] proposed a method by sensing the movement of the skin around the eye based on an infrared distance sensor attached to the inside of glasses. Using an infrared distance sensor attached to a position that detects a movement around the cheeks and outer corners of the eyes, Fukumoto et al. [15] proposed a method to recognize a smile in multiple stages.

Masai et al. [16] proposed using infrared distance sensors on glasses to detect eye and cheek movements to recognize eight facial expressions in everyday life, such as smiles and surprises. Masai et al. [17] also proposed using an infrared distance sensor attached to glasses to perform gesture input by rubbing the cheeks with the hands. In addition, methods to recognize both intentional blink gestures and natural blinks have also been proposed (Google Glass [18] and Dual Blink [19]). Dual Blink [19] also has a physical actuation to encourage blinking, such as hitting the eyelids with an air cannon. Furthermore, some studies have shown that the infrared distance sensor is suitable for wearable devices that are continuously used from multiple perspectives (e.g., power consumption) [16,19].

There is also a method that applies an infrared distance sensor to a head-mounted display. Yamashita et al. [20] proposed a method that recognizes touch gestures based on the movement of the cheeks by moving the skin of the cheeks by hand. Suzuki et al. [21] proposed a method to map the movement of the skin of the user's face obtained by the infrared distance sensor to the avatar's facial expression in the virtual space.

As mentioned earlier, some methods attach an infrared distance sensor to the earphone and recognize input gestures (e.g., pulling the ear [12], tongue movement [7]). There is also a method that applies an infrared distance sensor to a mouthpiece to recognize tongue gestures [22]. Ogata et al. [23] applied an infrared distance sensor to the ring's inside to recognize finger orientation and motion gestures. Fukui et al. [24] applied an infrared distance sensor to a wristband to recognize hand-shaped gestures. Matsui et al. [25] applied an infrared distance sensor to the inside of an eye mask to recognize the quality of sleep from eye movements.

These previous studies have shown that the infrared distance sensor can recognize skin movements with high accuracy and robustness and that the infrared distance sensor is suitable for wearable devices that are always used. From these previous studies, the infrared distance sensor is considered to be suitable for the proposed method.

2.3. Hands-Free Input Method

Many recognition methods have been proposed for hands-free input. The input method of speech recognition is common in devices equipped with a microphone. This input method is capable of linguistic input and detailed input, and it can be used by people who can handle language. Examples of utilization of this method are for text input [26] and navigation [27]. The disadvantages of this method are that the recognition accuracy of the method decreases in situations where the environmental sound is noisy [28] and that the psychological load of using the method in public places increases from the viewpoint of privacy and embarrassment. In addition, when the voice processing technology requires advanced data processing, data communication to an advanced processor or external device is required.

Other than speech recognition, there are many methods to recognize the movement of each part of the body that can be used for hands-free input. As mentioned above, recognition methods of facial parts (e.g., jaw, mouth, cheek and tongue) are used for hands-free input. Since gaze can also reflect the user's intention [29], gaze can be used for a hands-free gesture input method [9,30]. There are also methods that use the movement of the head, such as a method that can adjust the input value by rotating or tilting the user's head (HeadTurn [31]), a method to turn pages when browsing (HeadPager [32]), a method to operate the cursor (e.g., desktop device [33,34], mobile device [35]), and a method to select a target [36]. There are methods for gesture input that combine multiple factors, such as head movement and gaze [37,38]. In addition, postures can be used. For example, there is a method for navigation input that uses the inclination of wrist [39], head [40] and torso [41–44].

These previous studies have shown that various hands-free input methods expand the usage scenarios of information devices and make the use of information devices comfortable.

3. Method

This section describes that a method to recognize the gesture of the user's facial expression by using the infrared distance sensor attached to the ear accessory.

3.1. Flow of Proposed Method

The flow up to gesture recognition of the proposed method is shown in Figure 1. The ear accessories contain several infrared distance sensors. When users move their face (e.g., the cheeks or chin), the skin inside and around their ear canal physically moves, as does the distance between the ears and face. The temporomandibular joint moves when the mandibular condyle moves, as does the skin inside the ear hole (e.g., the ear canal) [8]. This skin movement is acquired using the proposed method based on the change in distance from the infrared distance sensor to the skin. Then, machine learning is applied to multiple sensor values to recognize facial gestures.

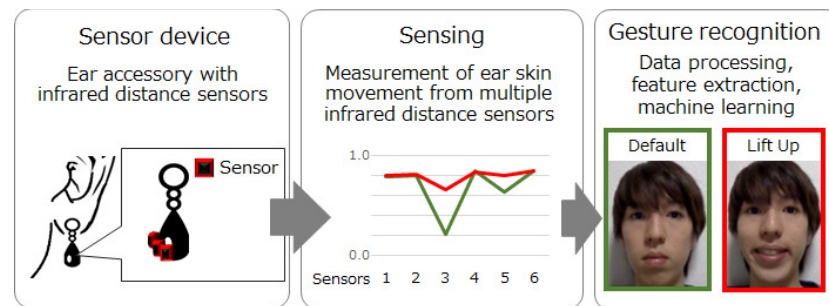


Figure 1. Flowchart of the proposed method. Reprinted/adapted with permission from Ref. [11] 2021, ACM.

3.2. Ear Accessory Design

The following three types of ear accessories were designed based on the proposed method. Since the shape and position of the ear differ to some extent from person to person, it was difficult to design a device that places sensors in the same position for everyone. In order to recognize skin movement regardless of these individual differences, multiple infrared distance sensors were attached to the entire ear accessory. Therefore, the detailed sensor positions may vary from person to person. Since it is difficult to describe which part the sensor is hitting, the figures show a certain range of skin area that is assumed to be hit by the sensor.

3.2.1. Ear-Root-Mounted Device

This device is worn around the root of the ear. Figure 2 shows the design drawing. This style is intended for application to ear accessories and other attachments (e.g., earrings, earphones and glasses) that have a part that touches the root of the ear (i.e., an ear hook part). Seven infrared distance sensors are placed on the side of the ear hook part, and seven infrared distance sensors are placed on the back of the ear hook part. The movement of the ear and the movement of the back of the ear are detected by this device. It is assumed that the ear movement occurs through the muscles that move the ear (e.g., anterior auricular muscle, superior auricular muscle and posterior auricular muscle) in conjunction with the facial expression movement.

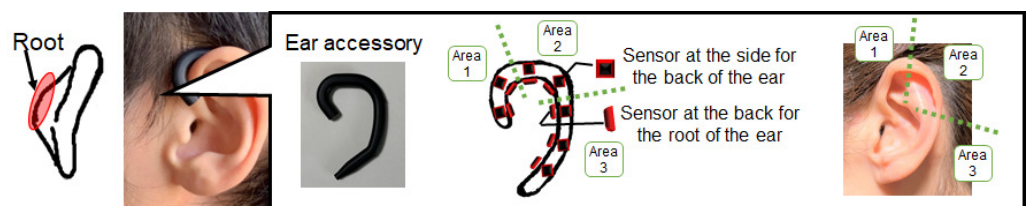


Figure 2. Ear-root-mounted device. Reprinted/adapted with permission from Ref. [11] 2021, ACM.

