Development of Motion Control Using the Kinect Sensor for the *Kansei* Communication Interface

Y. Sasaki [1*], S. Shibusawa[2], H. Negishi [3] and M. Emmi[3]
[1*]Faculty of Regional Environment Science, Tokyo University of Agriculture, 1-1-1, Sakuragaoka, Setagaya, Tokyo 156-8502, Japan, y3sasaki@nodai.ac.jp
[2]Dept. Environmental and Agricultural Engineering, Tokyo University of Agriculture and Technology, 3-5-8 Saiwai-cho, Fuchu, Tokyo 183-8509, Japan
[3]Graduate School of Agriculture, Tokyo University of Agriculture, 1-1-1, Sakuragaoka, Setagaya, Tokyo 156-8502, Japan.

ABSTRACT

Japanese farmers are aging and agricultural robots of the cooperative working type, which perform tasks together with humans, are required in particular. Additionally, in the case of the cooperative working type, compatibility with humans is important and the functions that can be controlled intuitively even by new agriculture workers and elderly people are desired. The function is *Kansei* communication, and we proposed agricultural robots, such as the *Kansei* Agri-robot ,which is equipped with the function, and the *Chinou* robot, which extracts tacit knowledge. We have been studying and developing them. In this paper, we built and evaluated an intuitive control part using motion, which is one of the core techniques. We built the system using the Kinect sensor, which can trace the skeleton information of a human. The Kinect sensor is a gaming device for the Xbox 360 and was released by Microsoft Corporation in 2010. It consists of an infrared light for distance sensor, video sensor, distance sensor and multi-array microphone. The target motion control was the “finger pointing” motion to provide the robot a working area or location for movement. As for the development environment, we used Windows 7 as the OS, OpenNI as the library, and NITE as the middleware, and we also used Visual Studio 2010, C++ language, for software development. The results are as follows: First, the skeleton information of a farmer could be extracted from various angles using the Kinect sensor. Next, an algorithm to calculate “finger pointing” points from the information of the joint coordinates of the shoulders, hands, and feet could be built. According to the verification experiment, the accuracy was high when compared to the assumed robot size and working area, and the control of a robot by hand pointing became possible. The estimation errors vary depending on the sensing angle of the robot toward the farmer, and the errors of sensing from behind the farmer were greater than those from other angles. It was also found that the Kinect sensor can be used even in the field in early morning and after late afternoon when the light intensity decreased and under artificial lighting.

Keywords: *Kansei*, Kinect sensor, motion control, computer vision, Japan

1. INTRODUCTION

Japanese farmers are aging. According to the Ministry of Agriculture, Forestry and Fisheries of Japan, the population of agricultural workers in the fiscal year 2010 is about 2.6 million and more than 60% of the workers are elderly persons at the age of 65 or older (World Census of Agriculture And Forestry In Japan 2010). In the next decade, 70% of agricultural workers will give up farming. Therefore, it is necessary to take action to compensate for the labor force to be lost. Additionally, when considering elderly people and new agriculture workers, agricultural machines and robots of the cooperative working type, which perform tasks together with humans, are required in particular. Although there are various challenges related to the agricultural robots (Kondo et al. 2004), in the case of the cooperative working type, compatibility with humans is very important, and the functions that can be controlled intuitively even by elderly people and new agriculture workers are desired. The function is *Kansei* communication, and we proposed agricultural robots, such as the *Kansei* Agri-robot, which is equipped with the function, and the *Chinou* robot, which extracts tacit knowledge. We have been studying and developing them. In this study, *Kansei* communication is defined as follows (Sasaki 2011): "In the communication between humans and machines/robots, they communicate mutually their intentions, emotions, feelings and health states, etc., through motions rather than the conventional one-way manner through operations/programs." As the specific communication assumed, as shown in Fig. 1, work instructions, emotions, feelings, health states, etc., are communicated from a worker to the robot through motions and facial expressions, while internal states (failure etc.), environmental information (environmental loads, health risks due to agricultural chemicals) and work proposals (harvest support, action plan, hazard notification) are communicated from the robot to a worker. It is thought this enables natural two-way communication. In this paper, an intuitive control part through motion, which is one of the core technologies for the robot, was built using the Kinect sensor, which can trace the skeleton information of a human, and the part was evaluated. The target motion was the "finger pointing" motion because it was thought that this motion is common to many persons when providing instruction to the robot. It was also checked whether the system using the Kinect sensor could be used outdoors.

