Proceedings of the 2020 IEEE/SICE
International Symposium on System Integration
Honolulu, Hawaii, USA, January 12-15, 2020

A System for Operating Home Appliances with Hand Positioning in a User-definable Command Space

Shixun Yan, Yonghoon Ji, and Kazunori Umeda

***Abstract*—In this paper, a novel system that operates home appliances at arbitrary positions in a room is proposed based on a command space associated with the operation of the home appliance. In the proposed system, a hand positioning gesture is used to operate home appliances. By installing a user-definable command space, it is possible to operate home appliances freely in an arbitrary place chosen by the user. Experiments were conducted to verify the operational accuracy and operation time for commands using the proposed system for detailed operations such as TV channel switching.**

I. Introduction

Home appliances indispensable for everyday life are becoming more multifunctional and high performing, which leads to complicated user operation. In recent years, there have been many intensive studies on the intuitive manipulation of products using human gestures that are familiar to users. For example, various methods for realizing the gesture interface have been proposed [1-4]. However, since these systems deal with too many gestures, it seems that they are still difficult in practical use. We also constructed intelligent rooms, based on a camera network, that can operate home appliances using simple hand gestures (e.g., hand waving) [5-7], as shown in Fig. 1. In order to solve the problem of elderly people not being good at operating small buttons, such as on a remote control, in a previous study, we proposed a system that can operate home appliances more intuitively. By using a pseudo-relative coordinate system, home appliances could be operated from any position [8]. However, there is a problem, in that it takes a long time to detect hand waving using an image-processing scheme. In addition, users may use a lot of physical strength because they have to wave their hands many times when operating home appliances, which can be especially difficult for elderly people. In this regard, there is a method of operating home appliances using eye-gaze information. However, it is difficult to detect eye gaze stably at any position in the room, and the time required for each operation has not been evaluated [9]. To overcome that limitation, in this study, we propose a novel system that can operate home appliances with a simpler hand gesture (i.e., hand positioning) with less burden on the user. Here, we define hand positioning gesture as putting a hand in one place in a short time within 5 s. The procedure for operating the proposed intelligent room system is as follows. First, the user performs a series of hand waving gestures twice at different positions to define the pseudo-relative coordinate system. Then, the command space is constructed based on the pseudo-relative coordinate system. Thereafter, the user can freely manipulate the home appliance through hand positioning in the command space.

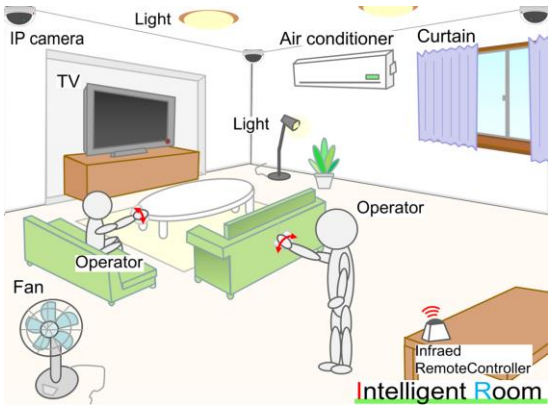

Figure 1. Conceptual diagram of an intelligent room.

Shixun Yan is with the Course of Precision Engineering, School of Science and Engineering, Chuo University, Tokyo, Japan (e-mail: karasu@sensor.mech.chuo-u.ac.jp).

Yonghoon Ji is with the Course of Precision Engineering, School of Science and Engineering, Chuo University, Tokyo, Japan (e-mail: ji@mech.chuo-u.ac.jp).

Kazunori Umeda is with the Course of Precision Engineering, School of Science and Engineering, Chuo University, Tokyo, Japan (e-mail: umeda@mech.chuo-u.ac.jp).

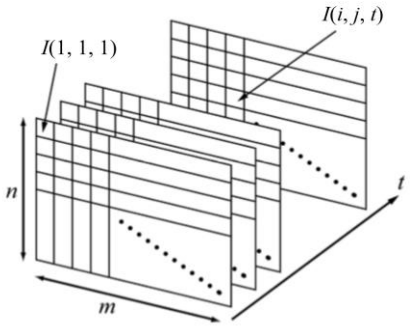

Figure 2. Temporal shading values of pixels.

The remainder of this paper is as follows. Section II presents the construction of command space based on the pseudo-relative coordinate system. Section III describes how to operate the home appliance. Experimental results are detailed in Section IV. Finally, Section V gives our conclusions and insight as to our future work.

II. Command Space Construction

A. Hand-waving detection

The hand-waving detection process is divided into the following steps. First, the resolution of each acquired grayscale image, $\boldsymbol{I}$, from the camera network is reduced to $m \times n$ pixels. Figure 2 depicts the definition of our image

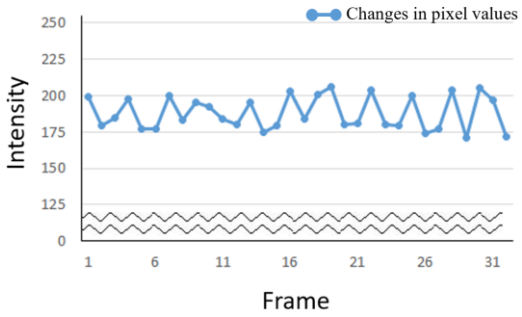

Figure 3. Changes in pixel values corresponding to hand waving.