3.2.2. Earlobe-Mounted Device

This device is worn on the earlobe. Figure 3 shows the design drawing. This is intended for application to ear accessories that are attached to the earlobe. Two infrared distance sensors are installed on the temporomandibular joint side, and two infrared distance sensors are installed on the cranial side. This device senses the change in distance between the earlobe and the vicinity of the temporomandibular joint. This device also senses the change in distance between the earlobe and the vicinity of the head. The temporomandibular joint movement and the ear movement as facial expressions move are thought to be the causes of this distance change.

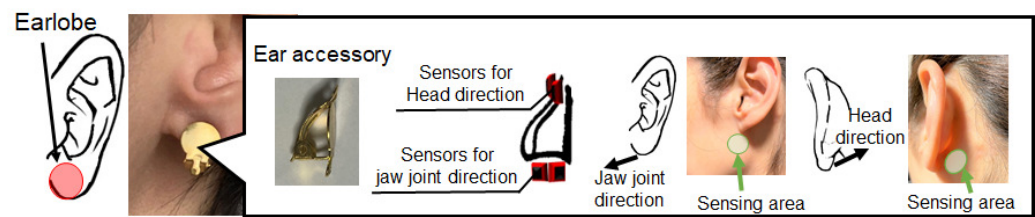


Figure 3. Earlobe-mounted device [11].

3.2.3. Tragus-Mounted Device

This device is attached to the tragus. Figure 4 shows the design drawing. This is intended for application to ear accessories and earphones attached to the tragus. Three infrared distance sensors are installed toward the side of the ear hole (e.g., the ear canal). This device senses the change in distance between the back of the tragus and the inner side of the ear canal. It is assumed that the movement of the tragus and the inner side of the ear hole corresponds to the movement of the facial expression. For example, the movement of each part of the face (e.g., temporomandibular joint) and facial expression is reflected in the change in the inside of the ear hole [8].

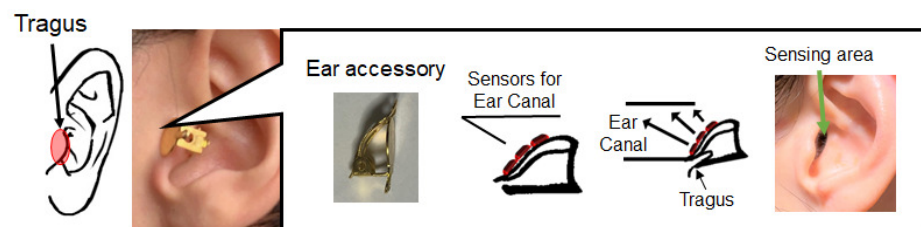


Figure 4. Tragus-mounted device. Reprinted/adapted with permission from Ref. [11]. 2021, ACM.

3.3. Sensor Selection and Method Advantages

The proposed method made use of an infrared distance sensor. We chose the infrared distance sensor for this study because previous research has shown that it is appropriate for wearable devices that are constantly used and that it can recognize skin movements robustly. The proposed method meets the following criteria.

(1) Cost of data processing and manufacturing: Due to the lightweight of the sensor data, the processor capacity and energy consumption for processing the sensor data are low. These lead to cost reduction and size reduction of batteries and processors. These allow the sensor device to be used for a long time. In addition, this system can be manufactured at a low cost because it can be configured only with an inexpensive infrared distance sensor and processor.

(2) Social acceptability and availability of always-on use: Ear accessories are likely to be accepted by society as controller devices for hands-free input because the shape and wearing of the accessories are socially accepted. The proposed method can be used in various everyday situations through ear accessories.

(3) Others: First, the proposed method can be used without obstructing the ear holes (i.e., ear-free). Although most hands-free input methods using ear-worn devices use canal-type earphones that close the ear hole, devices that close the ear hole throughout daily life are often inconvenient from multiple perspectives (e.g., the inhibition of environmental perception, poor hygiene of ear hole and the possibility of hearing loss). Secondly, the proposed method is expected to be adaptable and robust in recognizing skin movements. Previous research has shown that applying machine learning to multiple infrared distance sensor data makes it possible to recognize skin movements of various gestures adaptively and robustly. Thirdly, ear accessories equipped with the proposed method can be used in conjunction with other wearable devices because they do not interfere with wearing other wearable devices (e.g., glasses and earphones).

4. Implementation

The proposed method's prototype system was implemented. Figure 5 depicts the system configuration. The system consists of a sensor device, a personal computer and software for data recording and gesture recognition.

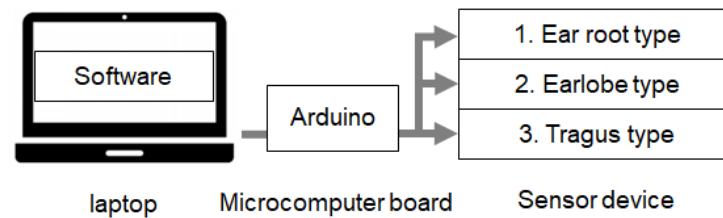


Figure 5. Prototype system configuration.

4.1. Sensor Device

Wearing sensor devices are depicted in Figure 6. The base of the ear-root-mounted device was made of silicone ear hooks. As the base of the tragus-mounted device and the earlobe-mounted device, ear accessories that sandwich the parts of the ear were used—the TPR-105F infrared distance sensor from SANYO Electric Co. was used. Two devices were created for each of the three devices to be worn on the left and right ears. The dimensions of the infrared distance sensor were 2 mm (width), 2 mm (length) and 1 mm (thickness).

The sensor measures a straight line distance from the front to the object. Figure 7 shows the characteristics of skin surface reflection captured by a sensor. We measured the sensor output value from a sensor that changes in accordance with the distance between the sensor and the skin surface. Five samples were collected at each position. The horizontal axis shows the distance between the sensor and the skin, and the vertical axis shows the normalized output value.

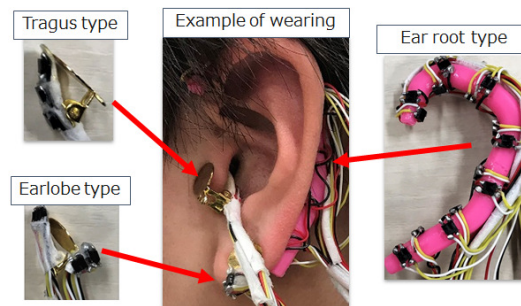


Figure 6. Example of wearing sensor devices. Reprinted/adapted with permission from Ref. [11] 2021, ACM.

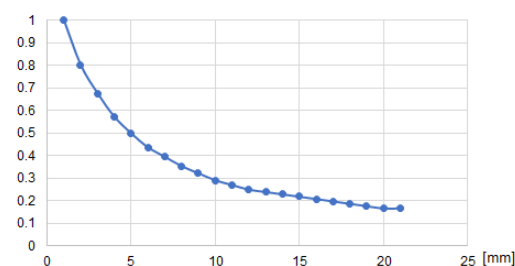


Figure 7. Characteristics of infrared distance sensor. The horizontal axis shows the distance between the sensor and the skin, and the vertical axis shows the normalized sensor output value.

