Effects of Immediate Feedback in Operating Information Device by Finger Tap Gesture

Kyosuke Watanabe, Makoto Oka(✉), and Hirohiko Mori(✉)

Tokyo City University, 1-28-1 Tamazutsumi, Setagaya-ku, Tokyo, Japan
{g2091402,moka,hmori}@tcu.ac.jp

Abstract. In this paper, we proposed a robust finger tap gesture recognition method and examined the feedback function when operating information devices using finger tap gestures. A finger tap gesture recognition method using hall sensors and gyro sensors achieved a high recognition rate without causing false detection. However, when operating information devices with finger tap gestures, there are the following two problems. The first is that it is recognized as a gesture different from the intended gesture, and the second is that the user's subjective gesture success or failure differs from the actual gesture success or failure. If the above phenomenon occurs, the user cannot operate the information device smoothly. In this paper, we aimed to make it possible for the user to quickly grasp the state of the information device and take appropriate actions by presenting the user with immediate feedback according to the type of gesture immediately after the successful gesture. As for the types of immediate feedback, tactile feedback and sound feedback were prepared and compared those feedback condition. As a result, it was found that by giving immediate feedback, the user can properly grasp the state of the information device and quickly per-form appropriate operations. Furthermore, when comparing tactile feedback and sound feedback, the operational performance tended to be higher for tactile feedback, and the subjective evaluation tended to be higher for sound feedback.

Keywords: Input interface · Eyes-free · Finger gesture

1 Introduction

In recent years, it has become possible to carry small information devices such as smartphones, smartwatches, and music players, and we are now in an era where we can benefit from computers anytime, anywhere. However, from the viewpoint of safety and social acceptability, it is difficult to operate information devices seamlessly in the following three situations:

- Situations where users are performing other tasks besides operating information device (For example, while walking or jogging).
- Situations where users can't move and speak something (For example, users are on a crowded train).

S. Yamamoto and H. Mori (Eds.): HCII 2022, LNCS 13305, pp. 166–182, 2022.
https://doi.org/10.1007/978-3-031-06424-1_13

- Situations where users are performing activities of daily living, including fine movements of fingers.

Gesture operation is one of the methods to operate information equipment without looking at it. In operating information devices seamlessly by gestures in the above three situations, it is necessary to fulfil all of the conditions <1> <2> <3> <4>, and, in addition, it is desirable to satisfy the condition <5>. However, none of the studies conducted satisfy all the following five conditions so far.

<1> There are no similar movements in daily life, and there is no need to switch between gesture recognition mode and gesture non-recognition mode by a specific operation.
<2> To make it possible for users to operate information device anytime regardless of the surrounding environment, the operation associated with the gesture should be compact.
<3> There is no need to watch the body part where the gesture is performed.
<4> Users can perform gestures while walking or jogging.
<5> To reduce the time and the effort required to operate information equipment smoothly, it is not necessary to acquire training data and build a classifier for each user.

In this paper, we propose a robust finger tap gesture recognition method that fulfill all above five conditions and examine the feedback function when operating information devices using finger tap gestures.

2 Related Works

The following are the studies that assume the operation of information devices by gestures while performing physical activities at the same time, and studies that examine false detection of gestures when performing various behaviors in daily life. Norieda et al. [1] proposed a gesture of tapping the arm with the other hand. Murao et al. [2] proposed a method for recognizing seven types of gestures (chop, throw, punch, draw a clockwise circle, draw a counterclockwise circle, jump, kick). Yoon et al. [3] proposed a method for recognizing gestures such as finger touch and swipe using a thin device attached to the index finger. However, we have shown that false detection can happen by jogging or finger flexion. Kawahata et al. [4] and Kerber et al. [5] investigated combinations of arm-based gestures that are less likely to cause false detection in activities of daily living. However, they have not made any proposals on how to recognize gestures. Yamamoto et al. [6] proposed a foot gesture that can be performed while jogging.

The following studies proposed gestures using only finger movements. Kubo et al. [7] proposed a method for recognizing 20 types of gestures, such as the posture that the finger is in contact with each other, by implementing piezo elements on the back of the hand. Chan et al. [8] proposed a method for recognizing gestures, such as the posture that the fingers are contacted with each other, by placing a fisheye lens on the hand. Saponas et al. [9] proposed a method for recognizing the state in which a force is applied