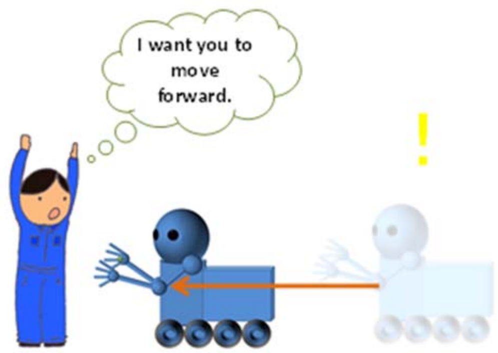

Figure 1. Image of *Kansei* communication

2. *KANSEI* COMMUNICATION FUNCTIONS AND EXPERIMENTAL EQUIPMENT & METHOD

2.1 Agricultural Tasks Utilizing Robots and Finger Pointing Control

The currently assumed *Kansei* Agri-robot and *Chinou* robot are those of compact size, which perform agricultural tasks with elderly people or women to support them and those which can be applicable for welfare agriculture. Mechanization has progressed mainly in rice cropping in Japan, but there are still many situations where humans have to perform the work. The robots are assumed to assist humans during the task between the mechanized tasks, which were performed by humans in the past or to perform advanced information sensing. The main tasks are as follows.
(1) Carrying task of fertilizers, farming implements, crops, etc.
(2) Sensing for precision farming
(3) Decision of optimal harvest timing and task support of harvest etc.
(4) Sensing and diagnosis of the fatigue degree of a worker
As for the carrying tasks, the robots carry the fertilizers and farming implements required for the task or the crops. The robots assist in the task between the mechanized tasks in particular. Then, they perform the sensing function for precision farming. The information such as weather, temperature, humidity, soil, plant, disease (Noé et al. 2011) and harmful insects are sensed in cooperation with the sensing information from field servers and satellites. For harvest support, it is assumed that the robots teach vegetables or fruits at their optimal harvest time and support the harvest task in high places, which is a burden to farmers. Finally, they perform the sensing of a farmer him/herself. In the future, sensing the relationship between the workload and degree of fatigue from the face and/or facial expression information and posture information of a farmer to diagnose the farmer will also be studied. In this paper, we studied the work instructions for robots through motions among the challenges to build the robots. Specifically, we focused on “instructing a work area” and “instructing a destination position” as highly common tasks for the tasks assumed above. Therefore, we decided to study “finger pointing” control that can be used in these tasks (Fig. 2). In our previous paper (Sasaki 2011), we studied simple robot control of forward move, backward move and stop through the motions, holding up both hands, holding up one hand and upright posture. As a result, we found that the gesture to move the robot forward varies from person to person. Compared to the previous study, it is thought that the difference among individuals will be small in the control using “finger pointing.”

2.2 Kinect Sensor and Development System

The Kinect sensor is a gaming device for the Xbox 360 and was released by Microsoft Corporation in 2010. It consists of an infrared light as the depth sensor, image sensor, and multi-array microphone. The infrared light as the depth sensor performs laser radiation of near infrared (830 nm) widely and the image sensor is an RGB camera of

