Agricultural Information Research 20(1), 2011. 13–18
Available online at www.jstage.jst.go.jp/

Original Paper

Detection of the Intention from Gestures of Workers for a *Kansei* Agri-Robot

Yutaka Sasaki*

Faculty of Regional Environment Science, Tokyo University of Agriculture, 1-1-1, Sakuragaoka, Setagaya, Tokyo 156-8502, Japan

Abstract

In this study, a next-generation *Kansei*-equipped agricultural robot is proposed and focuses on controlling the robot through body language. *Kansei* agri-robots are defined as agricultural robots equipped with *Kansei* communication and robotic functions. *Kansei* communication is defined as "two-way communication between humans and robots, or other machines, that supplements or supersedes traditional one-way (human-to-machine) input and operation". In *Kansei* communication, humans and machines interact in ways that take into account thoughts, emotions, etc. Thus, *Kansei* robotics can be defined as "robotics technology in which *Kansei* communication has been made possible". In this paper, the first step towards *Kansei* communication was the extraction of the outline of a human worker from the overall image observed by the robot. This was made possible in a closed environment using an algorithm combining the HSV color extraction and background finite difference methods. It was then investigated methods of controlling our *Kansei* agri-robot using body language and gestures extracted from still photos of agricultural workers. For this research, a LEGO Mindstorms robot equipped with computer vision was used. Commands instructing the robot to stop, advance, reverse, etc. became possible based on the ratio of the extracted worker area to the convex hull area.

Keywords

Kansei, Agri-robot, *Kansei* robotics, *Kansei* engineering, computer vision

Introduction

In order to perform efficiently in agricultural environments that include a wide variety of farm products, plants and animals, agricultural machinery and robots must be designed to take into consideration unique rules.

This may not always apply if the robot or machine is designed to operate in enclosed controlled areas, such as greenhouses and/or agricultural processing plants where artificial lighting is used and other environmental factors are controlled. However, robot operations in outdoor farm environments can be complicated because, in most cases, they involve conditions of fluctuating light, seasonal variations, and local weather conditions that can change without warning.

Based on the premise that a thorough understanding the complexity of such objects and environments is vital, advances in mechanization and other areas have been made that aim at assisting in the cultivation and harvesting of rice, vegetables, fruit and other agricultural products. As pointed out by Yukumoto (2009), the current fundamental key issues facing agricultural machinery are considered to be mechanization in currently non-mechanized sectors (such as vegetable and fruit cultivation/harvesting), targeted work in complex environments, and advanced computerization to enable increased productivity and quality of rice harvests, etc. Additionally, improvements to labor environments, enhancements to safety and comfort as well as device miniaturization based on the need to operate in restricted inter-mountain areas, are all necessary to support mechanized system tasks. Similarly, with regard to agricultural robots, advances are being pursued in areas that will facilitate further inclusion and support of agricultural machinery, as well as work in more complex environments and conditions, etc.

Aside from agricultural machinery, the entire robotics field can be expected to become the object of increased future research. Areas of interest include development of improved sensors for obtaining information about the environment and objects, the action of the brain in regard to information processing, recognition

* Corresponding Author
E-mail: y3sasaki@nodai.ac.jp

and decision making, development of robotic hands and feet to facilitate mobility and manipulation, and the development of communication functions (interface) with humans and other robots (Kondo et al. 2004). Furthermore, in Japan, the field of agriculture has other specific concerns, such as an increasingly aged workforce and the relative scarcity of people willing to become the next farming generation, which makes it necessary to consider work participation by women and the elderly.

In addition to the foregoing topics, it is necessary to examine the blueprint for future. It is begun this study by focusing on relations between human beings and agricultural machinery/robots, which can be roughly classified into three large categories: 1) “Beagle” automation, such as unmanned tractors and the autonomous robotic work performed in unmanned plants and factories, 2) Robotic accouterments such as Hybrid Assistive Limb (HAL), which is an artificially powered exoskeleton robot suit developed by Professor Yoshiyuki Sankai at Tsukuba University (CYBERDYNE: <http://www.cyberdyne.jp/>, browsed on May 12, 2009), along with other work accouterments as discussed by Toyama (2009), 3) The harmonization of working styles and environments where human beings and robots interact.