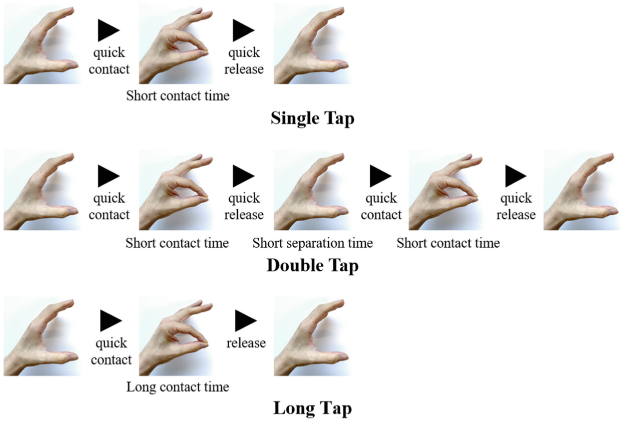

Fig. 1. Description of each gesture

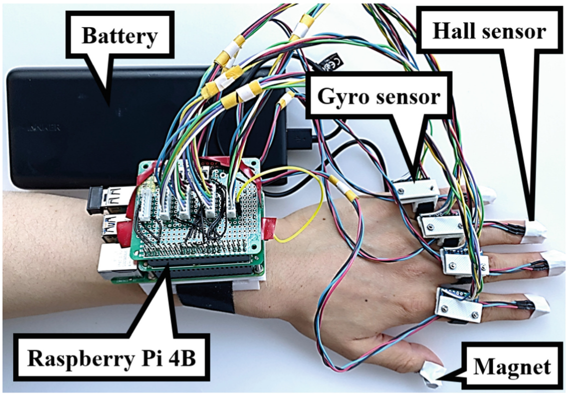

Fig. 2. Overall picture of gesture recognition device

Table 1. The state of satisfaction with conditions in related works

✓: meet the condition **–: unverified** **empty field: do not meet the condition**

Related works	<1> switching	<2> compact	<3> gaze	<4> walk/jog	<5> classifier
Norida et al. [1]	✓		✓	✓	
Murao et al. [2]	✓		✓	✓	
Yoon et al. [3]		✓	✓		✓
Kawahata et al. [4]	✓		✓	–	
Kerber et al. [5]	✓		✓	–	
Yamamoto et al. [6]	✓		✓	✓	
Kubo et al. [7]		✓		–	
Chan et al. [8]		✓	–	–	✓
Saponas et al. [9]		✓	–	–	
Zhang et al. [10]		✓	–	–	
Zhang et al. [11]		✓	–	–	
Dementyev et al. [12]		✓	–	–	
Nakamura et al. [13]		✓	✓	–	

to a finger by measuring the electromyogram of the forearm. Zhang et al. [10] proposed a method for recognizing gestures such as the posture in which fingers are in contact with each other and American Sign Language numbers by placing a transducer on the thumb and a microphone on the thumb and wrist. Zhang et al. [11] proposed a method in which a microphone and an inertial sensor are attached to the thumb to recognize the action of rubbing another finger with the thumb. Dementyev et al. [12] proposed a method of recognizing the posture in which the fingers are in contact with each other and the posture in which the palm is kept strongly spread by wrapping a band with a large number of pressure sensors around the wrist. Nakamura et al. [13] proposed a method of recognizing the gesture of tapping each segment of a finger other than the thumb with another finger by attaching an accelerometer to the base of each finger. Table 1 summarizes the condition satisfaction status from <1> to <5> mentioned in Sect. 1 regarding the studies mentioned in this chapter. From Table 1, there is no study that proposes gestures that satisfy all the conditions from <1> to <5> mentioned in Sect. 1 and can be used in all three situations mentioned in Sect. 1.

3 Finger Tap Gesture

3.1 The Types of Finger Tap Gesture

There are three types of finger tap gestures: single tap, double tap, and long tap (Fig. 1). Single tap is a gesture that performs the following three actions consecutively only once as the following processes:

1. Quickly bring the end segment of thumb and the end segment of another finger into contact with each other,
2. The time from contact to release is short,
3. Quickly release the fingers.

Double tap is a gesture of performing single tap twice consecutively.

Long tap is a gesture that performs the following three actions consecutively only once as following processes:

1. Quickly bring the end segment of thumb and the end segment of another finger into contact with each other.
2. Keep a long time from contact to release.
3. Release the fingers.

3.2 Sensors Used for Finger Tap Gesture Recognition

Linear output hall sensors (A1324LUA-T, hereafter hall sensor), a neodymium magnet (magnetic flux density 4200 mT, hereafter magnet), and gyro sensors (MPU9250) are used to recognize gestures. The hall sensor used in this paper has an output voltage of 2.5 V when the magnet is not nearby, and the output voltage increases as the magnet approaches (maximum 5 V). Measurement of sensor data and gesture recognition processing are performed by Raspberry Pi 4B.

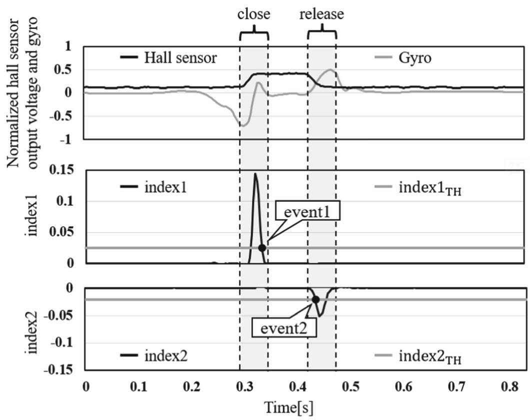

Fig. 3. Relationship between hall sensor output voltage, gyro, index1 and index2

The hall sensor is attached on the fingernail except the thumb, and the magnet is placed on the fingernail of the thumb. The gyro sensors are attached at base of the finger except the thumb (Fig. 2). The hall sensors measure the closeness of the thumb and other fingers, and the gyro sensors measure the speed of the finger. By attaching each sensor and magnet to the back side of the hand as shown in Fig. 2, it is not necessary to cover the palm side of the end segment, which is the most important when handling an object with the fingers.

