

## AnkleSens: Foot Posture Prediction Using Photo Reflective Sensors on Ankle

Kikui, Kosuke

Masai, Katsutoshi

Sasaki, Tomoya

Inami, Masahiko

他

<https://hdl.handle.net/2324/7330752>

---

出版情報 : IEEE Access. 10, pp.33111-33122, 2022. Institute of Electrical and Electronics Engineers (IEEE)

バージョン :

権利関係 : Creative Commons Attribution 4.0 International



Received January 31, 2022, accepted February 21, 2022, date of publication March 8, 2022, date of current version March 30, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3158158

# AnkleSens: Foot Posture Prediction Using Photo Reflective Sensors on Ankle

KOSUKE KIKUI<sup>1</sup>, KATSUTOSHI MASAI<sup>1,2</sup>, TOMOYA SASAKI<sup>3</sup>, MASAHIKO INAMI<sup>3</sup>,  
AND MAKI SUGIMOTO<sup>1</sup>

<sup>1</sup>Graduate School of Science and Technology, Keio University, Yokohama 223-8522, Japan

<sup>2</sup>NTT Communication Science Laboratories, Kanagawa 243-0198, Japan

<sup>3</sup>Department of Advanced Interdisciplinary Studies, The University of Tokyo, Tokyo 153-8904, Japan

Corresponding author: Maki Sugimoto (maki.sugimoto@keio.jp)

This work was supported by the Japan Science and Technology Agency (JST) The Exploratory Research for Advanced Technology (ERATO) under Grant JPMJER1701.

**ABSTRACT** Recognizing foot gestures can be useful for subtle inputs to appliances and machines in everyday life, but for a system to be useful, it must allow users to assume various postures and work in different spaces. Camera-based and pressure-based systems have limitations in these areas. In this paper, we introduce *AnkleSens*, a novel ankle-worn foot sensing device that estimates a variety of foot postures using photo reflective sensors. Since our device is not placed between the foot and the floor, it can predict foot posture, even if we keep the foot floating in the air. We developed a band prototype with 16 sensors that can be wrapped around the leg above the ankle. To evaluate the performance of the proposed method, we used eight foot postures and four foot states as preliminary classes. After assessing a test dataset with the preliminary classes, we integrated the eight foot postures into five. Finally, we classified the dataset with five postures in four foot states. For the resulting 20 classes, the average classification accuracy with our proposed method was 79.57% with user-dependent training. This study showed the potential of foot posture sensing as a new subtle input method in daily life.

**INDEX TERMS** Photo reflective sensor, wearable device, machine learning, foot posture prediction.

## I. INTRODUCTION

Foot posture and gesture recognition is an important field for understanding human behavior, and we can apply it to various human-computer interactions (HCI). Foot gestures can be unobtrusive, and they can be used for subtle, everyday interactions. Fukahori *et al.* define foot gestures suitable for media and cell phone operation, and propose a recognition method using pressure sensors [3]. Foot gestures have also been used to control wearable robotic arms because they are a hands-free input method [13]. Foot gestures have their merits, and if we can measure them with a method that is not restricted by situation or posture, we can enable a variety of inputs in daily life and expand the range of applications such as for people with upper limb amputations.

There are three main methods for recognizing foot postures and gestures. Pressure-sensitive devices, such as pressure-sensitive floors, are a traditional method for estimating foot states in human sensing. Using an RGB camera [1] or a depth camera [2] is another useful foot sensing method. A third

notable method is the use of a sensor device integrated into shoes or socks [3]–[5]. Floor-based sensing systems are very accurate, and camera-based methods are easy to set up, but these methods must be in a specific sensing area. Embedded shoes or socks can be used in different areas, but it is difficult to capture barefoot postures and gestures. In addition, foot sensors in the sole of the shoe or on the floor would not work if the user did not place their feet on the floor because there is no significant pressure distribution. These limitations can be problems in everyday scenarios where the user frequently changes his or her foot position and posture.

To overcome these problems, we propose an ankle-worn foot posture detection device with photo reflective sensors (Fig. 1). The device detects foot posture based on skin deformation at the ankle. Therefore, the ankle-worn approach can capture foot posture while the user keeps their feet in the air, even without shoes or socks, and without being limited to a specific sensing field. The main contributions of the ankle-worn sensing approach are as follows:

- We propose a novel foot posture-sensing device worn on the ankle and using embedded optical sensors.

The associate editor coordinating the review of this manuscript and approving it for publication was Lei Shu<sup>1</sup>.

- We validate the feasibility of the approach, which allows us to estimate foot posture even when the user is barefoot and does not place their foot on the ground without space constraint.
- We evaluate the classification accuracy with the actual implementation.

## II. RELATED WORK

### A. FOOT SENSING AND INTERACTION

One of the most common methods of foot sensing is a pressure-based approach. Fukahori *et al.* developed pressure sensors in socks for foot-based interactions. They installed eight pressure sensors with force-sensing resistance in a pair of socks [3]. Saunders *et al.* attached force-dependent resistors to the end of each insole. They synchronized these sensors with depth data from Kinect to detect foot gestures [4]. Shi *et al.* used a smart insole with three pressure sensors for virtual locomotion [5]. Ismail *et al.* analyzed fluctuations in gait and posture using force plates and an accelerometer [6]. Furthermore, optical fiber-based sensor technology can be used to monitor in-shoe pressures during the gait [7], [8]. These pressure-based foot interaction methods are easy to use, and they have high classification accuracy, but they cannot be used when users do not place their feet on the ground. External camera-based interaction methods also have been developed. For example, Yousaf *et al.* proposed a camera-based approach for a user interface in automobiles [1]. A depth camera has also been used to estimate foot gestures. Simeone *et al.* investigated the use of foot motion to support basic 3D interaction using Microsoft Kinect [2]. Müller *et al.* used an external IR camera to capture foot gestures [9]. Arami *et al.* evaluated a system with IR sensors and an inertial measurement unit using an IR-based motion capture system [10]. These systems are very accurate, but they require multiple IR cameras to capture motion, and the user must stay within the cameras' field of view. Inertial sensors can be used for gait analysis [11]; they are sensitive to foot movement, but they have limited ability to classify foot postures. These sensing methods do not allow use with floating or bare feet, and they do not recognize foot gestures that consider the state of the fingertips. Therefore, the situations in which these sensing methods can be used are limited.

Foot-based interaction can be observed in the HCI field and in interactions with robots. Velloso *et al.* analyzed foot-based interactions between users and systems [12]. Sasaki *et al.* developed a system to control a wearable robotic arm with a device to capture foot gestures [13]. We believe that more applications and interactions will become possible with our device.

### B. WRIST-WORN DEVICE FOR HAND POSE ESTIMATION

Hand pose estimation methods have been proposed using an IR camera, an RGB camera, a depth camera, etc. For example, one of the best-known commercially available devices is *Leap Motion* [16], which consists of two cameras and three IR

LEDs. These camera-based methods, however, are limited to the detection range.

To overcome this limitation, wrist-worn devices have been proposed. These devices attach electromyography sensors, accelerometers, and barometric pressure sensors to the user's wrist to estimate hand gestures and finger movements [14], [15]. These wrist-worn devices are not limited to the sensing area. They can also be attached to the user's ankle. That allows for the estimation of foot gestures and postures. Lv *et al.* proposed another method using electromyography and an accelerometer [14]. Their method allows the detection of finger gestures and the estimation of force based on a wrist-worn device. Shull *et al.* proposed another method for estimating hand posture based on a wrist-worn device that uses barometric pressure sensors to detect hand gestures and estimate finger angles [15]. In addition, a photo reflective sensor array was used to detect hand posture with a simple machine learning technique [17] or with a tomographic imaging technique [18]. Such wrist-worn hand posture detection devices inspire our ankle-worn method.