32-bit color VGA (640 x 480) at 30 fps. The depth sensor consists of a camera that takes a near-infrared pattern of the irradiated laser and has view angles of 57 degrees (horizontal) and 43 degrees (vertical) and a detection range of 1.2 – 3.5 m. The Kinect sensor can perform real-time skeleton tracing processing by estimating the joint locations from the information obtained by the near-infrared sensor. Since the skeleton information can be obtained without placing special markings on an object person, it is possible to analyze and accumulate agricultural tasks under natural conditions, and it is also possible to record agricultural tasks automatically. In this study, we used OpenNI (Natural Interaction) as the development environment for the Kinect sensor. OpenNI is an open source library mainly developed by PrimeSense, Ltd., which had developed the sensors on the Kinect sensor and consists of a sensor module that controls the hardware, such as the RGB camera, near-infrared camera and microphone on the Kinect sensor, and middleware that detects the postures, movements and gestures of a person using the input data from these hardware devices (Fig. 3). In this study, we used Windows 7 as the OS, OpenNI as the library, and NITE as the middleware, and we also used Visual Studio 2010 for the software development. Note that, in the three-dimensional coordinates of the Kinect sensor, the x-axis is right to left, the y-axis is bottom to top and the z-axis is back to front when viewed form the sensor.

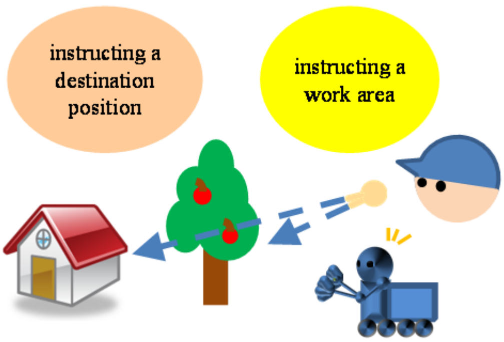

Figure 2. Control by pointing

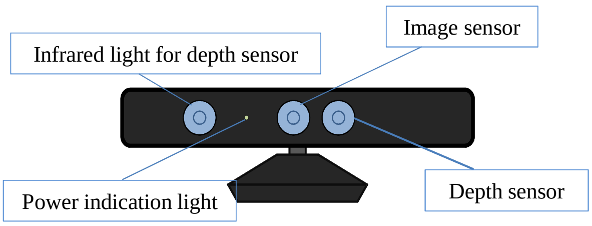

Figure 3. Appearance and configuration of Kinect

3. BUILDING AND EVALUATION OF MOTION CONTROL PART

3.1 Calculation of Finger Pointing Coordinates and Experimental Method

In this study, the coordinates of the dominant hand are expressed as $\overrightarrow{Hand}$, those of the shoulder of the dominant hand as $\overrightarrow{Shoulder}$, those of the foot on the dominant hand side as $\overrightarrow{Foot}$ and those of the calculated finger pointing position as $\vec{Q}$. The relationship of the coordinates in the upright posture is shown in Fig. 4. The (x, y, z) components are shown by Eq. (1) – (3).

$$\overrightarrow{Hand} = (Hand_x, Hand_y, Hand_z) \quad (1)$$

$$\overrightarrow{Shoulder} = (Shoulder_x, Shoulder_y, Shoulder_z) \quad (2)$$

$$\overrightarrow{Foot} = (Foot_x, Foot_y, Foot_z) \quad (3)$$

Where the coordinates $\vec{R}$ in Fig. 4 are shown by the following equation.

$$\vec{R} = (Shoulder_x, Hand_y, Shoulder_z) \quad (4)$$

On the XY- and ZY-axes, according to the similarity relationship between the triangle with its vertex at ($\overrightarrow{Shoulder}$, $\overrightarrow{Hand}$,$\vec{R}$) and that with its vertex at ($\overrightarrow{Shoulder}$, $\vec{Q}$,$\overrightarrow{\text{Foot}}$), the coordinates of $\vec{Q}$ are finally obtained by Eq. (5).

$$\vec{Q} = (\frac{(Hand_x - Shoulder_x) * (Foot_y - Shoulder_y)}{(Shoulder_y - Hand_y)} + Shoulder_x, Foot_y, \frac{(Hand_z - Shoulder_z) * (Foot_y - Shoulder_y)}{(Shoulder_y - Hand_y)} + Shoulder_z) \quad (5)$$

To verify the calculated finger pointing coordinates $\vec{Q}$, the following two experiments were conducted on December 7, 2012.