3.3 Recognition Method of Finger Tap Gesture

We define index1 (Eq. (3)) and index2 (Eq. (4)) to recognize finger tap gestures. The value of index1 changes greatly when the end segment of thumb and the end segment of other fingers are quickly brought into contact with each other. The value of index 2 changes greatly when fingers are quickly separated from each other.

$$\mathrm{Ne}(x) = \begin{cases} 1 & \text{if } x < 0 \\ 0 & \text{if } x \geq 0 \end{cases} \tag{1}$$

$$\mathrm{Po}(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases} \tag{2}$$

$$\mathrm{index1}_t = \max\left(H_{MA,t} - H_{MA,t-5}, 0\right)\max\left(\Omega_{MA,t} - \Omega_{MA,t-5}, 0\right)\mathrm{Ne}\left(\Omega_{MA,t-5}\right) \tag{3}$$

$$\mathrm{index2}_t = \max\left(H_{MA,t} - H_{MA,t-5}, 0\right)\max\left(\Omega_{MA,t} - \Omega_{MA,t-5}, 0\right)\mathrm{Po}\left(\Omega_{MA,t}\right) \tag{4}$$

Equation (1) is a function that becomes 1 when the argument value is negative and 0 when the argument value is non-negative. Equation (2) is 1 when the argument value is positive and 0 when the argument value is non-positive.

In Eqs. (3) and (4), H_{MA} is the moving average among 5 times of the normalized hall sensor output voltage. We normalize the hall sensor value [2.4 V, 3.0 V] to [0, 1] because the hall sensor output voltage value in being tapped by a man with a thick finger was about 2.6 [V], while the one in being tapped by a woman with a thin finger was about 2.9 [V]. Ω_{MA} is a moving average among 5 times of the normalized gyro sensor values. We normalized the gyro sensor values [−2000 [°/s], 2000 [°/s]] to [−1, 1] because the range of the gyro sensor used in this paper was the minimum value −2000 [°/s] and the maximum value 2000 [°/s]. t represents the sampling time.

When the thumb and another finger quickly is bringing closer, the slope of both the hall sensor output voltage value and the gyro become positive (Fig. 3, upper row), and index1 takes a positive value with a large absolute value (Fig. 3, middle row). When the thumb and another finger quickly released, the slope of the hall sensor output voltage value becomes negative and the slope of the gyro becomes positive (Fig. 3, upper row), and index2 takes a negative value with a large absolute value (Fig. 3, middle row).

We explain how to recognize the three gestures of single tap, double tap, and long tap on each finger using index1 and index2. We define two events; first event (event1) occurs when the fingers are brought into contact with each other quickly, second event (event2) occurs when the fingers are quickly separated from each other (Fig. 3).

Event1 is an event that occurs when all following three conditions are fulfilled. $\mathrm{Index1_{TH}}$ represents the threshold value for index1 and $\mathrm{index2_{TH}}$ represents the threshold value for index2.

- $\mathrm{index1_{t-1}} \geq \mathrm{index1_{TH}}$
- $\mathrm{index1_{t}} < \mathrm{index1_{TH}}$
- The maximum value of index1 in the past 0.17 s of the event target finger is the largest of other fingers.

Event2 is an event that occurs when all of the following two conditions are satisfied.

- $\mathrm{index2_{t-1}} \geq \mathrm{index2_{TH}}$
- $\mathrm{index2_{t}} < \mathrm{index2_{TH}}$

Event1 is an event that occurs when index1 falls below $\mathrm{index1_{TH}}$, and event2 is an event that occurs when index2 falls below $\mathrm{index2_{TH}}$. The third condition of event1 was set to avoid gesture recognition on unintended fingers.

Then, single tap, double tap, and long tap are defined as follows using event1 and event2.

- Single tap recognition occurs when event2 occurs less than 0.3 s after event1 occurs and event1 does not occur within 0.3 s after event2 occurs.
- Double tap recognition occurs when event2 occurs less than 0.3 s after event1, event1 occurs less than 0.3 s after event2 occurs, and event2 occurs less than 0.3 s after event1 occurs.
- Long tap starts when the hall sensor output voltage exceeds 2.51 [V] for 0.5 consecutive seconds after event1 occurs, and long tap ends when the hall sensor output voltage becomes 2.51 [V] or less.

