



Potential of Wearable Technology for Super-Aging Societies

Atsushi Shimada^(✉)

Faculty of Information Science and Electrical Engineering,
Kyushu University, Fukuoka, Japan
atsushi@ait.kyushu-u.ac.jp

Abstract. The paper discusses the potential of wearable devices and mainly focuses on smartwatches for super-aging societies. A smartwatch is designed to be worn on the wrist similar to a traditional watch. Additionally, smartwatches involve touch screens, software applications, and IMU (inertial measurement unit) sensors. Smartphones are mainly utilized for notifications related to phone calls, mails, SNS, or healthcare applications to measure and record heart rate and other vital signals. Conversely, smartwatches exhibit significant potential to generate new value in our daily life. In the paper, we introduce three applications with smartwatches published in extant studies and discuss the applicability of smartwatch applications to super-aging societies.

Keywords: Wearable device · Smartwatch · Application
Supporting system · Super-aging society

1 Introduction

Recently, wearable devices, such as smartwatches, constitute popular devices worldwide. A smartwatch is designed to be worn on the wrist similar to a traditional watch. Additionally, smartwatches include touch screens, software applications to notify phone calls, mails, and SNS. Smartphones also display an alternative aspect as sensors that measure daily activities such as location, heart rate, and motions. Given the sensory information, healthcare applications constitute a major area of application of smartwatches.

Several studies involve the use of smartwatches as sensors to measure human activities, to develop new applications that control smartwatch itself [11], or support daily life, working, learning, communication, and man–machine interaction [12] among others. Although several individuals potentially consider smartwatches as youth-focused devices, they are also useful for elderly individuals. For example, a gesture interface provides intuitive and memorable operations to control a TV or other digital equipment. A motion analysis and gesture recognition system examines an elderly individual’s daily life or records daily activities and reports them to his/her family. Therefore, smartwatches exhibit a significant potential to generate new value for super-aging societies.

In the study, we introduce previous studies related to applications using smartwatches and review the studies in terms of applicability to super-aging societies. In the following sections, man-machine interface based on gestures is first introduced [8]. The study aims at realizing user- and machine- friendly interactions. In terms of users, the important concept is ease of use and ease of recollection. Additionally, from the machine viewpoint, the ability to distinguish gestures from gestures used in daily activities is the most important issue. The study proposes a strategy to design good gestures for both aspects. Second, a gesture recording system is introduced [7]. The study measures a farm laborer's activities in a greenhouse and recognizes harvest gestures to automatically record the location and amount of yields. Third, an activity sensing and visualization system is introduced [5]. In the study, activities during group work are measured using sensors. A wearable camera and two smartwatches are used for activity sensing with respect to an individual. Group activities are visualized to understand the role of each individual, the contribution of each individual to the task, and the collaboration of individuals to achieve the goal of group work. Following the introduction of the studies, we describe the potential of smartwatches and their applicability to super-aging societies.

2 Man-Machine Interface Based on Gestures

2.1 Gestures Design

Man-machine interfaces play an important role in conveying intentions from a user to a machine. The keyboard and mouse traditionally provided an interface for personal computers because they are easy to use, and the inputs are easy for the computer to interpret. Recently, gesture interfaces based on computer vision techniques were examined [3, 13, 15, 16] to realize a perceptual user interface. For example, a computer vision-based method was examined, and it recognizes gestures from hand motions or shapes detected by an RGB camera [4]. A primary disadvantage of computer vision approaches is that the sensing area restricts the user to a narrow region where gesture recognition is available.

Several studies related to gesture interfaces mainly focus on the accuracy of gesture recognition and do not focus on gesture design. However, good gesture design is important to enhance usability and reduce false recognition of gestures. A user-customizable gesture interface for a TV control system was proposed in [14]. A user freely selects a gesture (e.g., hand shape and hand motion) and assigns it to an operation (e.g., volume up or volume down). The self-assignment of gestures makes it easier for a user to recollect the gestures. However, candidate gestures are designed by the researchers and are not assessed based on ease of recognition by a machine (system). A few gestures that involve motions that are used frequently in daily life can be falsely recognized as gestures assigned to an operation. Therefore, it is necessary to consider the ease of recollection and performing the gesture (user friendliness) as well as the ease of gesture recognition without false positives (machine friendliness).