Numerous studies are ongoing into autonomous robot work routines and worker accouterments, but it is difficult to judge if sufficient study and research is being conducted into harmonized working styles between humans and robots, which are considered critical for future agricultural work. This study focuses on *Kansei* as an important keyword for production in the next generation.

Various enterprises and industrial efforts incorporating *Kansei* have achieved increased social recognition, and the Ministry of Economy, Trade and Industry (METI) of Japan is now proposing “*Kansei* Value” as the fourth axis of value, surpassing the other three elements of “High Performance”, “Reliability”, “Lower Price” as targets for Japanese-made products, (METI: Kansei value creation initiative, <http://www.meti.go.jp/press/20070522001/kansei-gaiyou.pdf>, browsed on May 12, 2009).

One such field is *Kansei* engineering, which is a discipline dedicated to designing products and environments that incorporate *Kansei* values, and images into physical design elements.

With these goals in mind, two tasks specifically related to the proposed agri-robot were examined as part of *Kansei* communications: the extraction of worker gesture and body language information using computer vision, and operator control of robotic actions via body language.

Proposal of *Kansei* agri-robot

In this study, it is designed and proposed a *Kansei* agri-robot, while focusing on the communication functions and interface between human beings, machines, and robots in terms of affinity, cooperative task performance, mutual work support and *Kansei* values. A *Kansei* agri-robot is defined as an agricultural robot possessing *Kansei* communication functions and utilizing *Kansei* values in relations to agricultural machinery. To date, machinery and robot research efforts into communication between humans and robots (or other machinery) have been limited to information transfer and verbal communications. In many cases, this implied operations and programming, etc., that were only one-way from the human perspective. In contrast, Kato (2005) pointed out that *Kansei* is necessary for communication, and Hashimoto (2005) argued that *Kansei* communication is necessary for both humans and robots to interact smoothly.

This paper defines *Kansei* communication as communication interactions that embrace intentions, feelings, and health states among humans and robots, as well as other machines, instead of typical one-way actions such as human machine operation and programming. Also, *Kansei* Robotics serves to grasp humans’ feelings, the robots’ creations and expressions of its own *Kansei* being researched by Gotoh et al. (2006). As part of these efforts, Tokumaru et al. (2009) have been engaged in developing *Kansei* robots capable of color recognition that will enable them to perform as partners by sharing perceptions with human operators. Based on this vision, *Kansei* robotics is defined as “Robotics technology capable of *Kansei* communication.”

The characteristics and advantages of a *Kansei* agri-robot are explained below. The basic work objectives and environment are the same as those experienced by workers operating conventional agricultural machinery. Therefore, an agriculture robot designed with *Kansei* communication would be able to interact smoothly with, and display affinity for, its operators, thus facilitating the performance of cooperative tasks. Practical *Kansei* communication, as shown in Fig. 1, assumes that work instructions provided by means of gestures and expressions can convey information on feelings and health states from people to the *Kansei* agri-robot, while internal states (malfunctions, etc.) environmental information (health risks due to environmental loads or agricultural chemicals) work suggestions (harvest support, action plan/hazards) can be conveyed from the *Kansei* agri-robot to its operator.

As a result, it is thought that this will enable more natural two-way communication. Furthermore, the dual problems of an aging workforce and the scarcity of next generation farmers are issues of concern in contemporary Japanese agriculture. Thus, if a *Kansei* agri-robot can be fabricated that would receive its operational instructions in the form of gestures and expressions from its human operators, it would enable anyone (including elderly persons and women) to farm more intuitively.