3.4 Effectiveness Verification Experiment of Finger Tap Gesture Recognition Method

Prior to the experiment, $index1_{TH}$ and $index2_{TH}$ were set as follows: $index1_{TH}$ was set to a value slightly higher than the maximum value of index1 when the page turning operation.

- Pinch one sheet of A4 size paper with fingers so as to be easy to pick up
- Separate the picked papers and the unpicked papers as quickly as possible

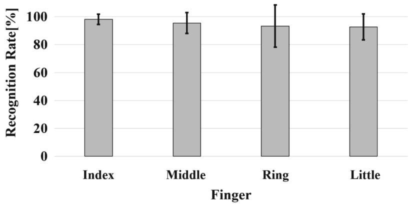

Fig. 4. Recognition rate for each finger

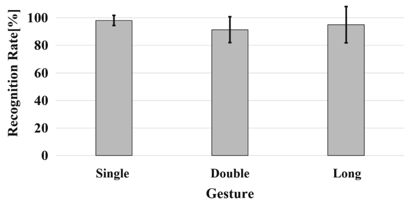

Fig. 5. Recognition rate for each gesture

Table 2. Recognition rate of gesture in related works

Related Works	Recognition Rate
Kubo et al.[7]	Entire: 84.4%
Chan et al.[8]	Entire: 84.75%
Saponas et al.[9]	Pinch state: 79% Travel mug gripping state: 85% Weight bag gripping state: 88%
Zhang et al.[10]	State of touching each phalanges: 93.77% American sign language: 95.64%
Zhang et al.[11]	Entire: 89%
Dementyav et al.[12]	Realtime recognition condition: 80.5%
Nakamura et al.[13]	Finger visible condition: 96.2% Finger invisible condition: 95.0%

We considered the page turning operation is the most similar to the finger tap gesture in activities of daily living. $Index2_{TH}$ was set to a value slightly higher than the minimum value of index2 when the finger was tapped quickly. In order to set each threshold, we collected data from two participants (participant 1: 25 years old male, participant 2: 54 years old female). The sampling frequency was set to 300 Hz. As a result of data collection, $index1_{TH}$ was set to 0.025 and $index2_{TH}$ was set to −0.008.

Two verifications were performed using the finger tap gesture recognition method described in Sect. 3.3. First, we verified whether false detections occurred during activities of daily life, and the second, we verified the recognition rate in performing gestures. The number of participants in the experiment was eight (6 male and 2 female).

As a result of the first verification, no false detections occurred even if the page turning motion and the pinch motion, which are activities of daily life similar to the finger tap gesture, were performed. Therefore, it can be said that it is a robust recognition method in daily life.

The explanation and results of the second verification are shown below. Assuming use while walking or jogging, we asked each participant to perform each gesture 30 times in each finger while walking on a spot without looking at the fingers. In addition, we also asked each participant to report whether gesture they subjectively feel failed or not. The average recognition rate for all fingers and gestures was 94.83% (SD: 10.01%). The average recognition rate for each finger is 98.06% (SD: 3.72%) for the index finger, 95.41% (SD: 7.44%) for the middle finger, 93.21% (SD: 15.08%) for the ring finger, and 92.63% (SD: 9.29%) for the little finger (Fig. 4). The average recognition rate for each gesture was 98.13% (SD: 3.78%) for single tap, 91.36% (SD: 9.54%) for double tap, and 95.00% (SD: 13.47%) for long tap (Fig. 5). Table 2 shows the recognition rate of finger gesture mentioned in Sect. 2.2. From Table 2, it can be said that the recognition rate of the finger tap gesture recognition method proposed by us is the highest level among finger gestures.

Though the recognition rate was high, two problems were shown. One is false recognition when user tried a double tap, a single tap was recognized by system (72 out of 957). The other is that there can be a difference between the subjective feelings of success or failure of a gesture and the actual success or failure of a gesture. For example, the user feels that the gesture is successful, but the system does not recognize it as the successful gesture. Of the total of 2866 trials of all experimental participants, the number of miss gestures judged by system was 149 times, and the number of miss gestures judged by the participants subjective feelings was 19 times. Of the 19 miss gestures that the participants judged subjectively, the number of miss gestures that the system judged was 7.