### C. WEARABLE PHOTO REFLECTIVE SENSING DEVICES

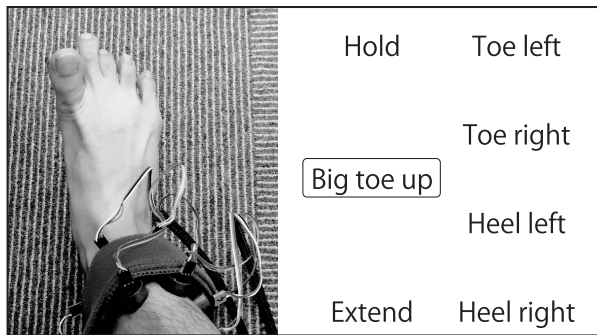
A photo reflective sensor is a tiny module that is widely used with HCI. It can be used not only for hand pose estimation (as mentioned in the previous section) but also for various human states, such as facial expression [19]. Yamashita *et al.* installed 20 photo reflective sensors on a head-mounted display (HMD) and detected the deformation of the cheek [20]. *EarTouch* enables users to interact with computing devices by the ear [21]. Kasahara *et al.* developed a wrist-worn device that recognizes four hand shapes and 14 wrist-bending states using photo reflective sensors [22]. Sugiura *et al.* proposed a device to recognize hand gestures by measuring skin deformation on the back of the hand [23]. *SenShoe* is a gait pattern recognition device that uses a photo reflective sensor attached to shoes [24]. FirstVR is a hand gesture input device that uses photo reflective sensors [25]. These methods mainly use such sensors to measure the change in skin deformation in response to the target human movement. We believe that it would be possible to measure the change in skin deformation corresponding to the movement of the ankle. We exploit this feature to develop a novel device for detecting foot posture.

## III. METHOD

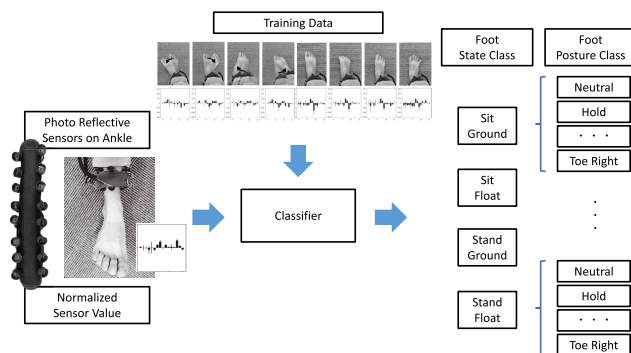
In this section, we describe the configuration of the device and the foot states and postures targeted in our experiment. *AnkleSens* detects various gestures by measuring the skin deformation of the ankle with photo reflective sensors. It can be worn barefoot, and detects gestures in the air. The system workflow of *AnkleSens* is shown in Fig. 2.

### A. PHOTO REFLECTIVE SENSOR

A photo reflective sensor consists of an infrared (IR) transmitter and a receiver (Fig. 3a). It detects a reflection and works as a proximity sensor. This sensor module is suitable



**FIGURE 1.** Concept of *AnkleSens*.



**FIGURE 2.** The system workflow of *AnkleSens*. Based on the training dataset, we train classifiers to identify the foot states and postures.

for wearable applications because it is tiny, easy to install, and has low processing costs.

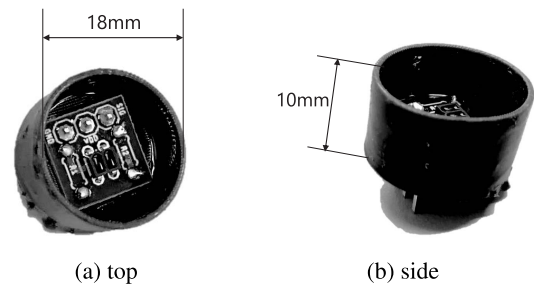
To estimate hand posture, measuring skin deformation associated with wrist tendons is a well-known approach. In this study, we use a similar approach to detect foot posture. When people move their toes or change their foot position, the movement deforms their ankles. Therefore, by measuring the skin deformation related to the tendons and muscle movements of the ankle with photo reflective sensors, we should be able to estimate foot posture.

We anticipate that this type of ankle-based sensing can be used to predict foot posture in various situations, even when the person is barefoot (e.g., daily indoor activities, gymnastics, ballet, kendo, judo), since the sensors do not need to be attached to the shoes, socks, or floor.

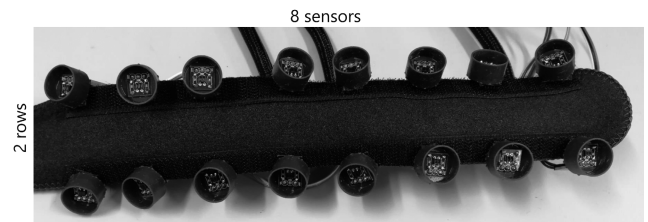
To place the sensor in the detection position and shut out the external light, we used a 3D printer to make sensor holders (Fig. 3b). Then, we installed the sensors on the holders and attached them to an ankle strap. We recorded a supervised learning dataset with preliminary classes and analyzed the classification accuracy by predicting the classes.

### B. DEVICE: *AnkleSens*

We installed photo reflective sensors in the sensor holders of a band and attached them to a user's ankle. The band is 45 mm

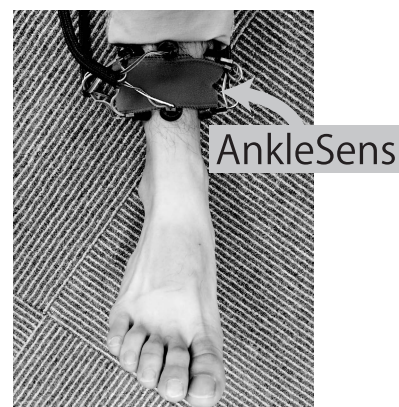


**FIGURE 3.** A view of a photo reflective sensor and holder.



**FIGURE 4.** *AnkleSens*, our novel device for detecting foot posture. We attached 16 sensors to the band in two rows of eight at nearly equal intervals. The band is 45 mm wide and 450 mm long. These sensors are connected to the multiplexer, and the Arduino Pro Mini controls this multiplexer and receives sensor values from the sensors.

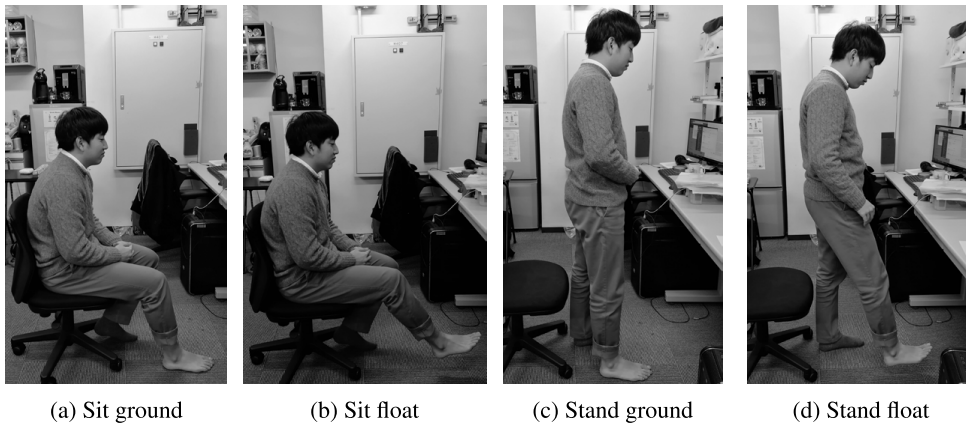
wide and 450 mm long. We placed 16 sensors in two rows of eight, each at nearly equal distances on the band (as shown in Fig. 4). The sensors were placed between the upper part of the ankle and the lower part of the shin. A snapshot of the device is shown in Fig. 5. These sensors were connected to a multiplexer. We used the Arduino Pro Mini to control the multiplexer and obtained values from the sensors.