4.2. Software

The sensor data were transmitted and saved in the laptop computer via a microprocessor of Arduino. The sampling frequency of the sensor data was set to 10 [HZ]. The sensor data were smoothed every three samples. In addition, the sensor values were normalized for each sensor. The instantaneous absolute value of the sensor was taken as the feature amount of sensor data applied to machine learning. SVM (RBF kernel, $C = 10$ and $\gamma = 1.0$) was used as a machine-learning classifier. We used SVM based on previous studies using multiple infrared distance sensors on wearable devices (e.g., glasses [16], wristbands [45]). These previous studies recognized skin movements in real-time using SVM. Python was used to create the software.

5. Evaluation 1

We verified the proposed method's recognition accuracy for facial expression gestures. The proposed method's recognition accuracy for nine different types of gestures was evaluated in this experiment. There were ten subjects in total. The subjects were Asian (seven males and three females, and the average age was 22 years (min: 20 and max: 23)). Subjects were selected through an open recruitment process within the university. There were no selection criteria for subjects, but the subjects took part in the experiment without ear accessories.

5.1. Gesture Content

The nine types of gestures shown in Figure 8 were used. These were determined based on previous studies [4,6,8]. The outline of gestures was ones that moving the facial parts of the chin, mouth, cheeks and eyes. The details are as follows. The Default is expressionless. Open Mouth is a gesture that opens the mouth wide. Slide Jaw Right is a gesture that moves the jaw to the right. Slide Jaw Left is a gesture that moves the jaw to the left. Lift Mouth Corner is a gesture that raises both cheeks. Blow Up Right is a gesture that swells the right cheek. Blow Up Left is a gesture that swells the left cheek. Lift Cheek Right and Wink Right is a gesture that raises the right cheek while closing the right eye. Lift Cheek Left and Wink Left is a gesture that raises the left cheek while closing the left eye.

The reason for these gestures is considered appropriate from the viewpoint of eye-free and public usability. The reason was based on the previous study [8]. The selected gesture is eye-free. Gestures should not harm the user's visual information. Examples of gestures that impair the user's visual information include gestures that close both eyes and gestures that turn the user's head sideways. When these gestures are made, the direction of the field of vision changes, or the field of vision is blocked during the gesture. In addition, the selected gesture is suitable for public use. Gestures should not make people feel uncomfortable. For example, the tongue-out gesture selected in previous studies may be noticed by others and make some people embarrassed to make a gesture.

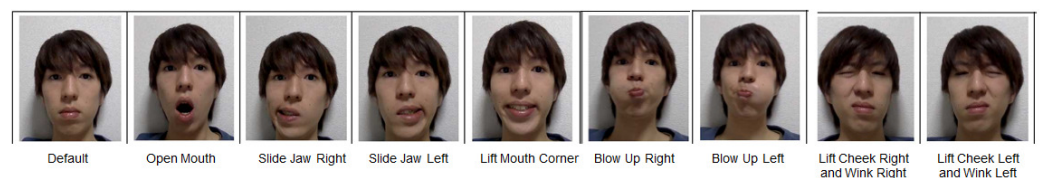


Figure 8. Types of gestures. Reprinted/adapted with permission from Ref. [11] 2021, ACM.

5.2. Data Acquisition Flow

The following is how the data were collected. Following an explanation of the experiment's content, the subjects wore all three types of devices. The gestures to be performed were explained and practiced. Then, they sat in the chair and performed the gesture task. One trial of the gesture task was as follows. First, the subject was in a natural state (i.e., Default in Figure 8) as the initial state. Next, the subject performed the same gesture as the model gesture displayed in front of them. The subject held the gesture for one second.

Finally, the subject stopped the gesture and returned to the initial state. This single trial was conducted for nine types of gestures. The order of the gestures within this single trial was randomized. Ten trials were performed. The device was not removed during these trials.

5.3. Verification

The recognition accuracy was evaluated for each individual. Since 10 trials were performed for each gesture, 10-fold cross-validation was performed. In other words, verification was performed 10 times for one gesture, using nine trials in the individual as training data and one trial as test data. The data in the same trial with the same gesture was not divided into training data and test data. The data for 0.2 s immediately after the start of the gesture and 0.2 s immediately before the end of the gesture was deleted. The reason for this is that the data may not reach the gesture.

5.4. Result

Figure 9 shows the average value for each device pattern. The following is the device pattern: Device 1 is a tragus-mounted device, Device 2 is an ear-root-mounted device, and Device 3 is an earlobe-mounted device. Device combinations are classified into seven types (i.e., individual use and a combination of each device). Precision is an indicator that shows the percentage of data predicted to be positive that is actually positive. Recall is an indicator that shows the percentage of those predicted to be positive out of those actually positive. The F-value is the harmonic mean of Precision and Recall and is an indicator used for overall evaluation. Figure 10 shows the average F-value for each gesture for each device pattern. Error bars indicate the standard error. Tables A1 and A2 in Appendix A shows raw data.

Tragus-mounted device (Device 1) had an F-value of 0.99 (min = 0.96 and max = 1.00) and the highest recognition accuracy among a single device. The F-values for Ear-root-mounted device (Device 2) and Earlobe-mounted device (Device 3) were 0.95 (min = 0.88 and max = 0.99) and 0.96 (min = 0.90 and max = 1.00). These results show that individual devices are capable of performing adequate recognition. Furthermore, these results show that the recognition accuracy can be improved by combining multiple devices. For example, the F-value of the device pattern consisting of Ear-root-mounted device and Earlobe-mounted device (Device 2 and Device 3) was 0.99 (min = 0.96 and max = 1.00), which was higher recognition accuracy than the pattern in which these two devices are used alone. Some devices have individual differences in recognition accuracy. Tragus-mounted device (Device 1) had high accuracy for all subjects. However, Ear-root-mounted device (Device 2) and Earlobe-mounted device (Device 3) had lower recognition accuracy for some subjects. For example, Subject 5 had an F-value of Ear-root-mounted device (Device 2) of 0.93 and Earlobe-mounted device (Device 3) of 0.90. This can be seen from the results of the min of F-value of Ear-root-mounted device (Device 2) and Earlobe-mounted device (Device 3).