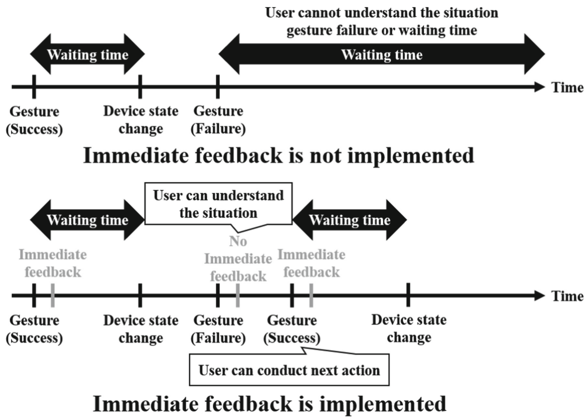

Fig. 6. The effect of immediate feedback

4 Giving Immediate Feedback

4.1 Problems and Their Improvement Plan

Fig. 7. State when the servomotor is installed

As we showed above, the gesture may be recognized different from the intended gesture, and that there may be a difference between the subjective feelings of success or failure of the gesture and the actual success or failure of the gesture. This problem becomes more serious when there is a time lag from the successful gesture to the change in the state of the information device. This time lag must occur the case such as amount of calculation of the information device are required and the Internet connection is unstable. When there is a time lag, the user cannot judge the success of the gesture and cannot accurately grasp the status of the information device (Fig. 6, top). Therefore, by giving feedback that represents only what type of gesture (hereinafter, immediate feedback) the system recognizes, the user must instantly and accurately grasp the state of the information device and take appropriate actions according to the state of information device (Fig. 6, bottom).

4.2 The Types of Immediate Feedback

We prepared two types of immediate feedback, sound feedback and tactile feedback.

Sound feedback uses earphones to present the user with the sound according to the types of gesture. A click sound is presented once when a single tap is successful, a click sound is presented twice when a double tap is successful, and a buzzer sound is presented when a long tap is successful.

Tactile feedback uses a servomotor (MG996R) to present the upper side of the user's wrist with tactile sensation according to the types of gesture (Fig. 7). A tactile sensation is presented once when a single tap is successful, a tactile sensation is presented twice when a double tap is successful, and the tactile presentation is continued for a certain period of times when a long tap is successful.

5 Experiment to Evaluate the Effect of Immediate Feedback

5.1 Procedure

Figure 8 shows the experimental procedure.

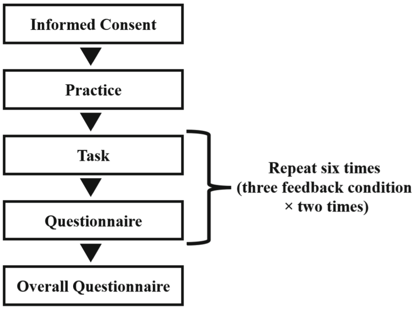

Fig. 8. Experimental procedure

Table 3. Relationship between gestures and functions

Finger	Gesture	Function
Index	Single Tap	Play / Stop
	Double Tap	15 second fast forward
	Long Tap	Next song
Middle	Double Tap	15 second rewind
	Long Tap	Previous song
Ring	Single Tap	Volume up
Little	Single Tap	Volume down

The phase of informed consent, we explained each participant the contents of the experiment based on the ethical guidelines for ergonomics research for humans [14] and obtained their consent.

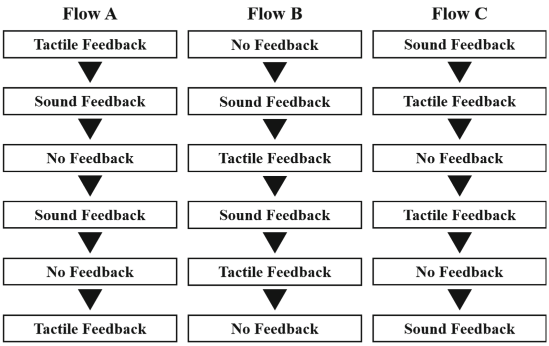

Fig. 9. Task flow

In the phase of training, we attached the experimental device to each participant and asked him/her to operate the virtual music player with gestures to get the knack of the gestures and to learn how to operate the music player (Table 3).

In the phase of task, assuming the operation of a music player, we asked each participant to perform the tasks under three feedback conditions (tactile feedback/sound feedback/no immediate feedback) twice. In order to remove the order effect, we prepared 3 order patterns as shown in Fig. 9 and asked each participant to perform one of the 3 patterns. A total of 26 operations (play and stop were performed 4 times each, and other operations were performed 3 times at random) were conducted for each task. In addition, we instructed each participant not to gaze the finger in performing the gesture and there were no visual feedback of the music player status.

In the phase of questionnaire, after each task, we asked each participant to answer the question of "Do you want to use it in your daily life?" in 5 stages.

In the phase of overall questionnaire, we asked each participant to answer the questions of "The condition that is the easiest to use", "The condition that is the most difficult to use" and "Which is better, the condition with immediate feedback or the condition without immediate feedback?".

16 participants were involved in the experiment (10 males and 6 females).

5.2 Experiment System

Figure 10 shows the overall picture of the experimental system. We used two RaspberryPi 4B, one for virtual music player (device1) and the other for gesture recognition and immediate feedback presentation (device2). The user performs a gesture, and device2 performs the gesture recognition process by the method described in Sect. 3.3. When gesture is recognized as a success, device2 sends a function command corresponding to the gesture to device1 by socket communication and presents immediate feedback to the user. When device1 receives an operation command from device2, it performs the function corresponding to the gesture after time lag (minimum 0.1 s, maximum 2 s, generated in uniform random). The lag time was implemented to mimic the situation where the two problems of the finger tap gesture may become serious mentioned in Sect. 3.4.