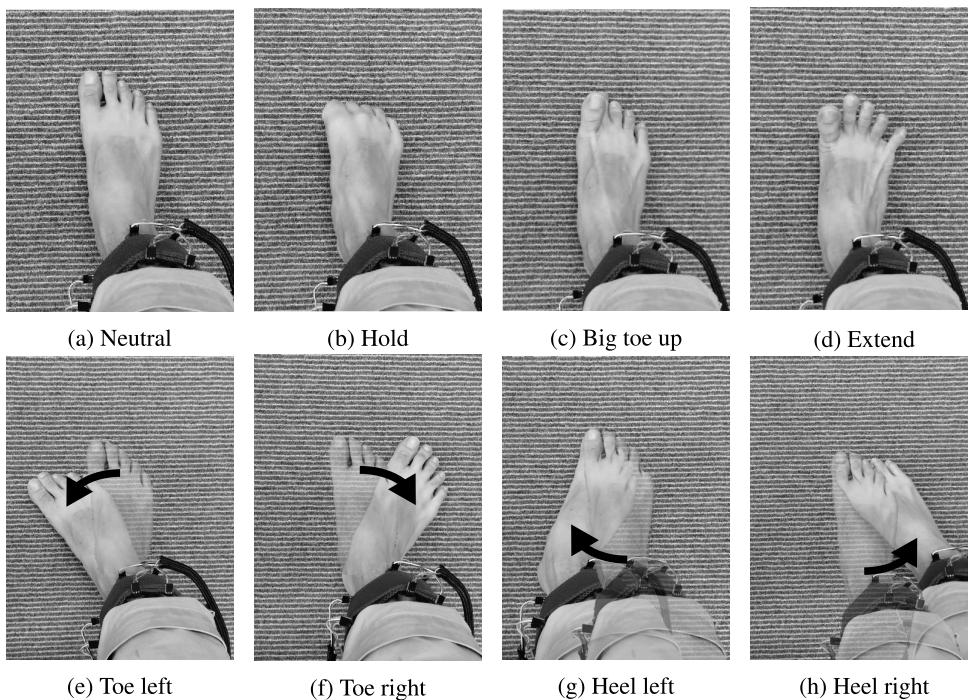
**FIGURE 5.** A snapshot of the device. We have placed *AnkleSens* between the upper part of the ankle and the shin.

We used Kodenshi SG-105 as a photo reflective sensors with a 3mm-6mm sensitivity to the skin. To keep the sensor within the sensitive distance from the skin, a custom sensor holder was fabricated with a 3D printer in a cylindrical shape (Fig. 3). The holder had a diameter of 18 mm and a height of 10 mm. This holder blocks the ambient light, which can be a disturbing factor for our purpose.





**FIGURE 6.** Four foot states.



**FIGURE 7.** Eight foot posture classes for data collection process.

We recorded a supervised learning dataset with preliminary classes and analyzed the classification accuracy by predicting the classes.

### C. FOOT STATES AND POSTURES

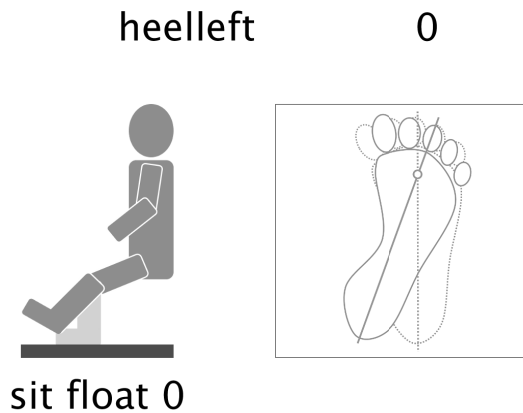
#### 1) FOOT STATES

We defined four foot states (Fig. 6) to cover the conditions in which the foot is standing on the ground or floating, and whether person is sitting or standing.

#### 2) FOOT POSTURES

Our device is designed to classify foot postures in which users keep their feet on the ground and lift their feet off the ground. To analyze the performance of the device, we defined

eight preliminary classes of foot postures as shown in Fig. 7. *neutral* is the initial posture from which users adopt other postures during data collection. In three postures, users change their toe movements, and in the others, they swing their feet. *extend*, *hold*, and *big toe up* are toe movements. *hold* requires users to flex their toes as if they were holding something. *extend* involves lifting the toes and creating gaps between them. *big toe up* requires users to lift only their big toe and lower the other toes. *Toe left*, *toe right*, *heel left* and *heel right* are foot-swinging postures derived from *toe rotation* and *foot rotation* defined by Scott *et al.* [26] and referred to by Eduardo *et al.* [12] in their survey. In *toe left* and *toe right*, the user pivots the foot around the heel. In *heel left* and *heel right*, the user pivots the foot around the toe. Users perform



**FIGURE 8.** An example of the instruction. Participants imitate the figure shown on the right side of the display. We ask them to maintain the target postures while the figure is shown on the display.

these postures by placing their feet on the ground and lifting them off the ground.

#### D. CLASSIFICATION

We used a support vector machine (SVM) to predict the target states and postures. We labeled four foot states and eight foot postures during data collection. We used Python and scikit-learn to calculate the classification accuracy using the 5-fold cross-validation method.

#### IV. DATA COLLECTION

This section describes the data collection protocol to validate the effectiveness of the device. We collected datasets with 13 subjects (12 males, 1 female) in 32 postures (7 postures and *neutral* in 4 states) 10 times each. The duration of each posture was recorded as 100 frames at 75 frames per second (FPS). This 100-frame data was used to classify the states and postures.

##### A. DATA ACQUISITION PROTOCOL

We collected seven types of postures using the following protocol for one trial, at 75 frames per second.

- 1) Put the foot in the neutral posture in the starting position (Fig. 7a).
- 2) Change the body position to the target posture (75 frames).
- 3) Maintain the target posture (100 frames).
- 4) Return to the neutral posture (50 frames).
- 5) Maintain the neutral posture (50 frames).

Before data collection, we provided instructions about data collection. First, we showed each target posture (as in Fig. 8) and asked them to imitate the posture. Then, we explained the data collection procedure. After these instructions, we attached *AnkleSens* to the participant and began the procedure. The target postures and states were indicated during

data collection by both illustrations and text on display on the desk.

During data collection, we randomized the order of all the postures and collected data in one state. After we obtained the data, the trial was repeated five times in each state. After obtaining a dataset with all four states, we repeated the data collection one more time. We collected 7 postures  $\times$  5 times  $\times$  4 states  $\times$  2 sessions = 280 trials. Each trial contained 100 frames of target posture data and 50 frames of *neutral* data. In total, we had 42,000 frames per participant available for analysis.

This study was conducted with signed consent forms from participants under an experimental protocol approved by the ethical committee of the faculty of science and technology, Keio University.

#### V. ANALYSIS

Based on the collected data, we tested whether gestures and postures could be classified using our device. Fig. 9 shows the distribution of sensor values of each posture in the four states. The Fig. 9 shows the difference between the average sensor value of *neutral* and each posture. In the following analysis, we trained an individual classifier using all the data from the seven postures and one-fifth of the data from *neutral*.

##### A. ANALYSIS ON FOUR FOOT STATE

First, we analyzed the classification accuracy of the four foot states shown in Fig. 6. Then, using these state labels, we calculated the classification accuracy of the state. We have 7,700 frames in each state, and we calculated the accuracy using the 5-fold cross-validation method for each participant. The parameters of the SVM are shown in Table 1.