(1) Evaluation of the accuracy of the calculated coordinates $\vec{Q}$

(2) Change of the accuracy of $\vec{Q}$ due to the change of angle

The background color for the imaging was standardized as white, and it was conducted in an environment where as few persons as possible appear in the images. The Kinect sensor was placed at a height of 45 cm from the ground. The posture of the farmer was an upright posture. For the evaluation of the accuracy of the calculated $\vec{Q}$, ten target points were set and their coordinates were measured using the Kinect sensor. And then the target and measured coordinates were compared to calculate the average and standard deviation of the error. For the evaluation of the accuracy due to the change of angle, the target point and the finger pointing posture of the farmer were fixed. and the sensing position was changed by 45 degrees at a time. Fig. 5 shows the algorithm to determine the coordinates of the finger pointing position $\vec{Q}$.

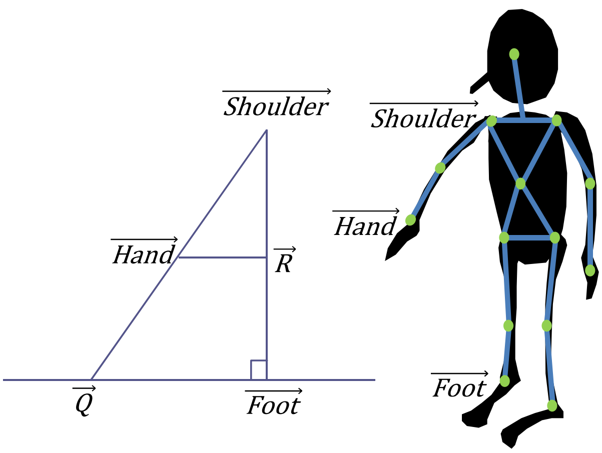

Figure 4. Skeleton rendition and a theoretical figure

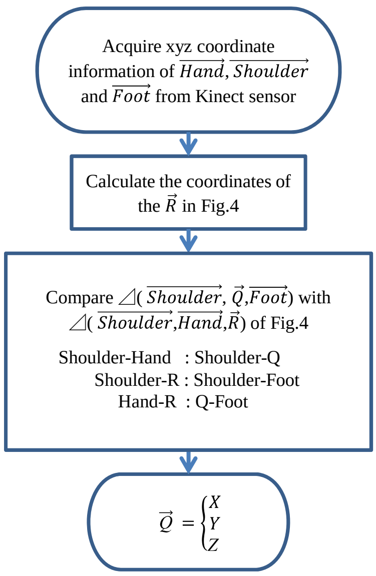

Figure 5. Calculation of Q

3.2 Result and Discussion of Finger Pointing Coordinate Calculation and Evaluation

First of all, the skeleton information could be extracted from various angles using the Kinect sensor (Fig. 6). For the accuracy of the calculated finger pointing points, the results of the average errors and standard deviations of the (x, y, z) axes are shown in Fig. 7. The errors are (18.2 ± 11.2 cm, 5.3 ± 3.1 cm, 22.9 ± 13.3 cm), and it is thought that the accuracy was high when compared to the assumed robot size (about 2 m x 2 m) and work space. Next, for the change of the estimated accuracy due to the change of angle, the comparison of the (x, y, z) components is shown in Fig. 8. The angle of 0 degrees corresponds to the front of the subject and the angle of 180 degrees corresponds to the back. The error of the z-component at 180 degrees is 123.9 cm, and this value is substantially higher than the others, so it is not shown in Fig. 8. As expected, the skeleton information could not be obtained correctly from the sensing from the back; therefore, the errors were greater than those of the measurements at other angles. The (x, y, z) components of the average error and standard deviation of all the angles are (7.3 ± 6.8 cm, 8.8 ± 5.7 cm, 31.1 ± 36.8 cm). According to the results and observation of the verification experiment, it was found that the accuracy of the finger pointing point to be calculated was high when all the skeleton coordinates were detected by the Kinect sensor and/or the sensing was performed at the angle near the front of the subject's body. Although it is assumed that the robot performs the sensing of the worker in various situations and work environments, it is thought that the robot must move or act in consideration of the accuracy of finger pointing points while understanding the sensing angle against the farmer.