A previous study [8,9] discussed gesture design from both user- and machine-friendliness perspectives. The proposed system first extracted low-false-positive gestures based on the random forest approach proposed by Kawahata et al. [9]. We assumed that each gesture consists of two primitive actions. Subsequently, extracted gestures were evaluated by the proposed criteria to create a recommendation ranking. The ranking was adjusted by controlling the weighting of the user-friendliness scores and the machine-friendliness score. We created three ranked lists of gestures that were subsequently presented to three groups of individuals. We investigated the effect of ranking on gesture selection, gesture types that are commonly selected, and the reasons given by respondents for selecting the gestures.

2.2 Gesture Recommendation System

The proposed system presents a user with a ranked list of gestures. In order to rank the gestures, it is necessary to define what improves a gesture for a user or for a machine. For a user, a smaller movement is considered as ‘better’ (and is assigned a lower score) relative to a longer movement. We defined seven primitive hand movement actions, namely “RIGHT”, “LEFT”, “PULL”, “PUSH”, “DOWN”, “UP”, and “ROLL”. We assume that a gesture consists of two successive primitive actions and define the score of successive actions A having smaller score with small movement. Specifically, the A consists of two scores, and interested readers can refer to the literature [8] for details.

Machines require a gesture that is easy to recognize. We directly utilize a recognition ratio as a recognition score R . It should be noted that the recognition ratio is stored in the system when the action database was created and tested for recognition accuracy of each action.

In order to control the weighting of the user-friendliness scores and the machine-friendliness score, A , and R are aggregated as E by using alpha blending as follows:

$$E = \alpha R + (1 - \alpha)A, \quad (1)$$

where the α denotes a parameter to control the weighting. If we assign a high value to α , machine friendliness is assigned a significantly higher weighting than user friendliness. The score E is calculated for all gestures to create a ranking. We prepared three gesture candidate lists by changing the weighting parameter α . Table 1 shows the three candidate lists generated by setting $\alpha = 0.1$ (user friendly), $\alpha = 0.5$ (neutral), and $\alpha = 0.9$ (machine friendly).

2.3 Result of User Study

We investigated how each gesture was selected and assigned to a command to commence or control a smartphone application. We assumed that a gesture interface operates applications on a smartwatch. In the experiments, we prepared two control categories including the operation of a music player and starting an

Table 1. Gesture candidate lists generated by $\alpha = 0.1$ (user friendly), $\alpha = 0.5$ (neutral), and $\alpha = 0.9$ (machine friendly)

Group U ($\alpha = 0.1$)		Group M ($\alpha = 0.5$)		Group S ($\alpha = 0.9$)	
Ranking	Gesture	Ranking	Gesture	Ranking	Gesture
1	DOWN_ROLL	1	DOWN_ROLL	1	DOWN_ROLL
2	ROLL_UP	2	ROLL_UP	2	DOWN_DOWN
3	DOWN_UP	3	DOWN_UP	3	DOWN_RIGHT
4	DOWN_RIGHT	4	DOWN_RIGHT	3	RIGHT_DOWN
4	RIGHT_DOWN	4	RIGHT_DOWN	5	ROLL_UP
6	RIGHT_UP	6	DOWN_DOWN	6	DOWN_UP
7	DOWN_LEFT	7	RIGHT_UP	7	RIGHT_UP
7	DOWN_PULL	8	DOWN_LEFT	8	DOWN_LEFT
7	LEFT_DOWN	8	DOWN_PULL	8	DOWN_PULL
7	PULL_DOWN	8	LEFT_DOWN	8	LEFT_DOWN
11	DOWN_PUSH	8	PULL_DOWN	8	PULL_DOWN
11	PUSH_DOWN	12	DOWN_PUSH	12	DOWN_PUSH
13	RIGHT_PUSH	12	PUSH_DOWN	12	PUSH_DOWN
14	LEFT_UP	14	RIGHT_PUSH	14	RIGHT_PUSH
14	PULL_UP	15	LEFT_UP	15	UP_UP
14	UP_PULL	15	PULL_UP	16	LEFT_UP
17	PUSH_UP	15	UP_PULL	16	PULL_UP
18	LEFT_PUSH	18	PUSH_UP	16	UP_PULL
19	DOWN_DOWN	19	UP_UP	19	PUSH_UP
20	UP_UP	20	LEFT_PUSH	20	LEFT_PUSH