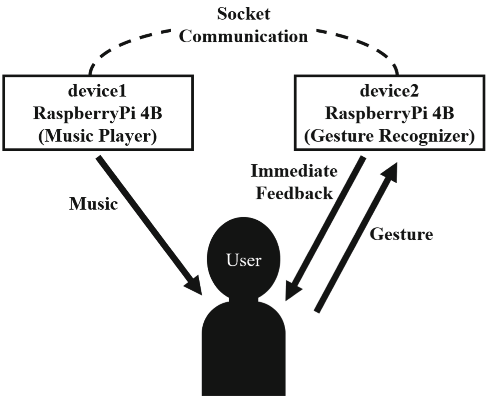

Fig. 10. Experimental system

5.3 Various Settings

The sampling frequency in device2 was set to 100 Hz. While the sampling frequency is different from that in the experiment described in Sect. 3.4, the values of $index1_{TH}$ and $index2_{TH}$ are the same as in Sect. 3.4 ($index1_{TH} = 0.025$, $index2_{TH} = -0.008$). No false detections occurred even when the high-speed page turning operation was performed, and each gesture was performed several times, but almost no false recognition occurred.

6 Results

6.1 Time Required Per Operation

The time for each operation was calculated to objectively evaluate whether each participant can operate the information device smoothly by finger tap gesture. It can be said that the shorter the time required for each operation is, the smoother the operation of the information device is. We standardized the value of the time in each operation of each participant and each gesture, because the time required for each operation differs depending on each participant and the time required for the next operation differs depending on the type of gesture. Hereinafter, the standardized value is defined as "time required per operation". In calculating the time required for each operation, we removed the waiting time from the time when device1 receives the function command to the time when the function corresponding to the gesture is performed. In addition, by comparing the video recorded in the experiment with the operation log measured during the experiment, we classified into two groups: one is the operation which is performed after the user judged that the operation is incorrect (hereinafter, after error), the other is ne is the operation which is performed after the user judged that the operation is correct (hereinafter, after non-error).

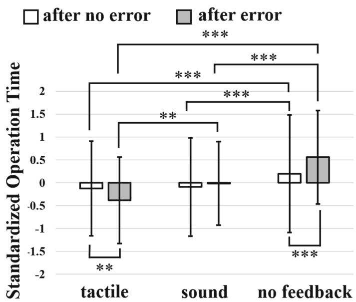

Fig. 11. The time required for each operation

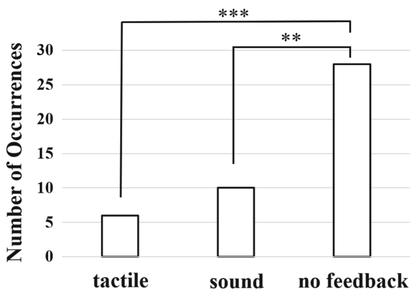

Fig. 12. The number of misunderstandings as success

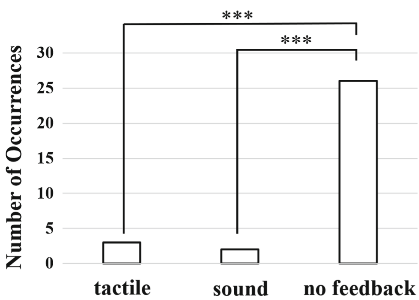

Fig. 13. The number of misunderstandings as failure

A between-subject two-way ANOVA was performed for the time required for each operation. The factors were feedback condition factors (3 levels: tactile feedback condition/sound feedback condition/no immediate feedback condition) and pre-operation factors (2 levels: after non-error/after error). The significance level was set to 5%.

Figure 11 shows the results of the ANOVA. As the result of ANOVA, an interaction between the two factors was observed ($p < 0.001$), a simple main effect test was performed. As the result of a simple main effect test for each level of the feedback condition factor, under tactile feedback condition, the time required for each operation is significantly shorter after error than after non-error ($p < 0.01$) and under no immediate feedback conditions, the time required per operation was significantly longer after error than after non-error ($p < 0.001$). As the result of a simple main effect test for each level of pre-operation factor, a simple main effect was observed both after non-error ($p < 0.001$) and after error ($p < 0.001$), so we conducted multiple comparisons by the Holm method. As a result of multiple comparisons after non-error, the time required for each operation of tactile feedback condition ($p < 0.001$) and sound feedback condition ($p < 0.001$) were significantly shorter than the condition no immediate feedback, respectively. As a result of multiple comparisons after the error, the time required for each operation of tactile feedback condition ($p < 0.001$) and sound feedback condition ($p < 0.001$) were significantly shorter than no immediate feedback condition. Furthermore, the time required for each operation of tactile feedback condition was significantly shorter than sound feedback condition ($p < 0.01$).

From the above results, it can be said that user can operate information device smoothly by giving immediate feedback immediately after a successful gesture in a situation where there is a waiting time. Furthermore, it can be said that by giving tactile feedback, user can deal with the operation error more quickly.

