



# Visualization of Farm Field Information Based on Farm Worker Activity Sensing

Daisaku Arita<sup>1</sup>(✉), Yoshiki Hashimoto<sup>2</sup>, Atsushi Shimada<sup>2</sup>,  
Hideaki Uchiyama<sup>2</sup>, and Rin-ichiro Taniguchi<sup>2</sup>

<sup>1</sup> Faculty of Information Systems, University of Nagasaki, Nagayo, Nagasaki, Japan  
arita@sun.ac.jp

<sup>2</sup> Graduate School of Information Science and Electrical Engineering,  
Kyushu University, Fukuoka, Japan

**Abstract.** Our research goal is to construct a system to measure farm labor activities in a farm field and visualize farm field information based on the activities. As the first step for the goal, this paper proposes a method to measure harvesting information of farm labors in a tomato greenhouse and to visualize the tomato yield distribution in the greenhouse, we call it a harvesting map, for supporting the farm managers making decisions. A harvesting map shows daily, weekly and monthly tomato yields in small sections into which the tomato greenhouse is divided.

**Keywords:** Smart agriculture · Farm work information  
Action recognition · Position estimation · Wearable sensor

## 1 Introduction

Farm managers are required to make some decisions when they manage their farm field or work plan. And, farm labors works according to the decisions by the farm managers. For example, a farm manager decides the times when they sow tomato seeds, when they harvest tomato, and when they cut dead leaves of tomato plant. When they make such decisions, information about the plants and the field environment, and so on is required. Recently, sensors for environmental information have been introduced into farm fields and helped farm managers taking decisions.

Although introduction of sensors to greenhouses have been proceeded, the field environmental condition is not spatially uniform even in a greenhouse. Due to the lack of uniformity, there are spatial variations of the tomato yields. A farm manager of a greenhouse conducting our experiments, does not know the spatial variation of the tomato yields.

The purpose of this research is to construct a system that automatically measures harvesting work of farm labors and visualizes the spatial distribution of tomato yields. By providing the information to farm managers, they can grasp the variation in the tomato yields, which leads to cultivation support.

## 2 Harvesting Map

A harvesting map is a map which visualizes the spatial distribution of yields in a greenhouse. The objective of generating harvesting maps is to inform farm managers about the farm conditions and to help them decide what farm work should be done.

The greenhouse where we conducted experiments has 21 passages where farm labors walk, and the width of the passage is 1.3[m] (Fig. 1). There are 20 ridges between passages where tomatoes are planted, and the length of a ridge and a passage is 45[m]. In three of the passages, 15 pillars which support the roof of the greenhouse are aligned, and the distance between two adjacent pillars is 3[m].

Each passage is divided into small sections constituting the units for measurement and visualization based on the ridges and pillars (Fig. 1(b)). To identify which section the farm labor works on, we have set the X-axis as the axis along the ridge and the Y-axis as the axis across ridges. The size of one section is 1.3[m]  $\times$  3[m], and the number of the section defined in the greenhouse is 336, because there are 21 passages and each passage has 16 sections.

## 3 Measurement of Farm Work

In this paper, we propose a system which measures the harvesting work to visualize the spatial distribution of tomato yields in the greenhouse. The system measures position and action information of farm labors with smart devices, and visualize the spatial distribution as a harvesting map.

### 3.1 Position Estimation

In this section, we present a method to estimate the position of a farm labor in a greenhouse. We have placed 150 beacons which broadcast Bluetooth UUID (Universal Unique Identifier: a 128-bit number used to identify the beacon) and each farm labor has a smartphone that receives these signals for position estimation. Based on the received signals, the system estimates the section where the farm labor is working in every one second.