Table 2. Eight operations of two categories

Category	Operation	Category	Operation
Music player	Play music	Application start	Weather
	Stop music		Music player
	Next		Map
	Previous		Pedometer

application. Each category contained four operations such that a total of eight operations are controlled by gestures as shown in Table 2.

We conducted an Internet-based survey to assess the gestures most likely to be assigned to each operation. A total of 351 individuals including old and young individuals responded to our survey. We assigned respondents to one of three groups with similar attributes to the maximum possible extent. We asked respondents to assign a gesture to each of the eight operations. The respondents

selected gestures based on the candidate list that they were shown; a different candidate list was shown to each group. A gesture list generated by $\alpha = 0.1$ was shown to Group U, the gesture list generated by $\alpha = 0.5$ was shown to Group M, and the gesture list generated by $\alpha = 0.9$ was shown to Group S.

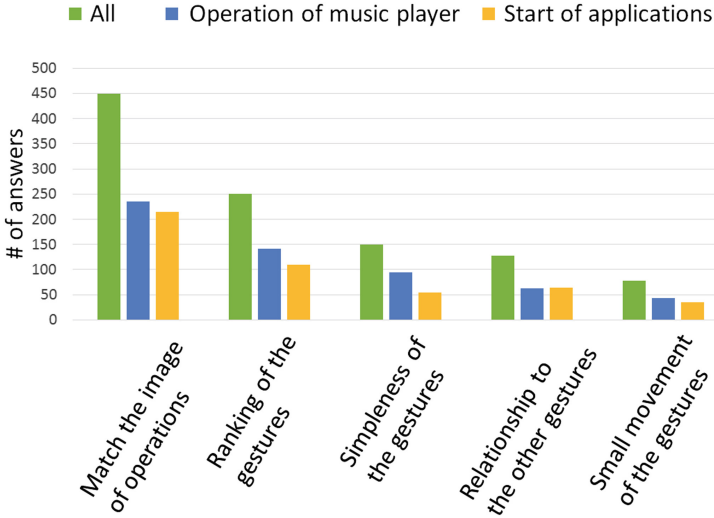


Fig. 1. Reason for gesture assignment. The vertical axis is quantity of answers. Multiple selection of answers was allowed.

We investigated the assignment results of each group, and they revealed that the results were not sensitive to changes in α . Given the page limitation, we recommend readers to refer to the literature [8] for more detailed results. In the study, we discuss the reason as why respondents selected and assigned the gesture to each operation. Figure 1 shows the questionnaire respondents’ reasons for assigning each gesture to an operation. The answers are listed on the horizontal axis, and the frequency of the answer (i.e., the number of respondents that selected the answer) is indicated on the vertical axis. The most common answer was “The gesture matches the operation”. Respondents tended to select gestures that were easily associated with the given operation. Even the most commonly assigned gestures were not selected by all respondents. Thus, the results indicate that respondents possessed their own impressions of each operation. The second most common answer was “The gesture is highly ranked”. Although the frequency is lower than that for “The gesture matches the operation”, the rank of the proposed gestures corresponded to a significant effect. The other three factors were not significantly considered by respondents.

Finally, the conclusions of the study were as follows.

- Individuals tend to select the highest ranked gestures from a list.
- It is unnecessary to assign strong weighting to user-friendly gestures that involve small movements.
- Individuals tend to select gestures that are easily associated with the operations.
- Individuals tend to select symmetric gestures for symmetric operations.

We recommend readers to refer to literature [8] for detailed reports.

3 Harvest Action Recognition in Tomato Yields

3.1 Measurement of Farm Work

Farm managers are required to make a few decisions when they manage their farm field or work plan. In the aforementioned decisions, information including that related to the plants and field environment is required. Recently, sensors for environmental information were introduced into farm fields and aid farm managers in decision making.