6.2 The Number of Times the User Could not Accurately Grasp the Status of the Information Device

We verified whether immediate feedback can suppress the errors that the user cannot accurately grasp the state of the information device.

We divided into the two types of the errors referring to the recorded video of experiment and the operation log measured during the experiment. One is the number of times that the users judged the operation was completed and moved on to the next operation despite correct operation was not completed (hereinafter, misunderstanding as success). The other is the number of times that the users judged the operation was not completed and performed the same operation again despite the correct operation was completed (hereinafter, misunderstanding as failure).

A chi-square test was performed to investigate whether there was a difference between each feedback condition regarding the number of misunderstandings as success (Fig. 12). As the result of a chi-square test and multiple comparisons by the Holm method, misunderstandings as success of tactile feedback condition ($p < 0.001$) and sound feedback condition ($p < 0.01$) were significantly less than no immediate feedback condition.

A chi-square test was performed to investigate whether there was a difference between each feedback condition regarding the number of misunderstandings as failure (Fig. 13). As the result of a chi-square test and multiple comparisons by the Holm method, misunderstandings as failure of tactile feedback condition ($p < 0.001$) and sound feedback condition ($p < 0.001$) were significantly less than no immediate feedback condition.

From the above results, it can be said that giving immediate feedback has the effect of suppressing the performance that the user cannot accurately grasp the state of the information device.

6.3 Results of Subjective Evaluations

In the experiment, each participant performed each feedback condition twice and the participant were asked the questionnaire "Do you want to use it in your daily life?" for each trial. The average value of the first and second answers for each feedback condition was used for the analysis. We performed a within-subject one-way ANOVA to investigate whether there is a difference for each feedback condition in the results of the questionnaire. The factor was feedback condition factor (3 levels: tactile feedback condition/sound feedback condition/no immediate feedback condition).

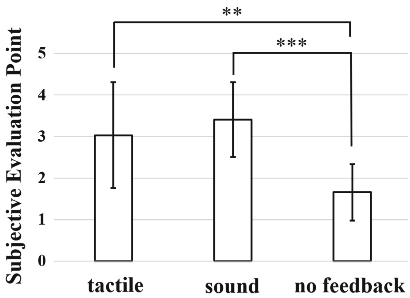

Fig. 14. Result of questionnaire item "Do you want to use it in your daily life?"

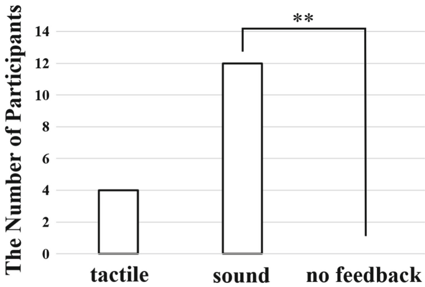

Fig. 15. Result of questionnaire item "The condition that is the easiest to use"

Figure 14 shows the results of the ANOVA. Since the main effect was observed ($p < 0.001$), multiple comparisons were performed by the Holm method. The point of tactile feedback condition ($p < 0.01$) and sound feedback condition ($p < 0.001$) were significantly higher than the immediate feedback condition. However, no significant difference was found between tactile feedback condition and the sound feedback condition (Fig. 15).

We conducted a chi-square test to investigate whether there are any differences between the feedback conditions regarding the response results of the "which is the easiest to use?" that were answered at the end of the experimental procedure. As a result of a chi-square test and multiple comparisons by the Holm method, the number of participants that prefer sound feedback conditions were significantly more than the no immediate feedback condition ($p < 0.01$), but no significant difference was observed between the other conditions.

At the end of the experiment procedure, we asked the participants "which is the most difficult to use?". In this question, all the participants answered that "No immediate feedback condition is the most difficult to use". Furthermore, for the question "Which is better, the condition with immediate feedback or the condition without immediate feedback?", The responses of 15 out of 16 participants were that "The condition with immediate feedback is better" and the remaining 1 responded that "Either is fine".

From the above results, it was found that the subjective comfort is improved by giving immediate feedback. In addition, many experimental participants answered that sound feedback was preferable to tactile feedback, but it can be said that there is no big difference.

7 Discussion

7.1 Discussion of Operational Performance

We found that giving the immediate feedback allows the users to reduce the operation time. It is considered that the reason is the user can predict the state of the change before the state change of the information device actually occurs and can prepare for the next operation in advance.

In addition, only the tactile feedback condition was found to shorten the time for the operation after the error. In the experiment, the time lag was set from 0.1 s to 2.0 s, but in the actual situation, the time lag is often longer. In such situations, the tactile feedback condition is the most effective feedback method.

It was found that the immediate feedback suppresses the error that the user cannot accurately grasp the state of the information device. Therefore, even when it is difficult for the user to determine the state change of the information device, it is possible to support the user to accurately grasp the state of the information device by giving immediate feedback. Since the experiment is conducted being supposed the operation of a music player, it was easy for the user to judge the state change of the information device. On the other hand, it should be difficult for the user to judge the change of state in the menu selection operation supposing the situation where the smartwatch or the smartphone is operated without visual information. In such a situation, it can be said that giving immediate feedback makes it easier for the user to grasp the state of the information device and the users can operate comfortably.