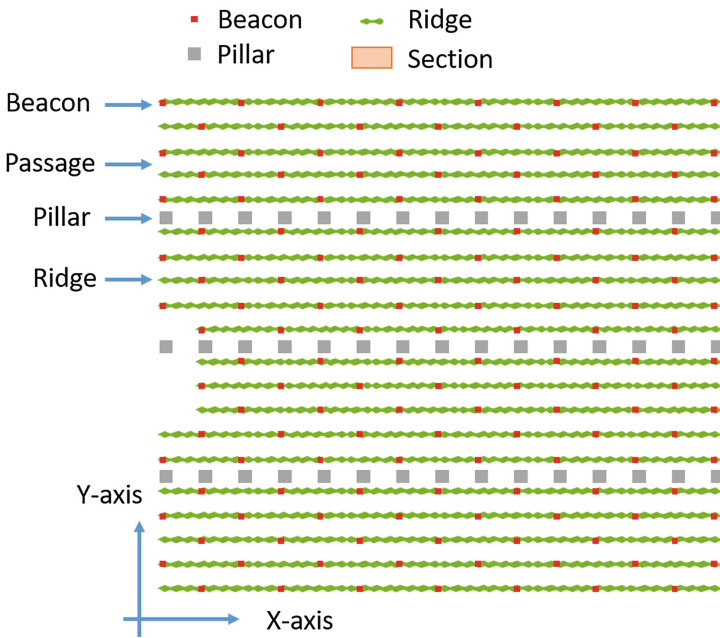
The method consists of three steps as follows. First, the farm labor's approximate position is estimated from signals broadcast by multiple beacons. Next, the X-position is smoothed with the mode function. The final step is smoothing the Y-position using the map matching technique. Finally, the system obtains a time series of a farm labor  $P_n = (x_n, y_n)$ , which indicates the position where the farm labor is working at discrete time  $n = \frac{t}{T^P}$  and  $T^P$  is the sampling interval of the beacon signal reception. The details of the position estimation method are described in [1].

### 3.2 Action Recognition

The system estimates the time when a farm labor harvests a tomato by recognizing specific actions made by the farm labor. Farm labors harvest tomatoes by repeating four actions listed below.

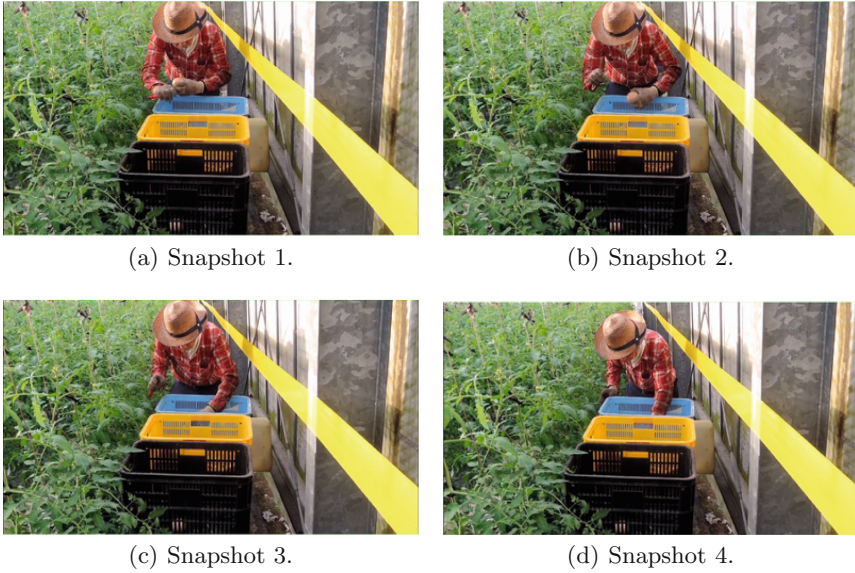


(a) Ridges, passages, and pillars of the greenhouse.



(b) Greenhouse configuration.

**Fig. 1.** The greenhouse where we conducted experiments.



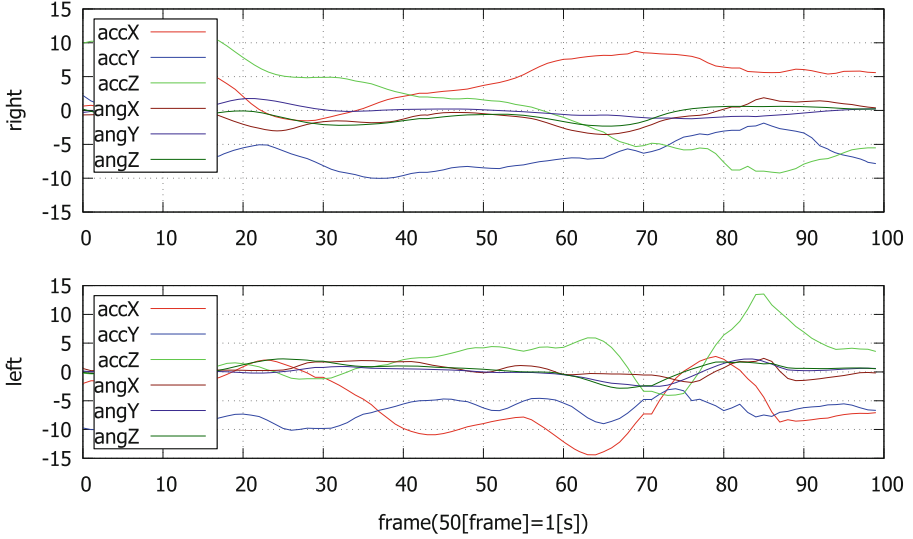
**Fig. 2.** A series of harvesting actions of a farm labor. He is putting a tomato in a container with left hand.