By understanding and reacting to sensory information displays indicating emotions such as displeasure, fear and/or surprise, such robots could lessen unpleasantness in labor environments and reduce work risks by performing actions such as modifying engine work speeds, or halting machinery operations, when

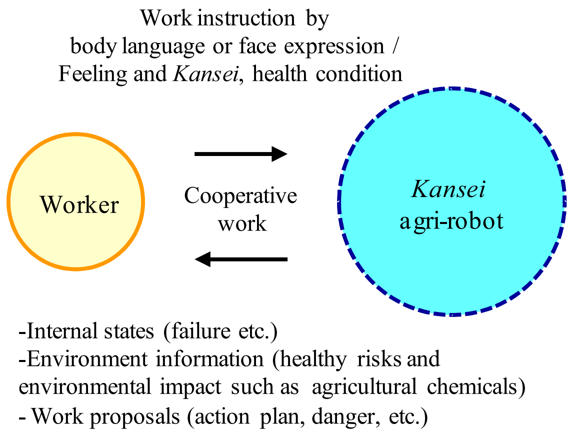


Fig. 1 Communication between worker and *Kansei* agri-robot

appropriate. Furthermore, the ability to monitor the health state of workers from their expressions and actions could enable intuitive control of work speeds and thus a reduction of work risks. In the other direction, suggestions from the *Kansei* agri-robot regarding optimal fruit harvest times, reports regarding problems with its own internal conditions, advisories on health risks to workers from chemicals and environmental loads, and other information necessary for efficient cooperative tasks and hazards, could reduce accidents and risks and thus enable efficient, environmentally conscious farm work.

The agri-robots created under the current study were designed to be small in order to support farm work by elderly persons and women, and may be applied to agricultural welfare and etc. It is expected that practical agricultural work will require robots capable of providing assistance with farm product harvesting, support removal of diseased plants and weeds, provide assistance with mechanized agricultural systems (the transport of crops and tools), among other chores.

In the future, it is also expected that efforts will continue towards development of agri-robot functions focusing on the workers and their actions, with special emphasis on cognitive extraction of exemplary farmer and management support behavior, and through automatic recording of actual work performance.

Experimental Method

Computer vision system and experimental environment

When designing and building *Kansei* agri-robots, recognition of workers is considered indispensable. As proposed in this paper, to ensure humans and robots work harmoniously, the robot first recognizes worker intentions from body language communications. In the future, it is believed that *Kansei* information will be extracted from facial expressions. Because outdoor environmental conditions, such as variations in light and wind, imposed significant difficulties, it is first attempted to extract worker outlines in an indoor environment using a computer vision system.

Computer vision is defined as visual recognition of information by computers/machines/robots. In this study, Open CV (Open Source Computer Vision Library) was used (Intel: Open Computer Vision Library, <http://www.intel.com/technology/computing/opencv/index.htm>, browsed on July 29, 2008).

Open CV is an open source C language image processing library. Its advantage is its multi-platform operability, which means that means it is operable on any OS free of charge, and can be used to design and build systems at minimum cost.

A PC camera, QCAM-200R (Logicool) with autofocus functions and a pixel count of 2,000,000, was connected to a laptop PC (DELL 1500 Vostro, w/Intel(R) Celeron(R) CPU @ 1.86 GHz, 2 GB RAM) via a USB port, and still images of workers were taken and studied. The Windows XP® operating system, and Visual Studio 2005, both produced by Microsoft Corporation®, were used for the development environment.

Image processing for worker gesture recognition was performed according to the algorithm shown in Fig. 2, which is a combination of HSV-color extraction and the background subtraction method.

First, the HSV-color extraction method is explained. The images were acquired from the PC camera and converted from RGB space color into HSV space color. HSV refers to Hue (coloring), Saturation (vividness), Value (brightness). The conversion formula from RGB is defined with expression (1) below:

$$H = \begin{cases} 60 \cdot \dfrac{G-B}{\max(R,G,B)-\min(R,G,B)} + 0 & \text{where } \max(R,G,B) = R \\ 60 \cdot \dfrac{B-R}{\max(R,G,B)-\min(R,G,B)} + 120 & \text{where } \max(R,G,B) = G \\ 60 \cdot \dfrac{R-G}{\max(R,G,B)-\min(R,G,B)} + 240 & \text{where } \max(R,G,B) = R \end{cases}$$

$$S = \frac{\max(R,G,B) - \min(R,G,B)}{\max(R,G,B)} \qquad (1)$$

$$V = \max(R,G,B)$$

The HSV-color extraction method samples the colors of a worker's skin, hair, and clothing, and using basic image processing, combines the three images. Afterwards, closing is applied to the lost areas. Closing is a technique in which an erosion process is applied to the subject images after a dilation process. It is also used during noise removal.