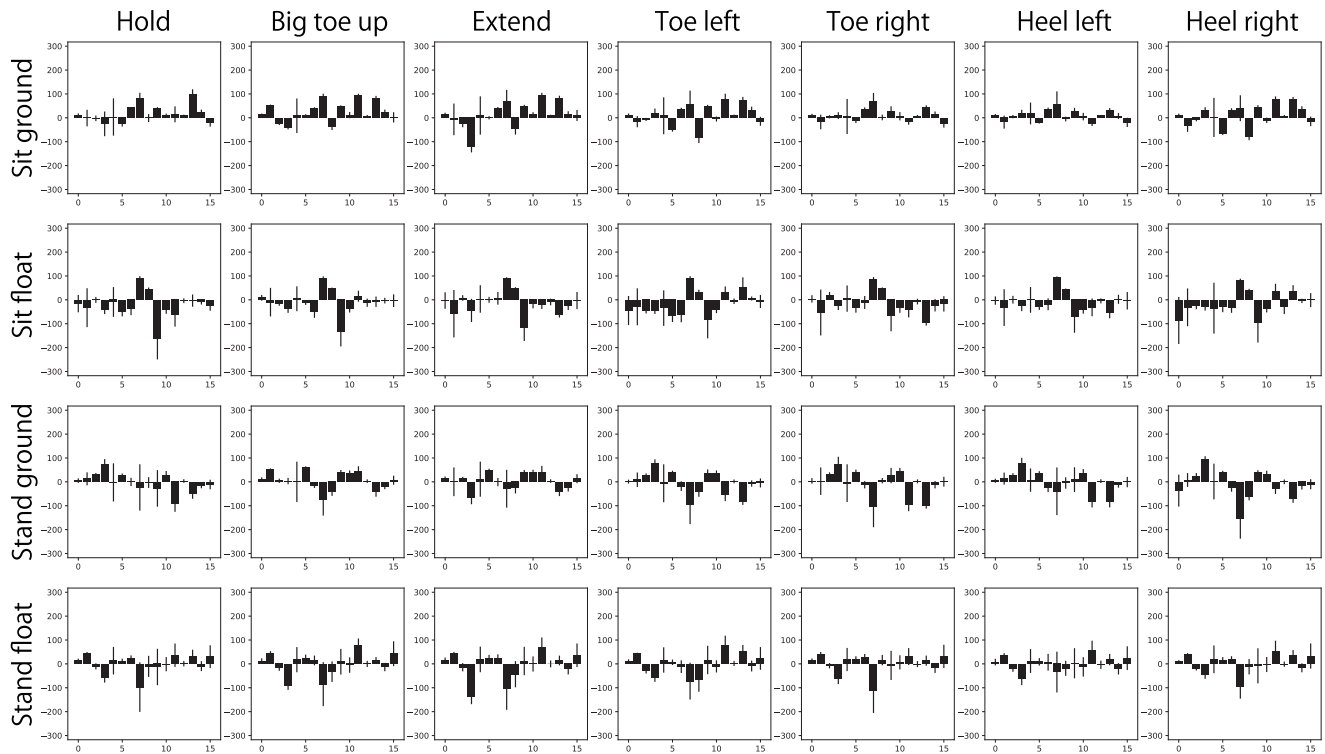
The average classification accuracy for the foot states was 94.21% for all participants. The highest rate was 99.02%, and the lowest was 88.53%. The foot state classification confusion matrix is shown in Fig. 10.

##### B. ANALYSIS OF PRELIMINARY EIGHT FOOT POSTURES IN EACH STATE

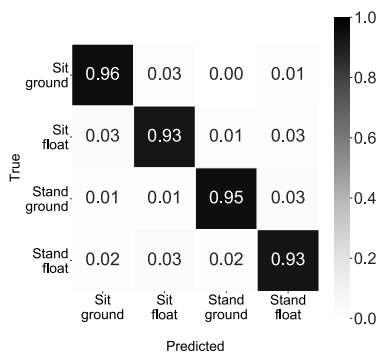
We analyzed the classification accuracy of posture in each state (Fig. 11). These classifiers were trained for each participant and each state. The parameters of SVM were the same as for state classification (Table 1), except that we changed  $C = 10$  in this evaluation. We also calculated the classification accuracy using the 5-fold cross-validation method (Table 2). The average classification accuracy was 73.28%. The classification accuracy of *sit float* was the highest among the four states. The classification accuracy was 80.20% in *sit float* state, and the lowest accuracy was 65.86% in *stand float* state.

##### C. ANALYSIS ON 32 CLASSES (PRELIMINARY EIGHT FOOT POSTURES IN FOUR STATES)

We had eight foot postures in the four foot states, yielding 32 foot classes for classifying both postures and states. Fig. 12 shows the confusion matrix of the 32 classes.



**FIGURE 9.** An example of the averaged sensor values in each foot state and posture class. The horizontal axis indicates the sensor IDs (0-15) and the vertical axis indicates relative sensor values to ones in the neutral posture. The error bar in each sensor indicates the standard deviation.



**FIGURE 10.** Confusion matrix of foot state classification.

We calculated the classification accuracy using the 5-fold cross-validation method. The average classification accuracy of all classes was 69.71%. As with the classification of individual states (section V-B), three pairs of postures were misclassified. However, fewer misclassifications were observed across states than misclassifications among postures in each state.

#### D. ANALYSIS WITH 20 CLASSES (INTEGRATED FIVE POSTURES BY FOUR STATES)

In the confusion matrix (Fig. 12) of the preliminary classes, several postures had misclassified each other. Class design is an essential parameter for machine learning. To assess the

**TABLE 1.** SVM parameters for state classification in scikit-learn SVC function.

Kernel	rbf
Gamma	$1.0 \times 10^{-6}$
C	100

performance of a novel user interface in a previous study, a follow-up analysis was performed that grouped existing classes [29]. In this study, we also examined the accuracy of an integrated posture class design. In the confusion matrix of the preliminary classes (Fig. 12), several postures misclassified each other. Fig. 13 shows the integrated foot posture classes. *Big toe up* and *extend* are similar because of joint coupling. When we float our foot, *toe left* and *heel right* and *toe right* and *heel left* are essentially the same postures, especially when we float our foot off the ground. Therefore, we analyzed the classification accuracy of the integrated classes of these foot postures. Fig. 13 shows the integrated foot posture classes. After these changes, *extend*, *toe left*, and *toe right* had twice the amount of data compared to the other classes. We used all data sets in this analysis. We used the 5-fold cross-validation method, just as in the analysis of the preliminary classes. Fig. 14 and Table 2 show the confusion matrix and the accuracy of the classification result for the five integrated postures in the four foot states. The average classification accuracy of the 20 classes with user-dependent training was 79.57%. After these changes, *extend*, *toe left*, and