1. Search a tomato to be harvested.
2. Pick the tomato from the tomato plant.
3. Cut the stem off with scissors.
4. Put the tomato in a container.

In this experiment, we focused on the 4th action, since all of farm labors in the greenhouse perform the action uniformly, and it is easy to know when a farm labor harvests a tomato. In addition the other actions are difficult to recognize, because the actions are performed uniquely in each farm labor and do not have specific motions to recognize. This action is defined as the harvesting action, and the other actions including not only three actions listed above but also unrepeated actions such as carrying a container, wiping the seat, and so on are defined as the normal action. A series of harvesting action of a farm labor is shown in Fig. 2.

In order to recognize harvesting action of farm labors, farm labors wear two smartwatches on both of their wrists. The smartwatch has an embedded IMU (Inertial Measurement Unit) sensors, which is able to measure a time series of triaxial accelerations and triaxial angular velocities. Figure 3 shows a time series data when the farm labor performed the harvesting action shown in Fig. 2. The farm labor put a tomato in a container with left hand, hence a time series data of the left hand changed more significantly than the right one.

The system classifies all actions in a harvesting work into the harvesting action and the normal action by the acceleration and the angular velocities. To classify actions in a harvesting work, first, raw time series data are smoothed

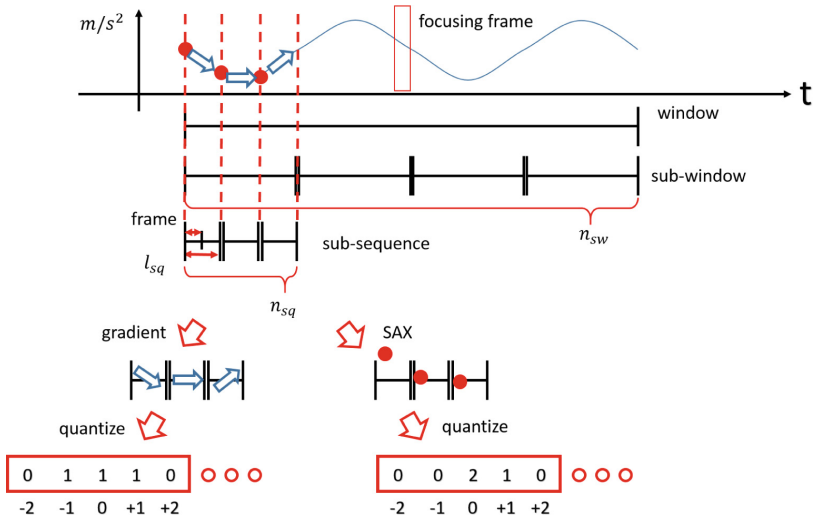


**Fig. 3.** Acceleration and angular velocities of the right and left wrists and of a farm labor.

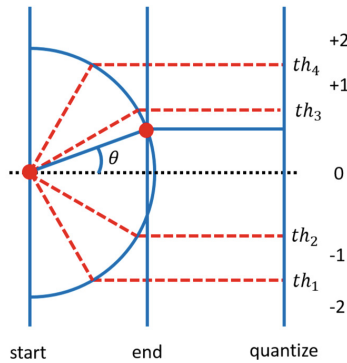
because the sensor data include a considerable amount of high-frequency noise in each axis, which hinders high recognition performance. To smooth the raw time series data, we apply a weighted moving average to each of triaxial accelerations and angular velocities.

A feature vector represents a time series of acceleration and angular velocity data in a fixed window size  $l_w$ , and, as shown in Fig. 4(a), the window-sized data is divided into multiple sub-windows, each of which is, again, divided into multiple sub-sequences. Here, we represent the number of sub-windows of a window is  $n_{sw}$ , the number of sub-sequences in a sub-window is  $n_{sq}$ , and the length of a sub-sequence, or the number of frames in a sub-sequence is  $l_{sq}$ . Therefore one window size  $l_w$  is defined as  $l_w = n_{sw} \times n_{sq} \times l_{sq}$ .