The background difference method analyses the background information, and then determines whether each pixel of the current image is background or foreground. Because of the simple differentials involved, it is vulnerable to variations in light such as flashes and changes in sunshine. Furthermore, this method often recognizes background items moving rotationally (such as plant

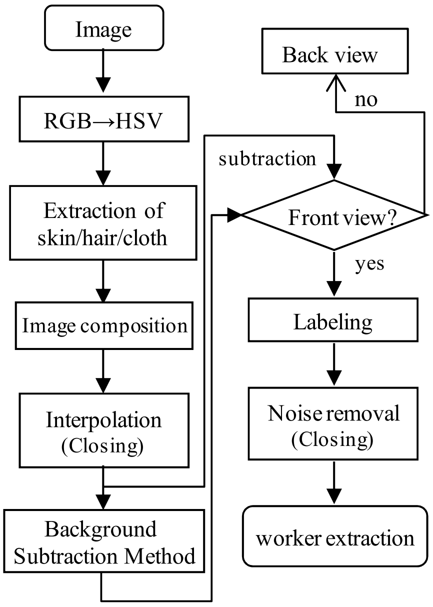


Fig. 2 Worker extraction algorithm

movements due to wind agitation) as foreground, and foreground objects with colors prevalent in the background could be regarded as background. Therefore, background distortion per pixel was modeled using normal distributions, after which the background differential method was applied to recognize the foreground. The normal distribution is computed using Eq. (2), where μ and σ represent the average and the dispersion per Pixel x out of the quantity N pixels.

$$f(x)=\frac{1}{\sqrt{2\pi\sigma}}\exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} \tag{2}$$

In this study, 100 images were first acquired using a 0.07 sec sampling cycle, after which a single Gaussian model was built for each image, and made up to twice the standard deviation as the background.

Additionally, it is possible to compute an image center of gravity and pinpoint the location of workers via labeling. Labeling is a process that attaches the same label (number) to adjacent pixels having the same value. This makes it possible to recognize sampled items.

The algorithm in Fig. 2 implements the RGB to HSV conversion for the extraction of skin, hair, working accouterments and any combination of the three, as well as performing interpolation on lost areas. Then, after the HSV-color extraction process and taking differentials with the input images, the algorithm implements labeling, noise removal (closing) and performs extraction of worker outlines.

As shown in Fig. 3, an experimental environment was created indoors using a green curtain, which was chosen to simulate plant color. A VITALITE fluorescent lamp, capable of high color rendering at 5,500 K, was chosen to provide illumination. This is

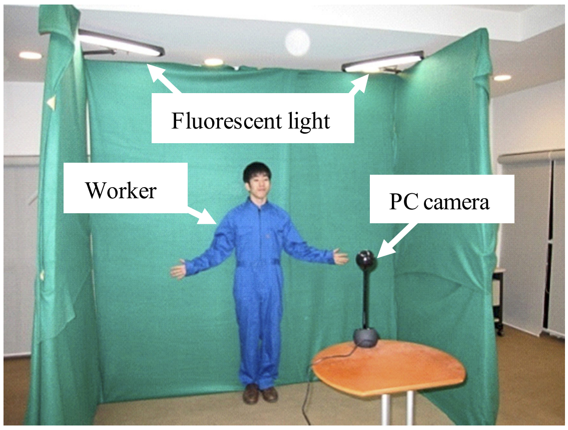


Fig. 3 Indoor experimental device

Table 1 Value of HSV (Min-Max)

Min-Max	H	S	V
Skin	0–20	25.5–255	25.5–255
Hair	0–360	25.5–102	25.5–102
Work cloth	90–170	25.5–255	25.5–255

a full spectrum lamp designed to approximate natural light (sunshine).

Table 1 shows the HSV values set for the indoor experiments. For the photo environment, a DT-1309 digital illuminometer was used, and measurements were taken within an optional 30 seconds. The illumination value was 250.0±4.9 lx and the image size acquired was 320×240 pixels. Worker outline data was extracted as shown in Fig. 4.