Our previous study examined an issue to create a system that automatically measures harvesting work of farm laborers and visualizes the spatial distribution of tomato yields. The system measured the position and action information of farm laborers with smart devices and visualized the spatial distribution as a harvesting map.

Position estimation is performed using beacons and smartphones. We placed 150 beacons that broadcast Bluetooth UUID (Universal Unique Identifier: a 128-bit number used to identify the beacon), and each farm laborer possesses a smartphone that receives the signals for position estimation. Based on the received signals, the system estimates the section where the farm laborer is working at any given second. The details of the position estimation method are described in [6].

The greenhouse where the experiments were performed includes 21 passages where farm laborers walk, and the width of the passage is 1.3 [m] (Fig. 2). Specifically, 20 ridges exist between passages where tomatoes are planted, and the length of the ridge is 45 [m]. Additionally, 15 pillars that support the greenhouse are aligned in a passage, and the length between two pillars is 3 [m].

3.2 Harvest Action Recognition

The system recognizes a harvest action performed by a farm laborer. The harvest procedure consists of the following four steps.

1. Select a tomato for harvesting.
2. Pick the tomato from the tomato plant.
3. Cut the stem with scissors.
4. Put the tomato in a container.

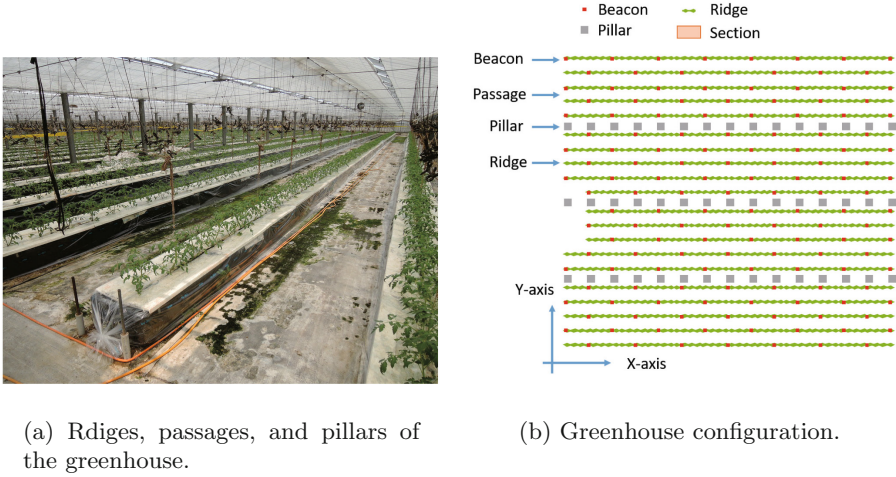


Fig. 2. The greenhouse where we conducted experiments.

In our implementation, we focus on the fourth step when a farm laborer puts a tomato in a container because it clearly specifies the timing when a tomato is harvested. Other steps involve ambiguity as to whether the farm laborer actually harvests a tomato.

A series of harvesting actions of a farm laborer is shown in Fig. 3. In order to recognize harvesting action of farm laborers, we attached two smartwatches on both wrists of the farm laborers. Figure 4 shows a time series data when the farm laborer performed the harvesting action shown in Fig. 3. The farm laborer put a tomato in a container with left hand, and thus a time series data of the left hand exhibited more significant changes when compared to that of the right hand.

The system classifies the harvest actions from others. Several procedures for action classification exists beginning with feature extraction from time-series data. Subsequently, a machine learning approach based on Random Forest [1, 10] is applied. In the recognition process, we address continuous recognition of time-series data by dividing the series into subsequences. The classifier investigates as to whether each subsequence is attributed to a harvest action. Please refer to the literature [7] for more detailed procedures.



Fig. 3. A series of harvesting actions of a farm labor. He is putting a tomato in a container.