Next, a sequence of acceleration and angular velocity data in a sub-sequence is transformed into a single quantized value, where the number of quantization level is five in this study. To achieve the quantization, we have used two representations: one is Symbolic Aggregate approxIMATION (SAX) [2]; the other is gradient of acceleration and angular velocity [3,4]. In SAX, a sequence is symbolically represented, and, here, the acceleration and angular velocity sub-sequence is transformed into a single constant value, which is quantized with a small number of quantization levels. The gradient of acceleration and angular velocity is calculated as the angle between the start and end values of the sub-sequence, and it is also quantized by simple thresholding shown in Fig. 4(b). Then, a feature histogram, or a histogram of the quantized data, is calculated in each sub-window as shown in Fig. 4(a). Finally, the histograms generated in all the sub-windows in a window are concatenated to represent the feature vector of the window.



(a) Window for calculating two features.

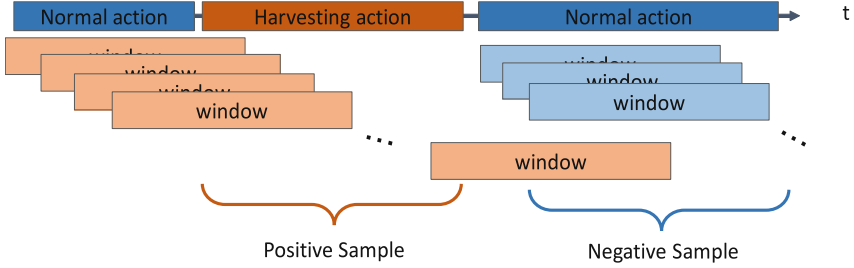


(b) Features quantized into five levels.

**Fig. 4.** Diagrammatic depiction of the calculation of a feature vector. To calculate a feature vector of the focusing frame, a fixed window is extracted, and the feature vector is calculated from the triaxial acceleration in the window. In each sub-sequence, two features (SAX and the gradient) are calculated, and in each sub-window, a histogram is created. (a) Two histograms are shown within a single sub-window. Next, eight histograms (four histograms for each SAX and the gradient feature) are created and the histograms are concatenated to represent the feature vector of the window. (b) Shows how feature value is quantized into five levels: for example, the gradient between start and end values is between  $th_2$  and  $th_3$ , therefore the gradient value is quantized into “0”.

Our goal is to recognize the harvesting actions from the entire time series data acquired by harvesting work. Therefore, the recognition is a two-class discrimination problem: the harvesting actions as a positive class and the normal actions as a negative class. In other words, a one-vs-rest strategy is applied [5], and the classification is performed in Random Forest [6, 7] which is one of the





**Fig. 5.** Windowed data are labeled based on the class labels. Windowed data which overlaps with the harvesting action are labeled as positive samples, and the others are labeled as negative samples.

machine learning method. Here, we extract windowed data, frame by frame, from the entire acceleration and angular velocity sequence of all the harvesting work, and each of windowed data is represented in a feature vector. Windowed data which overlaps with the harvesting action are labeled as positive samples, and the others are labeled as negative samples (Fig. 5). Then Random Forest is trained with the feature vectors of those positive samples and negative samples.

In a testing phase, we extract windowed data as same as train phase, and each of windowed data is represented in a feature vector. When the one-vs-rest strategy is applied, Random Forest produces a time series of a posteriori probability ( $0 \leq P_m \leq 1, 0 \leq m < M$ ) as the output for each class, in this experiment, for the harvesting action or the normal action within each frame, where  $m$  is discrete time  $m = \frac{t}{T^A}$  and  $T^A$  is the sampling interval of the IMU.

To locate the harvesting actions in the sequence, we have set the following rules.

1. Representative time of each harvesting action is decided by finding, in the sequence, local maximum of  $P_m$  which is greater than  $th_p$ .
2. If the difference between a representative time and its following one is smaller than  $2 \times l_w$  frames, the following one is ignored.
3. Based on the local maximum  $P_m$ , a harvesting action  $A_m$  is determined, which indicate a farm labor harvests a tomato or not at discrete time  $m$ .