Examination method of *Kansei* communication by body language

Our future plans include efforts to make *Kansei* communication possible with a higher level of body communication, and to build the specific agri-robots using *Kansei* information obtained from facial expressions as their operational controls.

A caterpillar-tread robot equipped with a camera was fabricated using a LEGO "Mindstorms NXT" robot kit, as shown in Fig. 5, and a part of *Kansei* communication was determined using the instruments shown in Fig. 3. The LEGO Mindstorms NXT was selected in order to provide actual proof of *Kansei* communication. The workers' information was recognized by the camera and processed with Open CV. The control systems of the NXT Robot were built using Visual Studio 2005.

Based on the assumption that the subject worker was facing towards the robot, the robot was programmed to perform the following actions:

1) When the worker is standing still, remain in one place.
2) When the worker raises one hand, move forward.
3) When the worker raises both hands, move backwards.

Original image (left)　　Worker extraction image (right)

Fig. 4　Recognition of worker

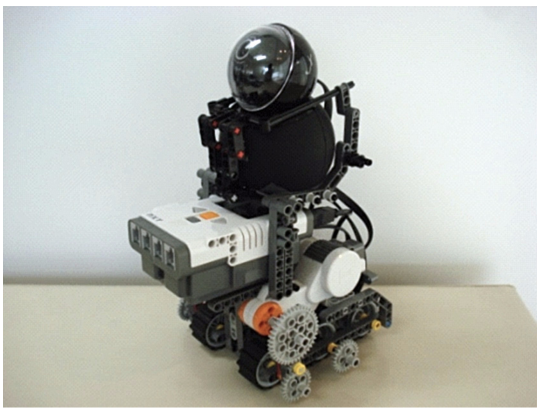
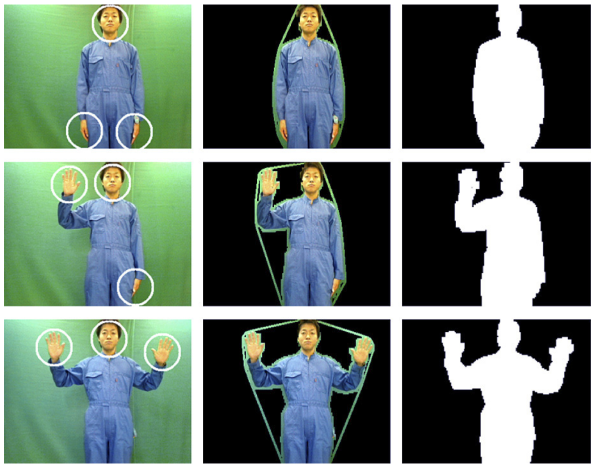


Fig. 5　LEGO mindstorms NXT Robot



(a) Recognition of the head and both hands　　(b) Convex hull image　　(c) Worker extraction image

Fig. 6　Convex hull and worker extraction image

For the control information, the ratio of the workers' extraction area out of convex hull area was applied. The convex hull is the minimum convex envelope that encloses a figure and is often used in image processing and pattern recognition. It can also be applied to moving images.

During our investigation, the distance from the worker to the camera was varied to be 1.0, 1.5, 2.0, or 2.5 m. Measurements were taken continuously for 30 seconds, after which the average and standard deviation were calculated. Using the worker gesture information extraction algorithm, HSV-color extraction was used with the value as: H: 0–20, S: 51–255, V: 25.5–255, and made closing three times.

Result and Discussion

The ratio of the worker's outline area to the convex hull area is shown in Table 2 and is based on the assumption that the worker is looking in the direction of the robot. The values recorded when the worker was standing straight, standing with one hand raised, and standing with both hands raised, were clearly different. The results for recognition of the head and hands, the convex hull and the worker outline are shown in Fig. 6. First, the ability to properly recognize the worker's head and hands was confirmed by the white circle in Fig. 6(a). The line circling the image of the worker is the convex hull (Fig. 6(b)). This area ratio to the worker's image sampled in (c) confirmed that *Kansei* communication is possible using simple body language. As shown in Fig. 7, when the most suitable threshold in the Hold-out method (H-method) was determined, it allowed us to distinguish between a person standing straight and then standing straight with one hand raised with 0.894 ratio, and standing straight with one hand and both hands up, with 0.726 ratio. The developed algorithm successfully enabled robots to stop, go forward, and go backward smoothly.