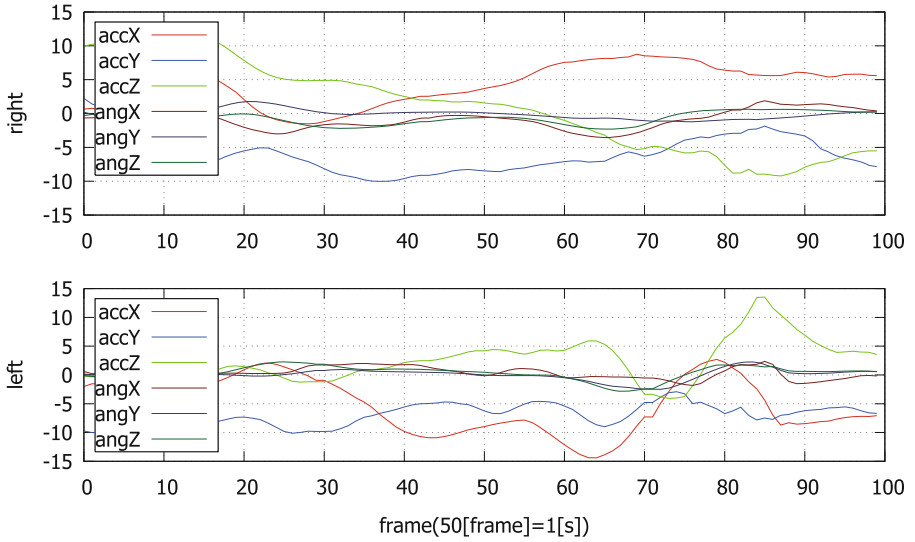


Fig. 4. Acceleration and angular velocities of the right and left wrists and of a farm labor.

3.3 Harvest Map

We conducted experiments for two weeks that included position estimation and action recognition, and each week includes four working days of harvesting tomatoes. Given the page limitations, we introduce results of estimation of yield amount based on action recognition and a harvest map generated by position estimation and action recognition.

Table 3. Counting the number of harvesting tomatoes. Three farm laborers (F1, F2 and F3) joined the experiments.

Day	F1	F2	F3
	Number (man/sys)	Number (man/sys)	Number (man/sys)
1	640/885	160/250	320/144
2	744/963	240/362	-
3	560/672	320/310	160/121
4	480/607	80/147	-
5	448/336	128/133	80/86
6	504/680	104/137	-
7	424/305	-	-
8	184/222	-	-
MAE	187	73	56

The results of yield amount are shown in Table 3. The number indicates the number of tomatoes that are harvested by each farm laborer, “man” denotes the number of tomatoes counted manually, and “sys” denotes the number of tomatoes counted by the proposed system. The system estimates the amount within approximately 20% error.

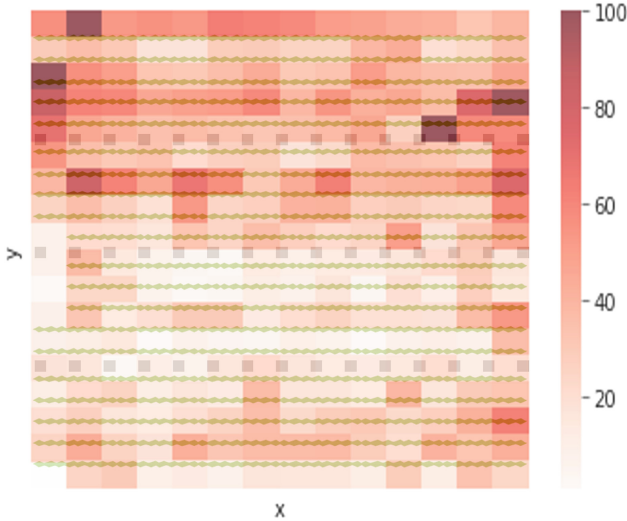


Fig. 5. Harvesting map over two weeks

Finally, Fig. 5 shows the harvest map generated by our system. We can infer the tendency of yield amount, and this indicated that the amount of upper side of the map exceeded those of middle and bottom of the map. Please refer to literature [7] for more details.