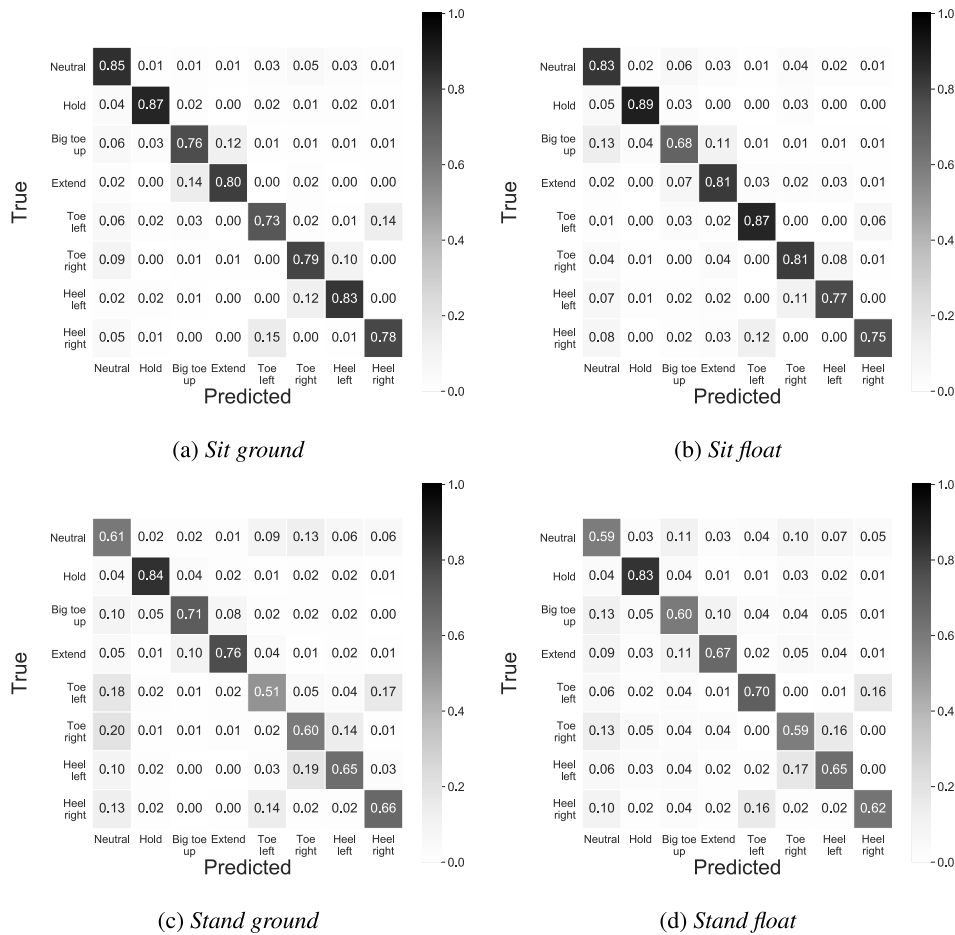


FIGURE 11. Average classification accuracy in the preliminary eight postures in the four states.

*toe right* had twice the amount of data compared to the other classes. We used all datasets in this analysis.

This change has improved the classification rate by about 10 points. It also reduced the number of cases of misclassification in certain postures. This indicates that the integrated class can be used to classify the foot postures.

In this analysis, we calculated the classification speed. It took about 3.5 seconds to train each subject's data. Since each frame can be classified in an average of 0.68 milliseconds, we can say that the system is sufficiently fast for some real-time occasions.

## VI. CASE TRIALS

### A. SEQUENTIAL STATES AND POSTURES

*AnkleSens* was designed to predict foot states and postures in daily situation. To this end, we collected time-series sensor data in sequential foot states and postures from one user. Fig. 15 shows the recorded situation and classification result, where the user changed his states and postures, starting at *sit ground* and following the order *sit float*, *stand ground*, *stand float*. In each state, five postures (*neutral*, *hold*, *extend*, *toe*

*left*, and *toe right*) were performed in order. The Fig. 15 shows the temporal transition of both postures and states.

### B. POSTURE ESTIMATION WITH VARIETY OF FOOTWEAR

As a further trial on our method, we also conducted a case trial with three types of footwear conditions (barefoot, indoor shoes, and sandals). Since our device is worn on the ankle, it is can to detect gestures while wearing shoes if the degree of freedom of the ankle and toes does not change significantly. Therefore, we investigated the possibility of classifying the foot gestures of a user who was wearing sandals and indoor shoes. The user performed five foot gestures in the following order: barefoot, wearing indoor shoes, and wearing sandals. The user repeated the procedure twice. After training the SVM classifier with the collected data, we visualized the results for the three types in Fig. 16.

## VII. DISCUSSION

For the posture classification, the two *sit* states had higher average classification rates than *stand* states. 90.34% in the *sit ground* state and 87.68% in the *sit float* state. In contrast, it was 79.62% in *stand ground* and 76.80% in *stand float*.



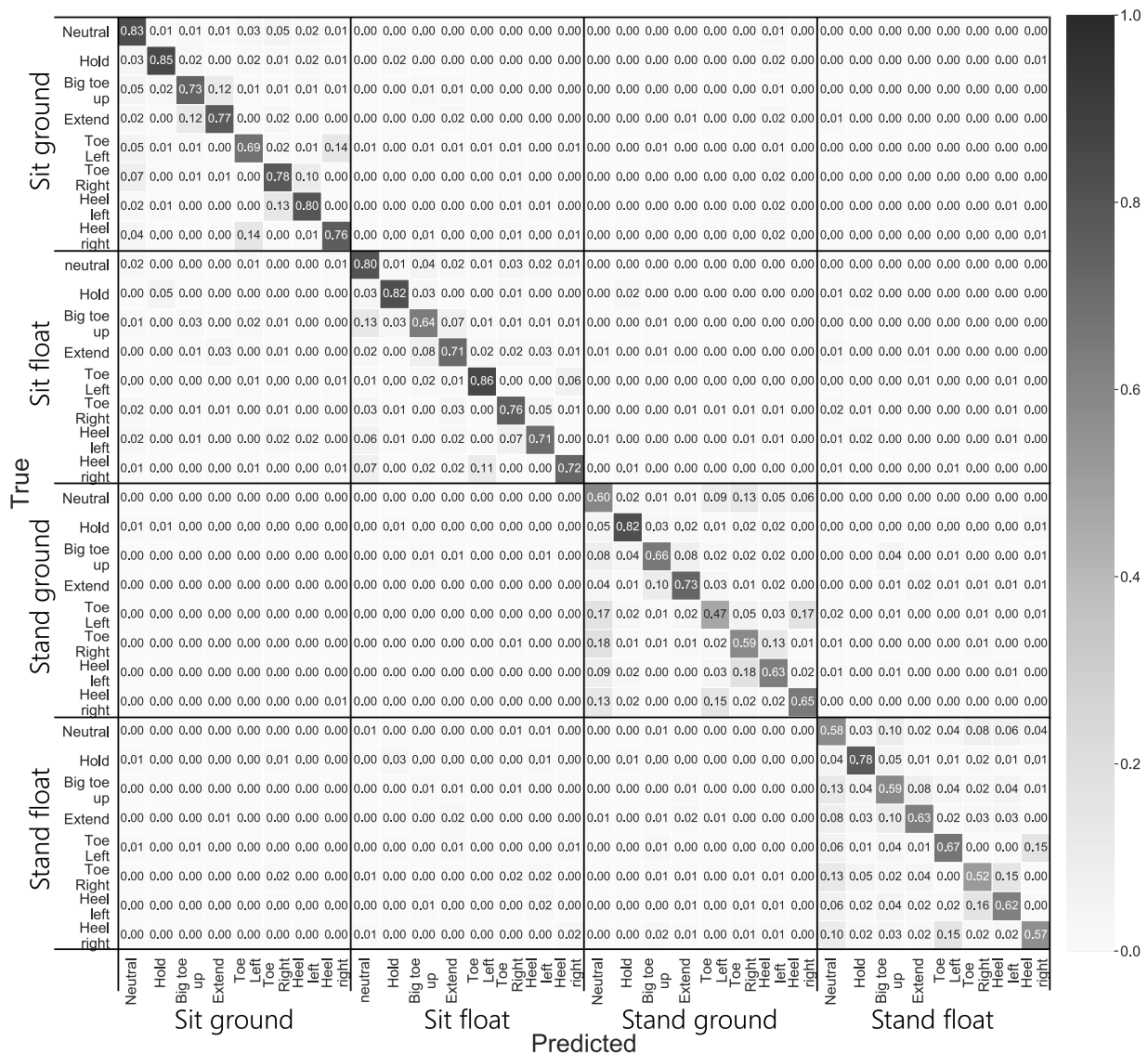


FIGURE 12. Confusion matrix of 32 classes (the preliminary eight postures by the four states).