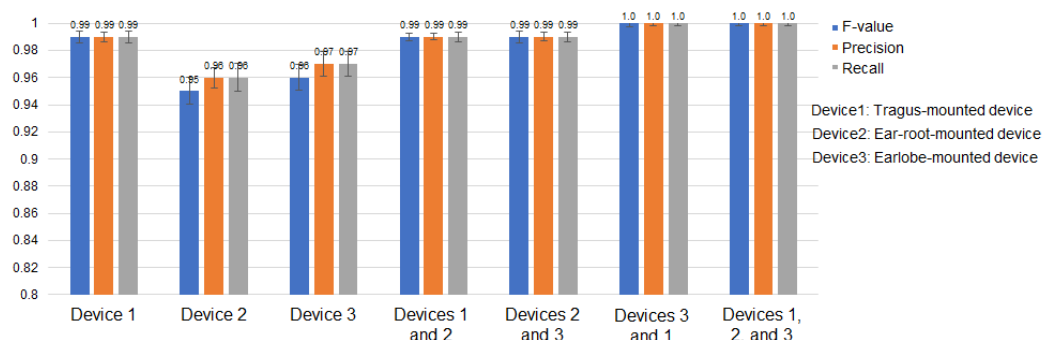


Figure 9. The results of Evaluation 1 using nine gestures. The average F-value for each device pattern. Device 1 is a tragus-mounted device, Device 2 is an ear-root-mounted device, and Device 3 is an earlobe-mounted device. Device combinations are classified into seven types (i.e., individual use and a combination of each device).

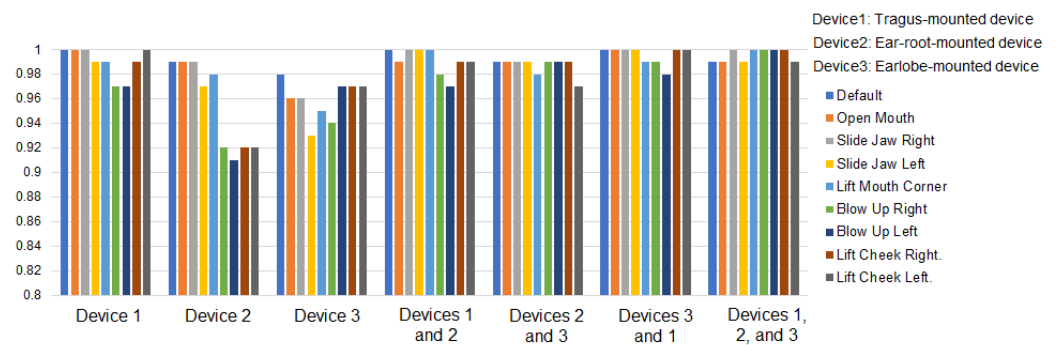


Figure 10. The results of each gesture in Evaluation 1 using nine gestures. The average F-value for each device pattern.

6. Evaluation 2

This evaluation assessed the proposed method's recognition accuracy for the gestures that were chosen while assuming the actual usage of the simple hands-free application. There were six different types of gestures. A player for media content is an example of a hands-free application (e.g., music, video and images). It is sufficient to be able to input about five types of data for these applications to function: play, stop, forward and back. In view of the previous work [8], examining six types of gestures is reasonable. The data set employed was comparable to Evaluation 1. The selected gestures were Default, Open Mouth, Slide Jaw Right, Slide Jaw Left, Lift Mouth Corner and Blow Up Left. The reason for selecting these was because they were eye-free.

Result

Figure 11 shows the average value for each device pattern. Error bars indicate the standard error. Figure 12 shows the average F-value for each gesture for each device pattern. Tables A3 and A4 in Appendix A shows raw data. The recognition accuracy of devices with a relatively low recognition accuracy of Evaluation 1 slightly improved because the number of gestures was reduced. For example, the F-value of Device 2 improved. Although the same gesture was chosen for all subjects in this evaluation, it is assumed that the recognition accuracy for each individual can be improved by selecting a gesture that is appropriate for each subject.

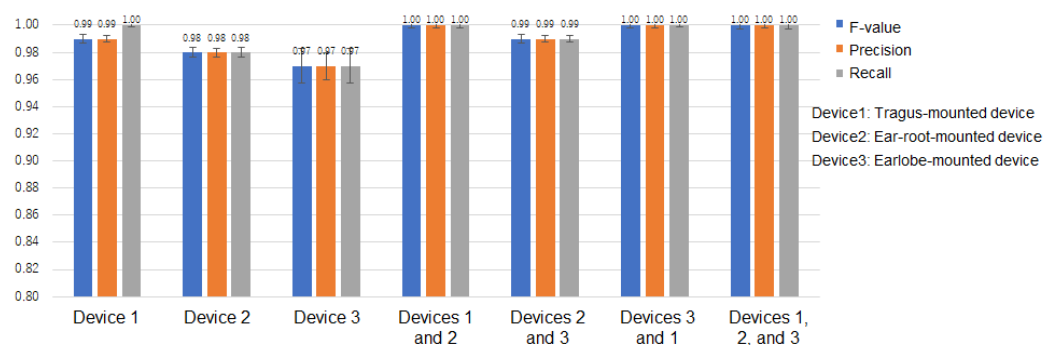


Figure 11. The results of Evaluation 2 using six gestures. The average value for each device pattern.

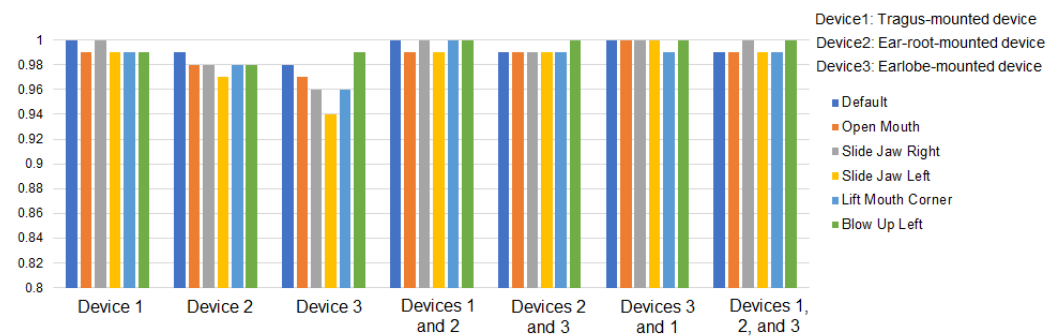


Figure 12. The results of each gesture in Evaluation 2 using six gestures. The average F-value for each device pattern.

7. Evaluation 3

We verified whether the user needs to prepare gesture learning data for each individual when using the proposed method. Leave-one-user-out was the strategy we used. The data of one person was used as test data, while the data of the other nine subjects were used as learning data for all of the subjects. The data set used in Evaluation 1 was the same. As in Evaluation 2, there were six types of gestures.

Result

Figure 13 shows the average value for each device pattern. Error bars indicate the standard error. Figure 14 shows the average F-value for each gesture for each device pattern. Tables A5 and A6 in Appendix A show the raw data. The recognition accuracy was lower than the results of Evaluation 1 and Evaluation 2, which used only the data within the individual. Although the recognition accuracy was improved by combining multiple devices, this recognition accuracy is considered to be insufficient for practical use. From these results, we found that the user needs to prepare their own gesture data in order to use the proposed method.