4 Activity Sensing in Collaborative Work

4.1 Sensing of Students' Activities

Collaborative learning is an efficient method to encourage students to explore and solve problems in conjunction with members who possess different abilities and thoughts, and it promotes various skills including oral communication and leadership. In the case of collaborative learning, formative assessment is increasingly important to evaluate learning process of each student. An important and expected evaluation in real world collaborative learning involves checking whether a student can share his/her opinions or ideas in conjunction with other students [2].

In order to develop technologies for measuring attention level relative to a speaker and activity synchronization level, we developed a system for automatic

sensing and visualization of real world activities in collaborative learning. In the experiment, the task involves group work for building a town diorama by using LEGO blocks. In order to measure subjects' attentions and activities, first person view camera and wearable activity sensor (smartwatch) are attached to each subject. In order to analyze group activities, all cameras and sensors were fully synchronized with manual adjustments.

4.2 Visualizer of Group Activities

We developed a visualization tool as shown in Fig. 6. The visualizer provides three sub windows showing a video sequence and two waveforms that are automatically synchronized by referring to the timestamp of the data.

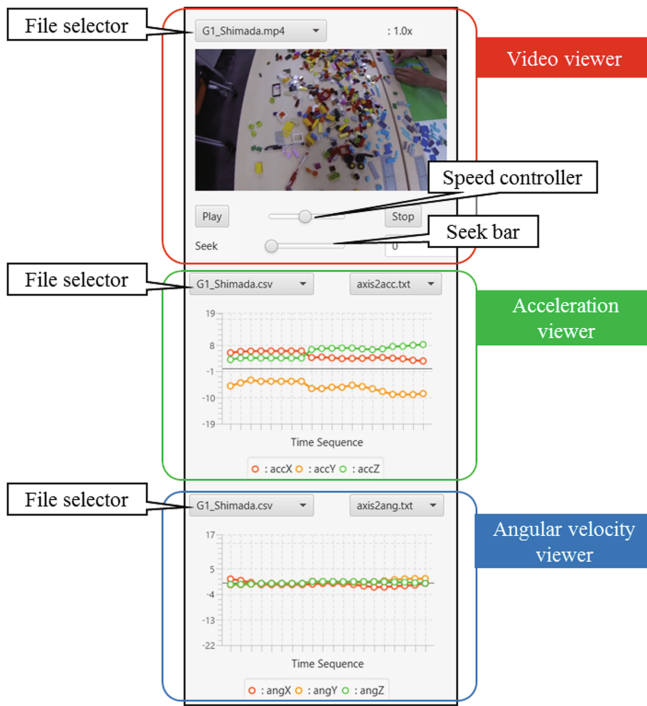


Fig. 6. A component of visualizers for camera and IMU sensors

We extended the above visualizer to compare the activities in the same group. The extended visualizer includes three components (a component corresponds to an individual). Figure 7 shows examples of visualization during the cooperative development work. All members of both group paid attention to the facilitator when the facilitator provided an explanation of the task of cooperative work with a whiteboard. In the period, acceleration and angular velocity values were



Fig. 7. Visualization of group activity a cooperative development work

almost flat, and thus hand motions were absent. In contrast, IMU sensor values were frequently changed when the members constructed LEGO blocks. Please refer to literature [5] for more detailed information.

5 Conclusion

The study examined the potential of smartwatches and their applicability to super-aging societies. Three studies were introduced as applications based on activity sensing, namely a gesture control system, an automatic recording system for farm laborers' work, and a monitoring system for group activities. The systems and their backbone technologies are useful in terms of enriching daily life, working, and analytics methodologies as well as realizing safety and security of society. Specifically, gesture interface technologies support seamless interaction between elders and sophisticated equipment to provide assistance to super-aging societies. Activity sensing and the monitoring of elders reassures family members living at far distances.

In a future study, the following issues should be continuously discussed to ensure the widespread use of smartwatches for super-aging societies.

- Low-cost technologies aimed at long-term sensing
- Real-time system to recognize activities
- Infrastructure providing high-speed communication between smartwatches and systems.

It is expected that the results of this study will overcome the aforementioned issues in the immediate future.