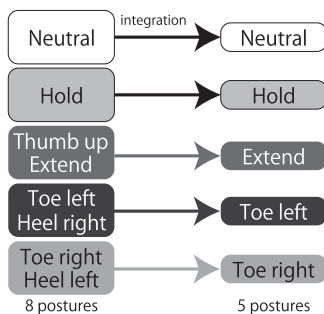


FIGURE 13. Integration of foot postures: We combined three pairs of similar postures. As a result, we have five foot postures with four foot states (20 classes in total).

In Sec. V, the results of the state classification were above 90.00% in all states. However, in the individual state and

TABLE 2. Average classification accuracy in the integrated eight/five postures in each state.

State	8 classes Avg.(%)	5 classes Avg.(%)
<i>Sit ground</i>	79.44	90.34
<i>Sit float</i>	80.20	87.68
<i>Stand ground</i>	67.10	79.62
<i>Stand float</i>	65.86	76.80
All states	73.2	83.61

posture classifications (Fig. 11), the accuracy rates were about 80.00% in the *sit* state and 70.00% in the *stand* state. In the classification of the postures, the two *sit* states had higher average classification rates than the *stand* states—90.34% in the *sit ground* state and 87.68% in the *sit float* state. In contrast, it was 79.62% in *stand ground* and 76.80% in *stand float*.

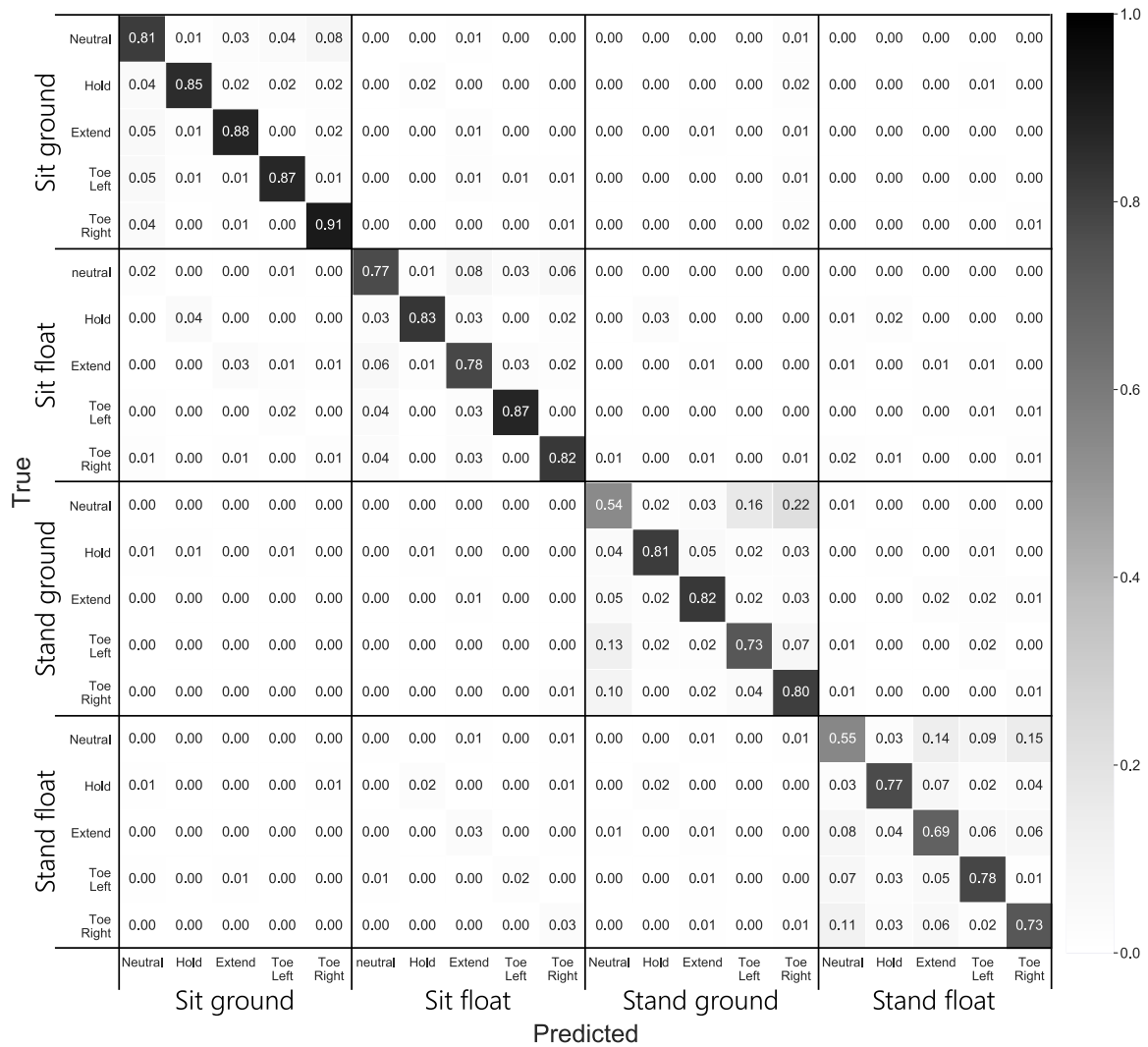


FIGURE 14. Confusion matrix of 20 classes (the integrated five postures by the four states).

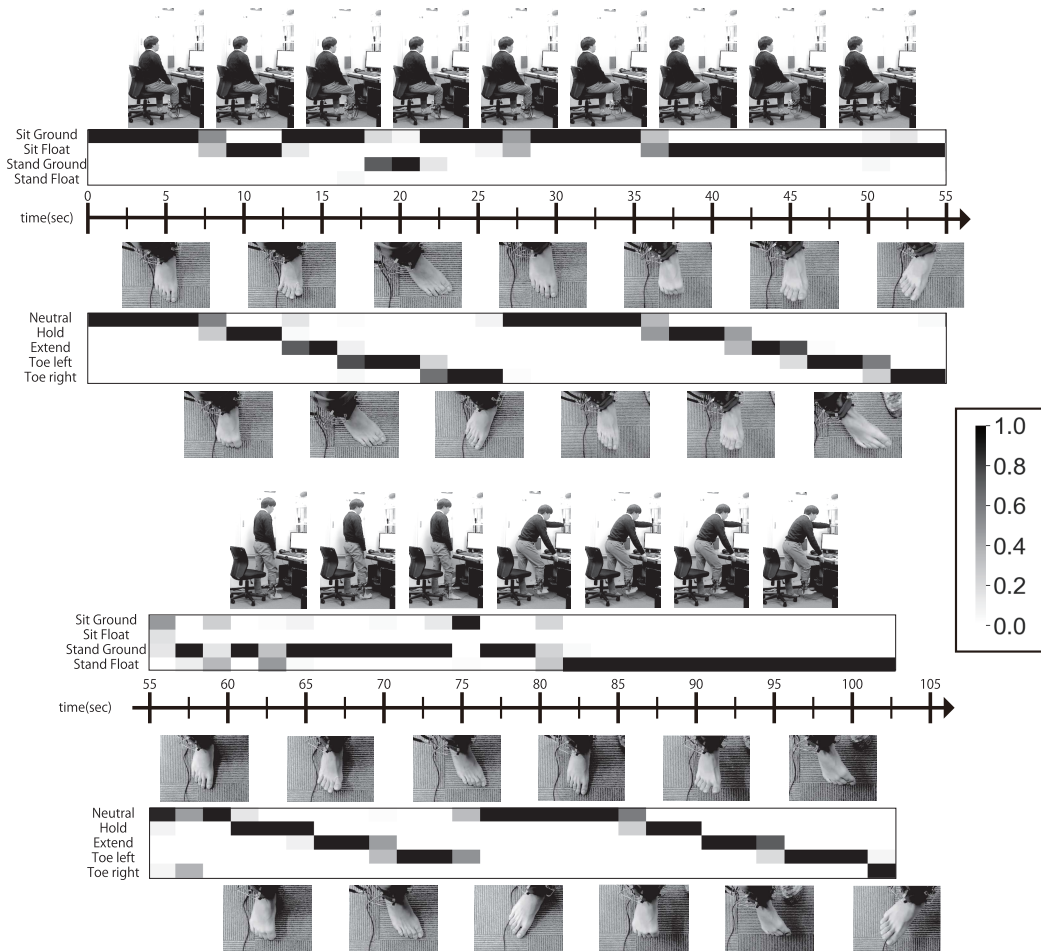
As described in the previous chapter, some pairs of postures were mixed with each other in the preliminary classes because of similarity. Therefore, we also analyzed the classification rates of the integrated classes. We treated *big toe up* as the same posture as *extend*. In the same way, we treated *heel right* as the same posture as *toe left* and *heel left* as the same posture as *toe right*. The average classification rate of the five integrated postures in each state was 83.61%. The average classification rate of 20 classes (5 postures in 4 states) was 79.57%.