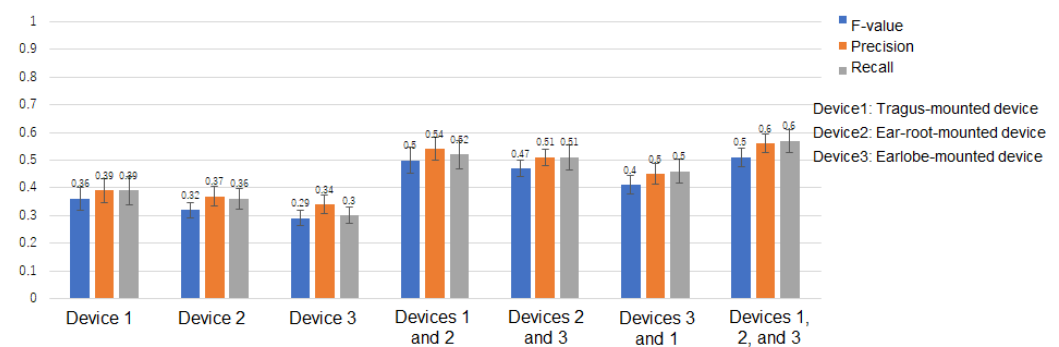


Figure 13. The results of Evaluation 3 testing with leave-one-user-out. The average F-value for each device pattern.

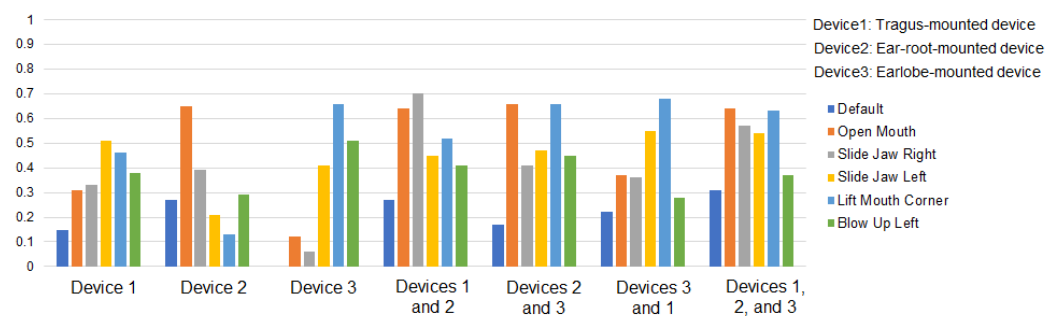


Figure 14. The results of each gesture in Evaluation 3 testing with leave-one-user-out. The average F-value for each device pattern.

8. General Discussion, Limitations and Future Works

8.1. Feasibility of the Proposed Method

The evaluation results showed the feasibility of the proposed method. Previous research using canal-type earphones found that the recognition accuracy of five facial expressions of methods using electric sensing technologies (e.g., EMG) was 90% [6] and recognition accuracy of six facial expressions of methods using active acoustic sensing was 85% accuracy [8]. Compared to the results of these previous studies, the proposed method has a recognition accuracy comparable to or better than the conventional methods.

The evaluation results also confirmed the effectiveness of the proposed method for three types of ear accessories with different styles (e.g., the wearing position and size). We considered that the proposed method can be used for various accessories that have a shape that is similar to the ones verified this time. For example, the proposed method is assumed to be used for glasses since the ear hook part of the glasses has a shape similar to that of the ear root-wearing device implemented in this study. On the other hand, some ear accessories were not verified in this study, such as ear accessories worn on a counter-helix part. Therefore, in the future, we plan to verify whether the proposed method can be applied to ear accessories other than the ones used in this paper.

8.2. Individual Difference That May Influence the Recognition Accuracy

The following factors may influence the recognition accuracy of the proposed method. Verification of these factors with subjects of various attributes is future work. (1) Obstructions between the front of the sensor and the skin may reduce the recognition accuracy of the proposed method. For example, it can be assumed that the recognition accuracy may be reduced if a long-haired user's hair bundle is in front of the sensor or if a large beard is in front of the sensor. (2) Skin color. Recognition accuracy may be reduced for people with very dark skin since infrared distance sensors often have difficulty reacting to black color. (3) Recognition accuracy may be reduced if there is little or no skin movement even when facial expression gestures are performed.

8.3. What to Do if the Mounting Position Is Significantly Misaligned

The proposed method can achieve the same accuracy as the results of this paper as long as the ear accessory is worn in roughly the same position. In this paper, the device's position did not move during the experiment since clips for common ear accessories were used, and such clips had a strong gripping force. However, if the device's attachment position deviates significantly after reattachment, it can be assumed that the recognition accuracy will decrease.

In this case, it is recommended to use the proposed method by acquiring the gesture training data again. As the gesture training data can be acquired in a short time (e.g., six gesture types and five trials need about 1 min 30 s), it can be assumed that the burden on the user is not significant. Since this paper did not verify the case of misalignment of the mounting position, this verification is also future work.

8.4. Challenges Related to the Learning Process of Machine Learning

Machine learning is assumed to work effectively when using data within an individual since personal characteristics (e.g., skin color, skin movement, hair, beard) do not change significantly. However, if personal characteristics change significantly over time, it can be assumed that previous training data cannot be used. In such cases, it is recommended that the training data is reacquired for a few minutes. On the other hand, as Evaluation 3 shows, training data is not available for other users due to individual characteristics differences. It is assumed to be difficult for machine learning to perform recognition regardless of individual characteristics differences. Therefore, it is recommended that only personal data should be used.

8.5. The Use of the Proposed Method in Real Life

In this study, the proposed method's recognition accuracy in real-world situations is not evaluated. It can be assumed that the proposed method can be effectively used in real life under conditions similar to those found in this experiment. For example, if the position of the attached device does not change significantly (e.g., when stationary and walking), the recognition accuracy of the proposed method can be assumed to be high. On the other hand, if the user shifts to the point where the position of the attached device changes significantly (e.g., dance and running), the proposed method's recognition accuracy may decrease. When using the proposed method in conditions where the position of the attached device changes significantly, it is assumed that some ingenuity is required, such as a mechanism for bringing the wearing device into close contact with the body.

8.6. Investigating Recognizable Gestures

Even though we conducted an experiment in which single gestures were performed independently, the proposed method is expected to recognize multiple gestures by combining single gestures. First, it is possible to perform a single continuous gesture, such as moving the jaw once to the left and then moving the jaw once to the right. Following that, it is possible to vary the degree of facial expression during the gesture, such as a gesture that moves the jaw slightly to the left and a gesture that moves the jaw very far to the left. In the future, we will verify whether the proposed method can recognize those various gestures.

9. Conclusions

In this study, we introduced a facial gesture recognition method using multiple infrared distance sensors attached to ear accessories for simple hands-free input using ear accessories. To verify the applicability of the proposed method to ear accessories with a different style, a prototype system was implemented for three ear accessories: the root of the ear, the earlobe and the tragus. The evaluation results for nine facial gestures (e.g., cheek lift and slide jaw) and 10 subjects showed that the F-value of a single device was 95% or more, and the F-value of the pattern combining multiple devices was 99% or more.

Compared to the method using canal-type earphones in the previous study, which had a recognition accuracy of about 90% for five facial gestures and 85% for six facial gestures using canal-type earphones, the proposed method is considered to have a recognition accuracy equivalent to or better than the previous method. These results showed that the proposed method is effective for facial gesture input using ear accessories.