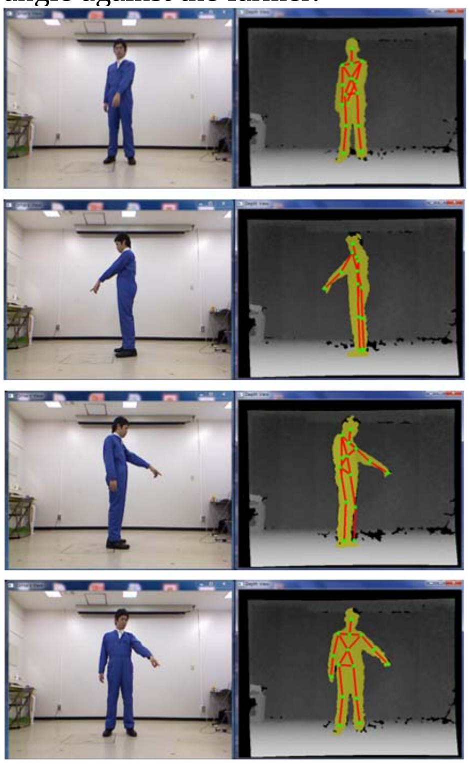

Figure 6. Example of detection

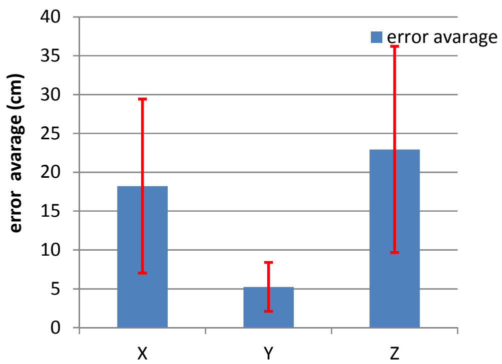

Figure 7. Error of detection

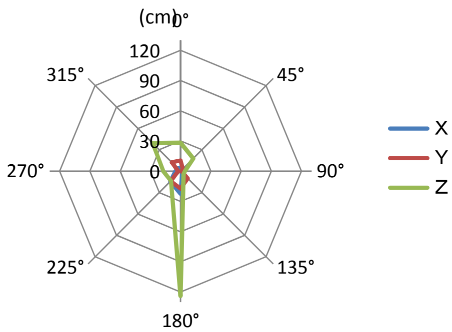

Figure 8. Error by angle

C0075
Y.Sasaki, S.Shibusawa, H.Negishi and M.Emmi. "Development of Motion Control Using the Kinect Sensor for the *Kansei* Communication Interface". EFITA-WCCA-CIGR Conference "Sustainable Agriculture through ICT Innovation", Turin, Italy, 24-27 June 2013.