We also performed one-user-leave-out cross-validation to check the generality of the device by learning between individuals using SVM. The accuracy of posture recognition was 35.32%, and that of state classification was 21.09%. This implies that individual learning is necessary for the practical use of the device, and we would like to explore more advanced algorithms in the future.

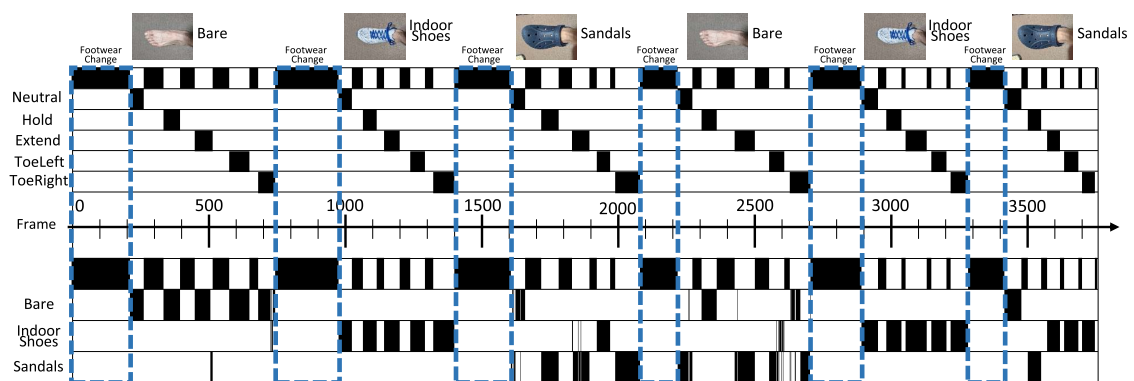
One reason for this result could be the reproducibility of the foot postures. In the *sit* states, the participants held

their foot on a stable chair. In the *stand* states, they had to maintain their body balance with their feet. Especially in the *stand float* states, they maintained their balance with a single foot. Therefore, this may be the reason that the *stand float* states had the lowest average accuracy. Moreover, in the *sit* states, they could view the instructions on display and control their posture with visual feedback. In the *stand* states, some participants looked at the instructions only on the display and kept their feet in the target posture without paying attention to their foot posture, which could be another cause for the results.

Users with large changes in sensor values also tended to have high identification rates. One possible reason for the large change in sensor values is the high flexibility of the ankle. On the other hand, one reason for the decrease in the accuracy of gesture classification is that the position of the sensor band shifts during repeated gestures. The sensor holder used in this study was not soft, and it had a small area that touched the ankle, so stability was an problem.



**FIGURE 15.** State and posture classification results of sequential data.



**FIGURE 16.** Footwear and posture classification trial with sequential footwear changes.

In addition, the range of sensor values may vary depending on the thickness of the ankle, but this can be normalized, and the difference between men and women is not likely to have a significant impact.

In Sec. VI-A, we tested our method with sequential postures and states. For most postures and states, a temporal transition can be seen. This suggests that our device can classify both the postures and the states correctly.

A few misclassifications in different states and postures were observed. For example, *toe right* in *stand ground* was classified as *neutral* in *sit ground* and *stand ground*. In this study, we used only the training dataset with five trials. If we increase the number of trials, we can obtain a more diverse training dataset to increase accuracy.

In Sec. VI-B, we show an experimental result of classifying posture and footwear states. Because of the drift of sensor values, the classification of the last condition of barefoot and sandals was not accurate, whereas the classification accuracy of posture states in each shoe class was high. Moreover, the classification of the foot posture class was almost perfect in this trial case.

## VIII. FUTURE WORK AND LIMITATIONS

Based on the experimental results, we would like to consider several ways to increase accuracy in future studies. For example, it would be effective to make the device lighter so that it can be stably fixed on the foot. This would optimize the sensitivity and placement of the sensor. We would also like to use machine learning or deep learning methods that take time series information into account.

Our foot-sensing device can be used barefoot. Some sports (gymnastics, judo, kendo, etc.) are performed barefoot. From a kinematic point of view, toes play some roles [27]. We expect that we will analyze toe movements with the help of *AnkleSens*. In the elderly, toe grasping strength is associated with light daily physical activity [28]. Our device may also be useful for training these people. As can be seen from the results in section VI-B, the drift of sensor values may be a problem if we use this approach for long-term trials. Therefore, we would consider using an online learning method for long-term use.

## IX. CONCLUSION

We proposed a novel foot posture sensing device called *AnkleSens*. This device can estimate users' foot posture even if they do not place their feet on the ground. In our analysis, the prototype device showed remarkable accuracy. Four states were classified at 94.21%, and 32 postures (4 states, 8 postures) were classified at 69.71%. We found that three pairs of postures were similar when users lifted their feet off the ground. Therefore, we combined the 32 classes into 20 (4 states, 5 postures). The average classification accuracy of the 20 classes with user-dependent training was 79.57%. In particular, our method showed significant classification accuracy of posture in the *sit ground* state (90.34%) and the *sit float* state (87.68%). This study introduced the possibility of a new approach to foot posture sensing that uses photo reflective sensors on a user's ankle, and it investigated the classification accuracy with the approach.