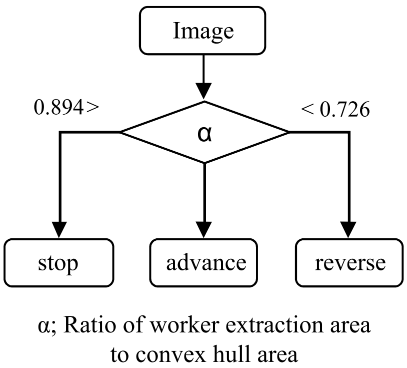


Fig. 7 Robot control algorithm by ratio of worker extraction area to convex hull area

Table 2 Ratio of worker extraction area to convex hull area

Distance to worker	Operation	Ratio of worker extraction area to convex hull area
1.0 m	Standing straight	0.939±0.006
	One hand	0.793±0.003
	Both hands	0.674±0.003
1.5 m	Standing straight	0.909±0.004
	One hand	0.776±0.006
	Both hands	0.649±0.007
2.0 m	Standing straight	0.939±0.004
	One hand	0.791±0.006
	Both hands	0.678±0.010
2.5 m	Standing straight	0.942±0.004
	One hand	0.855±0.004
	Both hands	0.700±0.006
Average	Standing straight	0.933±0.014
	One hand	0.805±0.031
	Both hands	0.677±0.020

Using these commands, a robot can be directed to the proper place for cooperative work by means of simple gestures.

In the future, it is believed that using information pertaining to the position (gravity) of hands and head (Fig. 6 (a)), further development of the algorithm could possibly be used to control the robot using more complex body language.

This time, in the interior of a room, it is thought that control by a *Kansei* agri-robot's body language was completed. In the outdoors, it is thought that extraction of the worker from the disturbance of light or the complexity of a background is difficult, and it becomes important how a worker can be extracted correctly. This considers a future big subject.

Conclusions and Further Plans

In this paper, a *Kansei* agri-robot was defined and proposed, and the following two items were performed:

1. A worker outline extraction algorithm was developed and its effectiveness was shown.
2. Using the ratio of worker's outline area to that of the convex hull, the ability of *Kansei* communication to stop, advance, and reverse a LEGO "Mindstorms" NXT Robot was demonstrated. Further robot control possibilities using the information pertaining to worker heads and hands were also discovered.

Acknowledgment

A portion of this work was supported by a Grant-in-Aid for Scientific Research (KAKENHI, 21780239).

References

Gotoh, M., M. Katoh, S. Kato, T. Kunitachi and H. Itoh (2006) Face Generation Using Emotion Regions for *Kansei* Robot, The Japanese Society for Artificial Intelligence, 21(1): 55–62.

Hashimoto, M. (2005) KANSEI Communication for Human-Robot Symbiosis, The 1st time Transdisciplinary Federation of Science and Technology Conference, 79–80.

Kato, T. (2005) Computational Modeling of KANSEI for Seamless Human Communication, The 1st time Transdisciplinary Federation of Science and Technology Conference, 61–66.

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

Tokumaru, M., Y. Nagata and N. Muranaka (2006) Development of the Kansei Robot with Color Sence, 22nd Fuzzy System Symposium, 603–608.

Toyama, S. (2009) AFP BB News, <http://www.afpbb.com/article/environment-science-it/science-technology/2556268/3667115>, browsed on July 15, 2009.

Yukumoto, O. (2009) The role which development, improvement, and research of an agricultural machine played in progress of farm mechanization, Agricultural machine Public relations Vol. 1, <http://www.agriworld.or.jp/nitinoki/column/column_43/column4303/index.html>, browsed on May 12, 2009.

Received July 26, 2010
Accepted December 6, 2010
Agro-informatics & Technology