### 3.3 Generating Harvesting Map

To generate a harvesting map, the system measured information about position and action of farm labors according to the method outlined previous subsections respectively. Position information is obtained as  $P_n^f = (x_n, y_n)$ , and it indicates the section in a greenhouse of a farm labor  $f$  at a discrete time  $n$ . Action information is obtained as  $A_m^f = \{0, 1\}$ , and it indicates that the farm labor harvests a tomato ( $A_m^f = 1$ ) or not ( $A_m^f = 0$ ) at a discrete time  $m$ . To generate a harvesting map, they had to be combined, because these two types of information are obtained separately.

First, the discrete time of position and action information is adjusted based on the time of the action information, in order to know the section  $P_n^f$  where

a farm labor harvests a tomato with the harvesting action at time  $m$ . Therefore position information  $P_n^f$  is converted to  $P_m^f$  by copying,  $P_m^f = P_n^f$ , where  $n = \lfloor m \frac{T^A}{T^P} \rfloor$ . The harvesting map of farm labor  $f$ ,  $H_p^f$ , is the 2-dimensional histogram of  $\{\mathbf{p} = P_m^f | A_m^f = 1\}$ , each bin  $\mathbf{p}$  of which indicates the number of tomatoes harvested in section  $\mathbf{p}$  by farm labor  $f$ . Finally, the harvesting map  $H_p$  is generated by the equation  $H_p = \sum_f H_p^f$ .

## 4 Experiment and Result

### 4.1 Confirming Action Recognition

In order to verify the proposed method on action recognition, time series data of acceleration and angular velocity of a farm labor's both wrists during harvesting work were measured. An experiment was conducted to see whether the time of harvesting action is correctly recognized. Three farm laborers' (F1, F2, F3) actions were measured three times, and the data is used for training and recognition. The smartwatch used for the measurement is moto 360 sport, and the measurement time is about 30 to 60 min. The measurement frequency of the smartwatch is 50[Hz]. The parameters in this experiment are used as  $T^P = 1[\text{sec}]$ ,  $T^A = 0.02[\text{sec}]$ ,  $n_m = 5$ ,  $n_s = 5$ ,  $l_s = 2$ ,  $th_s = \{-9, -6, -3, -1.5, 0, 1.5, 3, 6, 9\}$ ,  $th_p(F1) = 0.65$ ,  $th_p(F2) = 0.6$ ,  $th_p(F3) = 0.7$ .

The results of the action recognition is shown in Table 1. In the experiment, training of Random Forest was carried out with the first day of each farm labor,

**Table 1.** Result of the action recognition of three farm laborers.

Day		F1	F2	F3
1 (Train)	Precision	144/173	139/225	64/81
		0.83	0.62	0.79
	Recall	144/144	139/140	64/64
1.00		0.99	1.00	
	F-measure	0.91	0.76	0.88
2 (Test)	Precision	94/131	16/66	47/76
		0.72	0.24	0.62
	Recall	94/115	16/115	47/54
0.82		0.14	0.87	
	F-measure	0.76	0.18	0.72
3 (Test)	Precision	158/221	84/121	49/79
		0.72	0.69	0.62
	Recall	158/202	84/129	49/100
0.78		0.65	0.49	
	F-measure	0.75	0.67	0.54



and the test was conducted with the first, second and third data. The system can recognize the actions of farm labor F1 with stable accuracy. However, the accuracy of action recognition of farm labor F2 and F3 is not so high, and there seems to be two reasons for this low accuracy.

The first reason is how smartwatches were worn. Farm labors F2 and F3 were wearing their smartwatches over their clothes, therefore the smartwatches

**Table 2.** Counting the number of harvesting tomatoes.

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



(a) Snapshot 1.



(b) Snapshot 2.

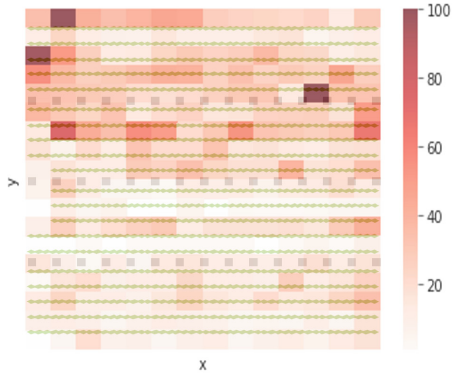


(c) Snapshot 3.