These results also showed that the proposed method can be used for the same facial gesture input across different styles of ear accessories. Although few ear accessories other than canal-type earphones have been verified for hands-free input devices, the proposed method is expected to enable various ear accessories with different styles to have a simple eye-free hands-free input function.

Author Contributions: Conceptualization, K.F., K.O. and K.M.; methodology, K.F., K.O. and K.M.; software, K.O.; validation, K.F. and K.O.; formal analysis, K.F. and K.O.; investigation, K.F. and K.O.; resources, K.F. and K.O.; data curation, K.F. and K.O.; writing—original draft preparation, K.F.

and K.O.; writing—review and editing, K.F. and K.O.; visualization, K.F. and K.O.; supervision, K.F. and K.M.; project administration, K.F. and K.M.; funding acquisition, K.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by JSPS(Japan Society for the Promotion of Science) KAKENHI Grant Number JP19K20330.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki. This study was approved by the research ethics committee of Ritsumeikan University (Permission number: 2021-067).

Informed Consent Statement: Informed consent was obtained from all subjects.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. The results of a single device for each gesture in Evaluation 1.

	Device 1			Device 2			Device 3		
	R	P	F	R	P	F	R	P	F
Default	1	1	1	0.99	0.99	0.99	0.97	1	0.98
Open Mouth	1	1	1	0.99	0.99	0.99	0.97	0.96	0.96
Slide Jaw Right	1	1	1	1	0.99	0.99	0.96	0.97	0.96
Slide Jaw Left	0.99	1	0.99	0.98	0.98	0.97	0.94	0.94	0.93
Lift Mouth Corner	0.99	0.99	0.99	0.98	0.98	0.98	0.96	0.96	0.95
Blow Up Right	0.96	0.98	0.97	0.93	0.93	0.92	0.97	0.94	0.94
Blow Up Left	0.98	0.96	0.97	0.91	0.93	0.91	0.98	0.98	0.97
Lift Cheek Right.	1	0.99	0.99	0.95	0.92	0.92	0.97	0.98	0.97
Lift Cheek Left.	0.99	1	1	0.92	0.94	0.92	0.98	0.97	0.97
Ave.	0.99	0.99	0.99	0.96	0.96	0.95	0.97	0.97	0.96

Table A2. The results of multiple device combinations for each gesture in Evaluation 1.

	Device 1 and 2			Device 2 and 3			Device 3 and 1			Device 1, 2 and 3		
	R	P	F	R	P	F	R	P	F	R	P	F
Default	1.00	1.00	1.00	0.99	1.00	0.99	1.00	1.00	1.00	0.99	1.00	0.99
Open Mouth	0.99	0.99	0.99	0.99	0.99	0.99	1.00	1.00	1.00	0.99	0.99	0.99
Slide Jaw Right	1.00	1.00	1.00	1.00	0.99	0.99	1.00	1.00	1.00	1.00	1.00	1.00
Slide Jaw Left	1.00	1.00	1.00	0.99	0.99	0.99	0.99	1.00	1.00	1.00	0.99	0.99
Lift Mouth Corner	1.00	1.00	1.00	0.99	0.99	0.98	1.00	0.99	0.99	0.99	1.00	1.00
Blow Up Right	0.97	0.99	0.98	0.99	0.99	0.99	0.99	1.00	0.99	1.00	1.00	1.00
Blow Up Left	0.98	0.97	0.97	0.99	0.99	0.99	0.99	0.98	0.98	1.00	1.00	1.00
Lift Cheek Right.	0.99	0.99	0.99	0.99	0.99	0.99	1.00	1.00	1.00	1.00	1.00	1.00
Lift Cheek Left.	0.99	0.99	0.99	0.98	0.97	0.97	1.00	1.00	1.00	1.00	0.99	0.99
Ave.	0.99	0.99	0.99	0.99	0.99	0.99	1.00	1.00	1.00	1.00	1.00	1.00

Table A3. The results of a single device for each gesture in Evaluation 2.

	Device 1			Device 2			Device 3		
	R	P	F	R	P	F	R	P	F
Default	1.00	1.00	1.00	0.98	1.00	0.99	0.97	1.00	0.98
Open Mouth	1.00	0.99	0.99	0.99	0.98	0.98	0.98	0.97	0.97
Slide Jaw Right	1.00	1.00	1.00	1.00	0.97	0.98	0.96	0.97	0.96
Slide Jaw Left	0.99	0.99	0.99	0.98	0.97	0.97	0.94	0.94	0.94
Lift Mouth Corner	0.99	0.99	0.99	0.98	0.98	0.98	0.97	0.97	0.96
Blow Up Left	1.00	0.99	0.99	0.98	0.99	0.98	0.99	1.00	0.99
Ave.	1.00	0.99	0.99	0.98	0.98	0.98	0.97	0.97	0.97

Table A4. The results of multiple device combinations for each gesture in Evaluation 2.

	Device 1 and 2			Device 2 and 3			Device 3 and 1			Device 1, 2 and 3		
	R	P	F	R	P	F	R	P	F	R	P	F
Default	0.99	1.00	1.00	0.99	1.00	0.99	1.00	1.00	1.00	0.99	1.00	0.99
Open Mouth	0.99	0.99	0.99	0.99	0.99	0.99	1.00	1.00	1.00	0.99	0.98	0.99
Slide Jaw Right	1.00	1.00	1.00	1.00	0.99	0.99	1.00	1.00	1.00	1.00	1.00	1.00
Slide Jaw Left	1.00	0.99	0.99	0.99	0.99	0.99	0.99	1.00	1.00	1.00	0.99	0.99
Lift Mouth Corner	1.00	1.00	1.00	1.00	0.99	0.99	1.00	0.99	0.99	1.00	0.99	0.99
Blow Up Left	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Ave.	1.00	1.00	1.00	0.99	0.99	0.99	1.00	1.00	1.00	1.00	1.00	0.99

Table A5. The results of a single device for each gesture in Evaluation 3.

	Device 1			Device 2			Device 3		
	R	P	F	R	P	F	R	P	F
Default	0.16	0.19	0.15	0.29	0.31	0.27	0	0	0
Open Mouth	0.44	0.32	0.31	0.76	0.68	0.65	0.11	0.17	0.12
Slide Jaw Right	0.34	0.35	0.33	0.41	0.44	0.39	0.1	0.04	0.06
Slide Jaw Left	0.59	0.52	0.51	0.2	0.29	0.21	0.39	0.48	0.41
Lift Mouth Corner	0.47	0.55	0.46	0.16	0.12	0.13	0.72	0.69	0.66
Blow Up Left	0.36	0.42	0.38	0.35	0.41	0.29	0.5	0.64	0.51
Ave.	0.39	0.39	0.36	0.36	0.37	0.32	0.3	0.34	0.29

Table A6. The results of multiple device combinations for each gesture in Evaluation 3.