## REFERENCES

- [1] M. H. Yousaf and H. A. Habib, "Adjustable pedals by gesture recognition: A novel approach for user interface in automotive," *Middle-East J. Sci. Res.*, vol. 11, no. 11, pp. 1575–1583, 2012.
- [2] A. L. Simeone, E. Velloso, J. Alexander, and H. Gellersen, "Feet movement in desktop 3D interaction," in *Proc. IEEE Symp. 3D User Interface (3DUI)*, Mar. 2014, pp. 71–74.
- [3] K. Fukahori, D. Sakamoto, and T. Igarashi, "Exploring subtle foot plantar-based gestures with sock-placed pressure sensors," in *Proc. 33rd Annu. ACM Conf. Hum. Factors Comput. Syst.*, 2015, pp. 3019–3028.
- [4] W. Saunders and D. Vogel, "Tap-kick-click: Foot interaction for a standing desk," in *Proc. ACM Conf. Designing Interact. Syst.*, 2016, pp. 323–333.
- [5] X. Shi, J. Pan, Z. Hu, J. Lin, S. Guo, M. Liao, Y. Pan, and L. Liu, "Accurate and fast classification of foot gestures for virtual locomotion," in *Proc. IEEE Int. Symp. Mixed Augmented Reality (ISMAR)*, Oct. 2019, pp. 274–285.
- [6] H. Ismail, "Gait and postural sway analysis, A multi-modal system," in *Proc. ACM Int. Conf. Multimodal Interact. (ICMI)*, 2015, pp. 629–633.
- [7] A. G. Leal-Junior, A. Frizera, L. M. Avellar, C. Marques, and M. J. Pontes, "Polymer optical fiber for in-shoe monitoring of ground reaction forces during the gait," *IEEE Sensors J.*, vol. 18, no. 6, pp. 2362–2368, Mar. 2018.
- [8] M. F. Domingues, N. Alberto, C. S. J. Leitão, C. Tavares, E. R. de Lima, A. Radwan, V. Sucasas, J. Rodriguez, P. S. B. André, and P. F. C. Antunes, "Insole optical fiber sensor architecture for remote gait analysis—An e-health solution," *IEEE Internet Things J.*, vol. 6, no. 1, pp. 207–214, Feb. 2019.
- [9] F. Müller, J. McManus, S. Günther, M. Schmitz, M. Mühlhäuser, and M. Func, "Mind the tap: Assessing foot-taps for interacting with head-mounted displays," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, 2019, pp. 1–3.
- [10] A. Arami, N. S. Raymond, and K. Aminian, "An accurate wearable foot clearance estimation system: Toward a real-time measurement system," *IEEE Sensors J.*, vol. 17, no. 8, pp. 2542–2549, Apr. 2017.
- [11] L. Tran, T. Hoang, T. Nguyen, H. Kim, and D. Choi, "Multi-model long short-term memory network for gait recognition using window-based data segment," *IEEE Access*, vol. 9, pp. 23826–23839, 2021.
- [12] E. Velloso, D. Schmidt, J. Alexander, H. Gellersen, and A. Bulling, "The feet in human-computer interaction: A survey of foot-based interaction," *ACM Comput. Surv.*, vol. 48, no. 2, p. 21, 2015.
- [13] T. Sasaki, M. Y. Saraiji, C. L. Fernando, K. Minamizawa, and M. Inami, "MetaLimbs: Metamorphosis for multiple arms interaction using artificial limbs," in *Proc. ACM SIGGRAPH Posters*, 2017, 55.
- [14] B. Lv, X. Sheng, W. Guo, X. Zhu, and H. Ding, "Towards finger gestures and force recognition based on wrist electromyography and accelerometers," in *Intelligent Robotics and Applications (ICIRA)*. Cham, Switzerland: Springer, 2017, pp. 373–380.
- [15] B. Peter Shull, S. Jiang, Y. Zhu, and A. X. Zhu, "Hand gesture recognition and finger angle estimation via wrist-worn modified barometric pressure sensing," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 4, pp. 724–732, Apr. 2019.
- [16] *Leap Motion*. Accessed: Mar. 28, 2021. [Online]. Available: <https://www.leapmotion.com/>
- [17] R. Fukui, M. Watanabe, T. Gyota, M. Shimosaka, and T. Sato, "Hand shape classification with a wrist contour sensor: Development of a prototype device," in *Proc. 13th Int. Conf. Ubiquitous Comput. (UbiComp)*, vol. 2011, pp. 311–314.
- [18] J. McIntosh, A. Marzo, and M. Fraser, "Sensir: Detecting hand gestures with a wearable bracelet using infrared transmission and reflection," in *Proc. UIST*, 2017, pp. 593–597.
- [19] K. Masai, Y. Sugiura, M. Ogata, K. Kunze, M. Inami, and M. Sugimoto, "Facial expression recognition in daily life by embedded photo reflective sensors on smart eyewear," in *Proc. 21st Int. Conf. Intell. User Interface*, Mar. 2016, pp. 317–326.
- [20] K. Yamashita, T. Kikuchi, K. Masai, M. Sugimoto, H. Bruce Thomas, and Y. Sugiura, "CheekInput: Turning your cheek into an input surface by embedded optical sensors on a head-mounted display," in *Proc. 23rd ACM Symp. Virtual Reality Softw. Technol.*, 2017, pp. 1–8.
- [21] T. Kikuchi, Y. Sugiura, K. Masai, M. Sugimoto, and H. B. Thomas, "EarTouch: Turning the ear into an input surface," in *Proc. 19th Int. Conf. Hum.-Comput. Interact. Mobile Devices Services*, 2017.
- [22] H. Karasawa, R. Fukui, M. Watanabe, and S. Warisawa, "Simultaneous recognition of hand shape and two-axis wrist bending using wearable wrist contour measuring device," in *Proc. IEEE/ASME Int. Conf. Adv. Intell. Mechatronics (AIM)*, Jul. 2019, pp. 1550–1555.
- [23] Y. Sugiura, F. Nakamura, W. Tawai, T. Kikuchi, and M. Sugimoto, "Behind the palm: Hand gesture recognition through measuring skin deformation on back of hand by using optical sensors," in *Proc. 56th Annu. Conf. Soc. Instrum. Control Eng. Jpn. (SICE)*, Sep. 2017, pp. 1082–1087.



- [24] K. Inaba, A. Murai, and Y. Sugiura, "Center of pressure estimation and gait pattern recognition using shoes with photo-reflective sensors," in *Proc. 30th Austral. Conf. Comput.-Hum. Interact.*, 2018, pp. 224–228.
- [25] E. Tamaki, S. Hosono, and K. Iwasaki, "FirstVR: A muscle deformation sensors array device to detect finger gestures and noise reduction case," in *Proc. 2nd Int. Conf. Electron., Commun. Control Eng.*, 2019, pp. 21–24.
- [26] J. Scott, D. Dearman, K. Yatani, and K. N. Truong, "Sensing foot gestures from the pocket," in *Proc. 23rd Annu. ACM Symp. User Interface Softw. Technol.*, Oct. 2010, pp. 199–208.
- [27] K. N. S. Gontijo, C. T. Candotti, G. D. S. Feijó, L. P. A. Ribeiro, and J. F. Loss, "Kinematic evaluation of the classical ballet step 'Plié,'" *J. Dance Med. Sci.*, vol. 18, no. 2, pp. 70–76, 2015.
- [28] T. Abe, R. S. Thiebaud, J. P. Loenneke, and N. Mitsukawa, "Association between toe grasping strength and accelerometer-determined physical activity in middle-aged and older women," *J. Phys. Therapy Sci.*, vol. 27, no. 6, pp. 1893–1897, 2015.
- [29] C. Harrison, D. Tan, and D. Morris, "Skinput: Appropriating the body as an input surface," in *Proc. SIGCHI Conf. Hum. Factors Comput. Syst. (CHI)*, 2010, pp. 453–462.



**KOSUKE KIKUI** received the M.S. degree from Keio University, Japan, in 2020.

His research interests include wearable devices and machine learning.



**TOMOYA SASAKI** received the M.S. degree from Graduate School of Media Design, Keio University, Japan. He is currently pursuing the Ph.D. degree with The University of Tokyo, Japan.

His research interests include robotics, haptics, wearable technologies, and virtual reality.



**MASAHIKO INAMI** received the Ph.D. degree from The University of Tokyo, Japan, in 1999.

He is currently a Professor with The University of Tokyo. His research interests include human augmentation technologies with bioengineering, HCI, and robotics. He is also the Director of JST ERATO Inami JIZAI Body Project.



**KATSUTOSHI MASAI** received the Ph.D. degree from Keio University, Japan, in 2018.

He worked as a JSPS Research Fellow (DC2), a Project Research Associate, a Research Associate at Keio University. Currently, he is a Research Associate at NTT Communication Science Laboratories. His research interest includes wearable interfaces for sensing human daily activities.



**MAKI SUGIMOTO** received the Ph.D. degree from The University of Electro-Communications, Tokyo, in 2006.

He is currently a Professor with the Department of Information and Computer Science, Faculty of Science and Technology, Keio University, Japan. His research interests include embedded optical sensing systems for AR/VR and wearable sensing systems with machine intelligence.

...