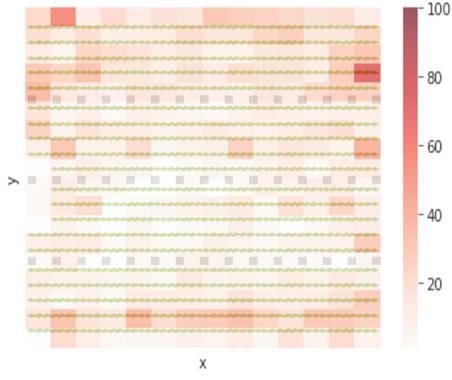


(d) Snapshot 4.

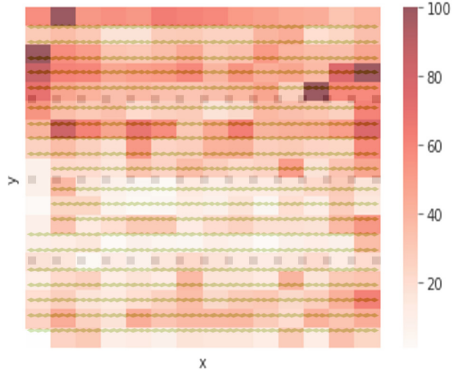
**Fig. 6.** Harvesting action accidentally detected. He took a tomato which is once classified as A quality, and put B quality.



(a) Harvesting map of the first week.



(b) Harvesting map of the second week.



(c) Harvesting map for two weeks.

**Fig. 7.** Generated harvesting map. (Color figure online)

can roll and slide easily. This makes the action recognition more difficult. On the other hand, farm labor F1 wears a smartwatch directly on his skin, and the position is considered to be stable.

The second reason is re-classification. It is necessary for farm labors to select a container from two containers in which they put each tomato according to the tomato quality. Farm labor F2 and F3 often re-classify a tomato after they put it in a container. This action is much similar to the harvesting action and causes over-recognition of the harvesting action. On the other hand, farm labor F1 rarely re-classifies tomatoes.

## 4.2 Creating a Harvesting Map

We conducted experiments for two weeks which included position estimation and action recognition, and each week has four working days of harvesting tomato. The results of position estimation is 86% (average of six farm labors to estimate the passage where they were working). The result of action recognition is shown in Table 2. In the table, number means how many tomatoes are harvested by each farm labor, and “man” means the number of tomatoes counted manually and “sys” means the number of tomatoes counted by the proposed system. The results of position estimation and action recognition is not high enough, and it is required to improve the accuracy.

The system combines position and action information into tomato yields information, and visualize them as harvesting maps for two weeks (Fig. 7). The color of the harvesting map shows the number of harvested tomatoes in each section. The spatial distribution of the tomato yields in the greenhouse is confirmed with the harvest maps.

## 5 Conclusion

This paper proposed a system visualizing the spatial distribution of tomato yields in a greenhouse, and two experiments were conducted. Firstly, a experiment to recognize the harvesting action of three farm labors is conducted. Next, harvesting maps for two weeks are generated based on a experiment to measure harvesting work of three farm labor. By visualizing the number of tomato yields as the harvesting map, it is confirmed that the spatial distribution of tomato yields in the greenhouse.

As for the next task, since the subjects of action recognition were three in this experiment, it is necessary to recognize the action of the remaining one farm labor. The information of the farm labor is necessary to generate a harvesting map for the system, because there usually four farm labors harvest tomatoes. In addition, it is required to calculate the accuracy of the harvesting map, and also to be evaluated by the farm manager. We aim to visualize information obtained by the system as a harvesting map over a long period of time, and to provide these information to farm managers for making decisions.

**Acknowledgement.** This work was partially supported by JSPS KAKENHI Grant Numbers JP15H01695 and JP17H01768.

We would like to thank Shunsuke Hamachi and his labors, a tomato farm labor in Fukuoka, for their cooperation in obtaining farm work information in his tomato greenhouse.

## References

1. 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, pp. 200–205 (2016)
2. Lin, J., Keogh, E., Lonardi, S., Chiu, B.: A symbolic representation of time series, with implications for streaming algorithms. In: Proceedings of the 8th ACM SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery, pp. 2–11. ACM (2003)
3. Zhang, M., Sawchuk, A.A.: Motion primitive-based human activity recognition using a bag-of-features approach. In: Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium, IHI 2012, pp. 631–640. ACM, New York (2012)
4. 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
5. Weston, J., Watkins, C.: Multi-class support vector machines. Technical report, Citeseer (1998)
6. Breiman, L.: Random forests. *Mach. Learn.* **45**(1), 5–32 (2001)
7. Liaw, A., Wiener, M.: Classification and regression by randomforest. *R News* **2**(3), 18–22 (2002)