	Device 1 and 2			Device 2 and 3			Device 3 and 1			Device 1, 2 and 3		
	R	P	F	R	P	F	R	P	F	R	P	F
Default	0.3	0.28	0.27	0.17	0.19	0.17	0.21	0.3	0.22	0.27	0.38	0.31
Open Mouth	0.71	0.61	0.64	0.77	0.65	0.66	0.47	0.35	0.37	0.73	0.61	0.64
Slide Jaw Right	0.73	0.74	0.7	0.5	0.46	0.41	0.38	0.4	0.36	0.62	0.63	0.57
Slide Jaw Left	0.48	0.56	0.45	0.44	0.56	0.47	0.68	0.56	0.55	0.66	0.59	0.54
Lift Mouth Corner	0.56	0.56	0.52	0.79	0.66	0.66	0.74	0.77	0.68	0.68	0.7	0.63
Blow Up Left	0.37	0.51	0.41	0.41	0.55	0.45	0.3	0.3	0.28	0.44	0.44	0.37
Ave.	0.52	0.54	0.5	0.51	0.51	0.47	0.46	0.45	0.41	0.57	0.56	0.51

References

- Fejtová, M.; Figueiredo, L.; Novák, P.; Štěpánková, O.; Gomes, A. Hands-Free Interaction with a Computer and Other Technologies. *Univers. Access Inf. Soc.* **2009**, *8*, 277. [[CrossRef](#)]
- Cabral, M.C.; Morimoto, C.H.; Zuffo, M.K. On the Usability of Gesture Interfaces in Virtual Reality Environments. In Proceedings of the 2005 Latin American Conference on Human-Computer Interaction, Cuernavaca, Mexico, 23–26 October 2005; pp. 100–108.

3. Stoakley, R.; Conway, M.J.; Pausch, R. Virtual Reality on a WIM: Interactive Worlds in Miniature. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Denver, CO, USA, 7–11 May 1995; pp. 265–272.
4. Ando, T.; Kubo, Y.; Shizuki, B.; Takahashi, S. Canalsense: Face-Related Movement Recognition System Based on Sensing Air Pressure in Ear Canals. In Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology, Quebec City, QC, Canada, 22–25 October 2017; pp. 679–689.
5. Bedri, A.; Byrd, D.; Presti, P.; Sahni, H.; Gue, Z.; Starner, T. Stick It in Your Ear: Building an in-Ear Jaw Movement Sensor. In Proceedings of the Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers, Osaka, Japan, 7–11 September 2015; pp. 1333–1338.
6. Matthies, D.J.; Strecker, B.A.; Urban, B. Earfieldsensing: A Novel in-Ear Electric Field Sensing to Enrich Wearable Gesture Input through Facial Expressions. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, Denver, CO, USA, 6–11 May 2017; pp. 1911–1922.
7. Taniguchi, K.; Kondo, H.; Kurosawa, M.; Nishikawa, A. Earable TEMPO: A Novel, Hands-Free Input Device That Uses the Movement of the Tongue Measured with a Wearable Ear Sensor. *Sensors* **2018**, *18*, 733. [\[CrossRef\]](#)
8. Amesaka, T.; Watanabe, H.; Sugimoto, M. Facial Expression Recognition Using Ear Canal Transfer Function. In Proceedings of the 23rd International Symposium on Wearable Computers, London, UK, 9–13 September 2019; pp. 1–9.
9. Manabe, H.; Fukumoto, M.; Yagi, T. Conductive Rubber Electrodes for Earphone-Based Eye Gesture Input Interface. *Pers. Ubiquitous Comput.* **2015**, *19*, 143–154. [\[CrossRef\]](#)
10. Röddiger, T.; Clarke, C.; Wolfram, D.; Budde, M.; Beigl, M. EarRumble: Discreet Hands-and Eyes-Free Input by Voluntary Tensor Tympani Muscle Contraction. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, Yokohama, Japan, 8–13 May 2021; pp. 1–14.
11. Futami, K.; Oyama, K.; Murao, K. A Method to Recognize Facial Gesture Using Infrared Distance Sensor Array on Ear Accessories. In Proceedings of the 23rd International Conference on Information Integration and Web Intelligence, Linz, Austria, 29 November–1 December 2021; Association for Computing Machinery: New York, NY, USA; pp. 650–654.
12. Kikuchi, T.; Sugiura, Y.; Masai, K.; Sugimoto, M.; Thomas, B.H. EarTouch: Turning the Ear into an Input Surface. In Proceedings of the 19th International Conference on Human–Computer Interaction with Mobile Devices and Services, Vienna, Austria, 4–7 September 2017; pp. 1–6.
13. Lissermann, R.; Huber, J.; Hadjakos, A.; Nanayakkara, S.; Mühlhäuser, M. EarPut: Augmenting Ear-Worn Devices for Ear-Based Interaction. In Proceedings of the 26th Australian Computer-Human Interaction Conference on Designing Futures: The Future of Design, Sydney, Australia, 2–5 December 2014; pp. 300–307.
14. Futami, K.; Tabuchi, Y.; Murao, K.; Terada, T. A Method to Recognize Eyeball Movement Gesture Using Infrared Distance Sensor Array on Eyewear. In Proceedings of the 23rd International Conference on Information Integration and Web Intelligence, Linz, Austria, 29 November–1 December 2021; Association for Computing Machinery: New York, NY, USA; pp. 645–649.
15. Fukumoto, K.; Terada, T.; Tsukamoto, M. A Smile/Laughter Recognition Mechanism for Smile-Based Life Logging. In Proceedings of the fourth Augmented Human International Conference, Stuttgart, Germany, 7–8 March 2013; pp. 213–220.
16. Masai, K.; Sugiura, Y.; Ogata, M.; Kunze, K.; Inami, M.; Sugimoto, M. Facial Expression Recognition in Daily Life by Embedded Photo Reflective Sensors on Smart Eyewear. In Proceedings of the 21st International Conference on Intelligent User Interfaces, Sonoma, CA, USA, 7–10 March 2016; pp. 317–326.
17. Masai, K.; Sugiura, Y.; Sugimoto, M. Facerubbing: Input Technique by Rubbing Face Using Optical Sensors on Smart Eyewear for Facial Expression Recognition. In Proceedings of the ninth Augmented Human International Conference, Seoul, Korea, 7–9 February 2018; pp. 1–5.
18. Crook, J. The Google Glass Wink Feature Is Real. *TechCrunch*, 9 May 2013.
19. Dementyev, A.; Holz, C. DualBlink: A Wearable Device to Continuously Detect, Track, and Actuate Blinking for Alleviating Dry Eyes and Computer Vision Syndrome. In Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, Maui, HI, USA, 11–15 September 2017; Volume 1, pp. 1–19.
20. Yamashita, K.; Kikuchi, T.; Masai, K.; Sugimoto, M.; Thomas, B.H.; Sugiura, Y. CheekInput: Turning Your Cheek into an Input Surface by Embedded Optical Sensors on a Head-Mounted Display. In Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology, Gothenburg, Sweden, 8–10 November 2017; pp. 1–8.
21. Suzuki, K.; Nakamura, F.; Otsuka, J.; Masai, K.; Itoh, Y.; Sugiura, Y.; Sugimoto, M. Recognition and Mapping of Facial Expressions to Avatar by Embedded Photo Reflective Sensors in Head Mounted Display. In Proceedings of the 2017 IEEE Virtual Reality (VR), Los Angeles, CA, USA, 18–22 March 2017; pp. 177–185.
22. Hashimoto, T.; Low, S.; Fujita, K.; Usuni, R.; Yanagihara, H.; Takahashi, C.; Sugimoto, M.; Sugiura, Y. TongueInput: Input Method by Tongue Gestures Using Optical Sensors Embedded in Mouthpiece. In Proceedings of the 2018 57th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE), Nara, Japan, 11–14 September 2018; pp. 1219–1224.
23. Ogata, M.; Sugiura, Y.; Osawa, H.; Imai, M. IRing: Intelligent Ring Using Infrared Reflection. In Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology, Cambridge, MA, USA, 7–10 October 2012; pp. 131–136.
24. Fukui, R.; Watanabe, M.; Gyota, T.; Shimosaka, M.; Sato, T. Hand Shape Classification with a Wrist Contour Sensor: Development of a Prototype Device. In Proceedings of the 13th International Conference on Ubiquitous Computing, Beijing, China, 17–21 September 2011; pp. 311–314.