3.3 Skeleton Information Extraction in the Field and Discussion of Its Application

The basic performance of the Kinect sensor was measured and investigated in the field between 5:40 a.m. and 5:40 p.m. on October 24, 2012 (weather: fine). The Kinect sensor was placed at a height of 45 cm form the ground and at a distance of 180 cm form the subject. The illuminance was measured along with the extraction results of the Kinect sensor. The measurement position of a light meter is at the center of the subject (and a height of 150 cm) and the light meter was directed to the Kinect sensor. The light meter used was CEM DT-1309. During the time of measurement, the illuminance changed between 5 and 22,000 lux. They finger pointing posture was measured between 5:40 a.m. and 7:20 a.m. at every 20 minutes. Additionally, it was measured at 0:40 p.m., 4:30 p.m., 5:00 p.m., and 5:40 p.m. (after sunset). For the measurement at 5:40 a.m., the angle was changed by 45 degrees at a time to investigate if the skeleton tracing is possible for each posture. It was confirmed that the skeleton information could be obtained from all directions (360 degrees) when the posture was detected correctly. Under this measurement condition, the skeleton coordinates could be detected up to the illuminance of 9,600 lux (at 7:10 a.m.), but the skeleton information could not be detected in the next measurement because the illuminance was 18,000 lux (at 7:20 a.m.). The skeleton information could not be detected during the day when the illuminance is high. At 4:30 p.m., the illuminance dropped to 2,400 lux and the skeleton information could be detected again. After sunset, even when the subject could not be seen with the RGB camera, the skeleton information could be detected. Based on the results shown above, although it may depend on optical disturbances due to weather, season and measurement conditions, it is thought that the sensing in the field is possible using the Kinect sensor when agricultural tasks are performed in the early morning, after late afternoon, at night, and under artificial lighting. The actual agricultural tasks, which are carried out in the early morning or after late afternoon, include the following cases.

- In open-field culture, sprinkling is carried out during cool time in the early morning or after late afternoon.
- Harvesting is carried out during the time when the temperature is low to prevent quality loss due to high temperature as much as possible.
- Sowing and planting are carried out during cool time in the early morning or after late afternoon as an effort to do the tasks at an optimal timing.
- Both in facility and open-field culture, sprinkling is carried out amply in the early morning or after late afternoon to prevent growth disturbance such as bloom delay due to high temperature and/or dryness.
- Chemical spraying is carried out in the early morning or after late afternoon when the temperature is low, because chemical antagonism might occur if it is done during the time when the temperature is high.
- Harvesting is carried out in the early morning when the temperature is low to prevent quality loss such as wilting due to high temperature as much as possible.

4. CONCLUSIONS

In this study, for the *Kansei* communication function used for the *Kansei* Agri-robot and Chinou robot, the control part through finger pointing motion was built and evaluated in particular. First of all, the skeleton information of a farmer could be extracted from sensing at various angles using the Kinect sensor. Next, an algorithm to calculate a finger pointing point from the joint coordinate information of shoulder, hand and foot could be built. The verification experiment of the algorithm was conducted and it was found that the accuracy was high when compared to the assumed robot size and work range, and the control of a robot through finger pointing motion became possible. The estimation errors varied depending on the angle of the robot against the farmer to be sensed, and the errors of sensing from the back of the farmer were higher than those from other angles. It was also found from the investigation that the Kinect sensor could be used even in the field in the early morning and after late afternoon when the light intensity decreased and under artificial lighting. The future challenges include improvement of the Kinect sensor to enable sensing through an entire day and building of the other motion control parts.

5. **ACKNOWLEDGMENT**

Part of this work was supported by Grant-in-Aid for scientific research (KAKENHI A, 23248043)

6. REFERENCES

2010 World Census of Agriculture And Forestry In Japan. 2010. Ministry of Agriculture, Forestry and Fisheries, the agricultural population in Japan

Kondo,N., M.Monta and N.Noguchi. 2004. Agri-Rbot (I) ‑Fundamentals and Theory, CORONA PUBLISHING CO., LTD.

Sasaki, Y. . 2011. Detection of the Intention from Gestures of Workers for a *Kansei* Agri-Robot, Agricultural Information research, Vol.20 (1):13-18.

Velázquez-López,N, Y.Sasaki, K.Nakano, J.M.Mejía-Muñoz, E.Romanchik. Kriuchkova. 2011. DETECTION OF POWDERY MILDEW DISEASE ON ROSE USING IMAGE PROCESSING WITH OPEN CV, Revista Chapingo Serie Horticultura,   Vol.17 (2):151-160.