7.2 Discussion on Subjective Evaluation

The tactile feedback condition shows the highest in the operation performance, while the sound feedback condition tended to be high in the subjective evaluation. We are considering that the device that gives tactile feedback causes the discrepancy between operational performance and subjective evaluation. The servomotor used in the experiment is relatively large and heavy (Fig. 7). This may have hindered the user's comfort. Therefore, tactile feedback can be excellent in both operation performance and subjective evaluation by reducing the size and weight of the device that presents tactile feedback and devising the mounted position.

8 Conclusion

In this paper, we proposed a robust finger tap gesture recognition method that fulfill all above five conditions and examined the feedback function in operating information devices using finger tap gestures. A finger tap gesture recognition method using hall sensors and gyro sensors achieved a high recognition rate without causing false detection. However, when operating information devices with finger tap gestures, there are the following two problems: one is that it is recognized as a gesture different from the intended gesture, and the other is that the user's subjective gesture success or failure differs from the actual gesture success or failure. To address the above problems, we have added a function to present user with immediate feedback according to the type of gesture immediately after the successful gesture. Tactile feedback and auditory feedback were prepared as the types of immediate feedback. In the experiment, supposing the operation of a music player, we com-pared the three conditions of tactile feedback condition, sound feedback condition and no immediate feedback condition. As the result, it was found that by giving the immediate feedback, the user can properly grasp the state of the information device and quickly perform appropriate operations. Furthermore, when comparing tactile feedback and auditory feedback, the operational performance becomes higher for tactile feedback, and the subjective evaluation becomes higher for sound feedback. By reducing the size and weight of the device that presents tactile feedback and devising the mounting position, it should be able to be an excellent immediate feedback method in terms of both operational performance and subjective evaluation. Since the occurrence of misrecognition is a common problem in operating information device by gesture, the results of this paper will be able to be applied not only to finger tap gesture but also to operating information device by other gestures.

References

1. Norieda, S., Mitsuhashi, H., Sato, M.: ArmKeypad: a new input interface by tapping on user's arm. Trans. Hum. Interface Soc. **13**(4), 315–322 (2011)
2. Murao, K., Terada, T.: A motion recognition method by constancy decision. IPSJ J. **52**(6), 1968–1979 (2011)
3. Yoon, S., Huo, K., Ramani, K.: Wearable textile input device with multimodal sensing for eyes-free mobile interaction during daily activities. Pervasive Mob. Comput. **33**, 17–31 (2016)
4. Kawahata, R., Shimada, A., Yamashita, T., Uchiyama, H., Taniguchi, R.: Design of a low-false-positive gesture for a wearable device. In: International Conference on Pattern Recognition Applications and Methods, pp. 581–588 (2016)
5. Kerber, F., Schardt, P., Löchtefeld, M.: WristRotate - a personalized motion gesture delimiter for wrist-worn devices. In: The 14th International Conference on Mobile and Ubiquitous Multimedia, pp. 218–222 (2015)
6. Yamamoto, T., Terada, T., Tsukamoto, M., Yoshihisa, T.: A FootStep input method for operating information devices while jogging. IPSJ J. **50**(12), 2881–2888 (2009)
7. Kubo, Y., Koguchi, Y., Shizuki, B., Takahashi, S., Hilliges, O.: AudioTouch: minimally invasive sensing of micro-gestures via active bio-acoustic sensing. In: Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services, no. 36, pp. 1–13 (2019)

8. Chan, L., Chen, Y., Hsieh, C., Liang, R., Chen, B.: CyclopsRing: enabling whole-hand and context-aware interactions through a fisheye ring. In: Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology, pp. 549–556 (2015)
9. Saponas, T., Tan, D., Morris, D., Balakrishnan, R., Turner, J., Landay, J.: Proceedings of the 22nd Annual ACM Symposium on User Interface Software and Technology, pp. 167–176 (2009)
10. Zhang, C., et al.: FingerPing: recognizing fine-grained hand poses using active acoustic on-body sensing. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, no. 437, pp. 1–10 (2018)
11. Zhang, C., et al.: FingOrbits: interaction with wearables using synchronized thumb movements. In: Proceedings of the 2017 ACM International Symposium on Wearable Computers, pp. 62–65 (2017)
12. Dementyev, A., Paradiso, J.: WristFlex: low-power gesture input with wrist-worn pressure sensors. In: Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology, pp. 161–166 (2014)
13. Nakamura, Y., Sakai, T., Yazaki, K.: PhKey: an input interface for wearable devices using phalanges as keys. IPSJ J. **62**(2), 701–712 (2020)
14. Japan Human Factors and Ergonomics Society: The ethical guidelines for ergonomics research (2009)