25. Matsui, S.; Terada, T.; Tsukamoto, M. Smart Eye Mask: Sleep Sensing System Using Infrared Sensors. In Proceedings of the 2017 ACM International Symposium on Wearable Computers, Maui, HI, USA, 11–15 September 2017; pp. 58–61.
26. He, J.; Chaparro, A.; Nguyen, B.; Burge, R.; Crandall, J.; Chaparro, B.; Ni, R.; Cao, S. Texting While Driving: Is Speech-Based Texting Less Risky than Handheld Texting? In Proceedings of the Fifth International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Eindhoven, The Netherlands, 28–30 October 2013; pp. 124–130.
27. Feng, J.; Sears, A. Using Confidence Scores to Improve Hands-Free Speech Based Navigation in Continuous Dictation Systems. *ACM Trans. Comput.-Hum. Interact. (TOCHI)* **2004**, *11*, 329–356. [[CrossRef](#)]
28. Hirsch, H.-G.; Pearce, D. The Aurora Experimental Framework for the Performance Evaluation of Speech Recognition Systems under Noisy Conditions. In Proceedings of the ASR2000-Automatic Speech Recognition: Challenges for the New Millenium ISCA Tutorial and Research Workshop (ITRW), Paris, France, 18–20 September 2000.
29. Monty, R.A.; Senders, J.W. *Eye Movements and Psychological Processes*; Routledge: London, UK, 2017; Volume 22.
30. Jacob, R.; Stellmach, S. What You Look at Is What You Get: Gaze-Based User Interfaces. *Interactions* **2016**, *23*, 62–65. [[CrossRef](#)]
31. Nukarinen, T.; Kangas, J.; Špakov, O.; Isokoski, P.; Akkil, D.; Rantala, J.; Raisamo, R. Evaluation of HeadTurn: An Interaction Technique Using the Gaze and Head Turns. In Proceedings of the Ninth Nordic Conference on Human–Computer Interaction, Gothenburg, Sweden, 23–27 October 2016; pp. 1–8.
32. Tang, Z.; Yan, C.; Ren, S.; Wan, H. HeadPager: Page Turning with Computer Vision Based Head Interaction. In Proceedings of the Asian Conference on Computer Vision, Taipei, Taiwan, 20–24 November 2016; Springer: Cham, Switzerland, 2016; pp. 249–257.
33. Gorodnichy, D.O.; Roth, G. Nouse ‘Use Your Nose as a Mouse’ Perceptual Vision Technology for Hands-Free Games and Interfaces. *Image Vis. Comput.* **2004**, *22*, 931–942. [[CrossRef](#)]
34. Varona, J.; Manresa-Yee, C.; Perales, F.J. Hands-Free Vision-Based Interface for Computer Accessibility. *J. Netw. Comput. Appl.* **2008**, *31*, 357–374. [[CrossRef](#)]
35. Crossan, A.; McGill, M.; Brewster, S.; Murray-Smith, R. Head Tilting for Interaction in Mobile Contexts. In Proceedings of the 11th International Conference on Human–Computer Interaction with Mobile Devices and Services, Bonn, Germany, 15–18 September 2009; pp. 1–10.
36. Esteves, A.; Verweij, D.; Suraiya, L.; Islam, R.; Lee, Y.; Oakley, I. SmoothMoves: Smooth Pursuits Head Movements for Augmented Reality. In Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology, Québec City, QC, Canada, 22–25 October 2017; pp. 167–178.
37. Jalaliniya, S.; Mardanbegi, D.; Pederson, T. MAGIC Pointing for Eyewear Computers. In Proceedings of the 2015 ACM International Symposium on Wearable Computers, Osaka, Japan, 7–11 September 2015; pp. 155–158.
38. Jalaliniya, S.; Mardanbeigi, D.; Pederson, T.; Hansen, D.W. Head and Eye Movement as Pointing Modalities for Eyewear Computers. In Proceedings of the 2014 11th International Conference on Wearable and Implantable Body Sensor Networks Workshops, NW Washington, DC, USA, 16–19 June 2014; pp. 50–53.
39. Crossan, A.; Williamson, J.; Brewster, S.; Murray-Smith, R. Wrist Rotation for Interaction in Mobile Contexts. In Proceedings of the tenth International Conference on Human Computer Interaction with Mobile Devices and Services, Amsterdam, The Netherlands, 2–5 September 2008; pp. 435–438.
40. Tregillus, S.; Al Zayer, M.; Folmer, E. Handsfree Omnidirectional VR Navigation Using Head Tilt. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, Denver, CO, USA, 6–11 May 2017; pp. 4063–4068.
41. Beckhaus, S.; Blom, K.J.; Haringer, M. ChairIO—the Chair-Based Interface. *Concepts Technol. Pervasive Games Read. Pervasive Gaming Res.* **2007**, *1*, 231–264.
42. Probst, K.; Lindlbauer, D.; Haller, M.; Schwartz, B.; Schrempf, A. A Chair as Ubiquitous Input Device: Exploring Semaphoric Chair Gestures for Focused and Peripheral Interaction. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Toronto, ON, Canada, 26 April–1 May 2014; pp. 4097–4106.
43. de Haan, G.; Griffith, E.J.; Post, F.H. Using the Wii Balance Board™ as a Low-Cost VR Interaction Device. In Proceedings of the 2008 ACM Symposium on Virtual Reality Software and Technology, Bordeaux, France, 27–29 October 2008; pp. 289–290.
44. Wang, J.; Lindeman, R.W. Silver Surfer: A System to Compare Isometric and Elastic Board Interfaces for Locomotion in VR. In Proceedings of the 2011 IEEE Symposium on 3D User Interfaces (3DUI), Singapore, 19–20 March 2011; pp. 121–122.
45. Ogata, M.; Sugiura, Y.; Makino, Y.; Inami, M.; Imai, M. SenSkin: Adapting Skin as a Soft Interface. In Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology, St. Andrews, Scotland, UK, 8–11 October 2013; pp. 539–544.