Acknowledgements. This work was partially supported by JST PRESTO Grant Number JPMJPR1505, JSPS KAKENHI Grant Number JP16K12786, JP15H01621 and JP15H01695, Japan.

References

1. Breiman, L.: Random forests. *Mach. Learn.* **45**(1), 5–32 (2001)
2. Damsa, C.I., Kirschner, P.A., Andriessen, J.E.B., Erkens, G., Sins, P.H.M.: Shared epistemic agency: an empirical study of an emergent construct. *J. Learn. Sci.* **19**(2), 143–186 (2010)
3. Chen, M.Y., Mummert, L., Pillai, P., Hauptmann, A., Sukthankar, R.: Controlling your TV with gestures. In: *Proceedings of the International Conference on Multimedia Information Retrieval, MIR 2010*, pp. 405–408. ACM, New York (2010). <http://doi.acm.org/10.1145/1743384.1743453>
4. Chen, Q., Georganas, N.D., Petriu, E.M.: Real-time vision-based hand gesture recognition using Haar-like features. In: *Instrumentation and Measurement Technology Conference Proceedings, 2007, IMTC 2007*, pp. 1–6. IEEE (2007)
5. Deguchi, D., Kondo, K., Shimada, A.: Subjective sensing of real world activity on group study. In: *The Eighth International Conference on Collaboration Technologies (CollabTech 2016)*, pp. 5–8, September 2016
6. Hashimoto, Y., Arita, D., Shimada, A., Okayasu, T., Uchiyama, H., Taniguchi, R.: Farmer position estimation in a tomato plant green house with smart devices. In: *Proceedings of International Symposium on Machinery and Mechatronics for Agriculture and Biosystems Engineering (ISMAB) 2016*, pp. 200–205, May 2016
7. Hashimoto, Y., Arita, D., Shimada, A., Uchiyama, H., Taniguchi, R.: Visualization of spatial distribution of tomato yields based on action recognition. In: *The International Workshop on Frontiers of Computer Vision (FCV) 2018*, February 2018
8. Kawahata, R., Shimada, A., Taniguchi, R.: What are good design gestures? - towards user- and machine-friendly interface-. In: *23rd International Conference on Multimedia Modeling*, January 2017
9. Kawahata, R., Shimada, A., Yamashita, T., Uchiyama, H., Taniguchi, R.: Design of a low-false-positive gesture for a wearable device. In: *5th International Conference on Pattern Recognition Applications and Methods*, pp. 581–588, February 2016
10. Liaw, A., Wiener, M.: Classification and regression by randomforest. *R News* **2**(3), 18–22 (2002)

11. Park, T., Lee, J., Hwang, I., Yoo, C., Nachman, L., Song, J.: E-gesture: a collaborative architecture for energy-efficient gesture recognition with hand-worn sensor and mobile devices. In: Proceedings of the 9th ACM Conference on Embedded Networked Sensor Systems, pp. 260–273. ACM (2011)
12. Ruiz, J., Li, Y., Lank, E.: User-defined motion gestures for mobile interaction. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 197–206. ACM (2011)
13. Shan, C.: Gesture control for consumer electronics. In: Shao, L., Shan, C., Luo, J., Etoh, M. (eds.) Multimedia Interaction and Intelligent User Interfaces. Advances in Pattern Recognition. Springer, London (2010). https://doi.org/10.1007/978-1-84996-507-1_5
14. Shimada, A., Yamashita, T., Taniguchi, R.: Hand gesture based TV control system -towards both user- & machine -friendly gesture applications-. In: The 19th Japan-Korea Joint Workshop on Frontiers of Computer Vision, pp. 121–126, January 2013
15. Wachs, J.P., Kölsch, M., Stern, H., Edan, Y.: Vision-based hand-gesture applications. Commun. ACM **54**, 60–71 (2011). <http://doi.acm.org/10.1145/1897816.1897838>
16. Zabulis, X., Baltzakis, H., Argyros, A.: Vision-based hand gesture recognition for human-computer interaction, pp. 1–30. Lawrence Erlbaum Associates, Inc. (LEA) (2009). http://www.ics.forth.gr/~xmpalt/publications/papers/zabulis09_book.pdf