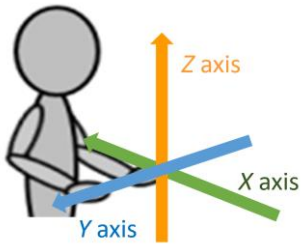

Figure 4. Pseudo-relative coordinate system based on the position of hand waving.

coordinate system. We define $\boldsymbol{I}(i, j, t)$ at time t as the grayscale value where $i = 1, 2, ..., m$ and $j = 1, 2, ..., n$ represent pixel coordinates. By reducing the image resolution, the computational burden of subsequent processing is reduced, and noise is also suppressed. In addition, the change in the grayscale pixel value of the image by the hand-waving gesture becomes smooth; thus, it is possible to obtain a change in the pixel value close to the sine wave. Fast Fourier transform (FFT) analysis of time series for all pixel values is performed to detect hand waving in the image data. This process is valid, given that the pixel values are clearly changed between the hand region and the background region, as shown in Fig. 3. Next, if the calculated frequency values from the FFT exceed a certain threshold, the pixel is detected as a hand-waving gesture region in the image. The above-mentioned process is performed for all images from multiple cameras in order to find pixels satisfying the epipolar constraints for where the hand-waving gesture is performed under an assumption that the hand-waving gesture occurs only in one place at any one time. Finally, a 3D position for where the hand-waving gesture was performed is acquired based on the principle of stereo measurement.

B. *Pseudo-relative coordinate system*

The pseudo-relative coordinate system in a 3D space is defined by two consecutive hand waving gestures at different positions by the user, and is fixed at the initial position, even if the user moves to another place. The command space is constructed based on the pseudo-relative coordinate system that the user sets by the hand-waving gesture. Thus, it is possible to operate the home appliances freely in an arbitrary place chosen by the user by installing the user-definable command space mentioned above. Figure 4 shows a conceptual image of the pseudo-relative coordinate system. The 3D position extracted by the first hand-waving gesture is defined as the origin of the pseudo-relative coordinate system and the center of the command space. The axis connecting the 3D positions extracted by the first and second hand-waving gestures and the direction perpendicular to the ground are defined as the Y axis and the Z axis, respectively. Therefore, the X axis is defined by the outer product of the Y axis and the Z axis. The relationship between absolute coordinates and pseudo-relative coordinates as shown in Fig. 5 is as follows:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \mathbf{R}\begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} \quad (1)$$

$$\mathbf{R} = \begin{pmatrix} \frac{y_2-y_1}{\sqrt{(x_2-x_1)^2+(y_2-y_1)^2}} & \frac{x_1-x_2}{\sqrt{(x_2-x_1)^2+(y_2-y_1)^2}} & 0 \\ \frac{x_2-x_1}{\sqrt{(x_2-x_1)^2+(y_2-y_1)^2}} & \frac{y_2-y_1}{\sqrt{(x_2-x_1)^2+(y_2-y_1)^2}} & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (2)$$

where $\boldsymbol{O}(x_1, y_1, z_1)$ and $\boldsymbol{Q}(x_2, y_2, z_2)$ denote the 3D positions of the first and second hand waving, respectively. Here, z_2 is set equal to z_1 to expand the command space horizontally. $\mathbf{R}$ is a rotation matrix between the absolute coordinate system and the pseudo-relative coordinate system.

III. Home Appliance Operation

A. *Hand position detection*

After constructing the command space, the operation of the home appliance is performed based on a simple hand gesture, which is hand positioning instead of hand waving. Background subtraction and frame subtraction are used to

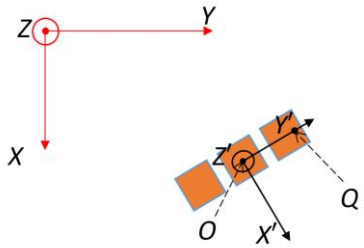

Figure 5. Relationship between an absolute coordinate frame and a pseudo-relative coordinate frame. Here, orange rectangles represent the command spaces.

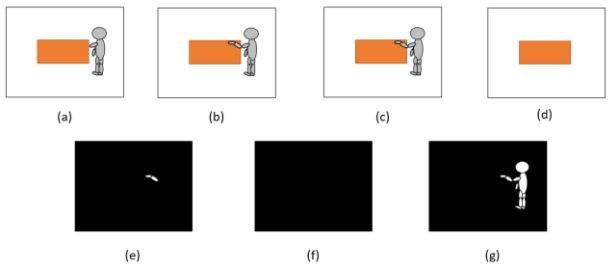

Figure 6. Illustration of background subtraction and previous frame subtraction: (a) image of t=k-1, (b) image of t=k, (c) image of t=k+1, (d) background image, (e) result of frame subtraction between the image of t=k-1 and the image of t=k, (f) result of frame subtraction between the image of t=k and the image of t=k+1, and (g) background subtraction result on the image of t=k+1.

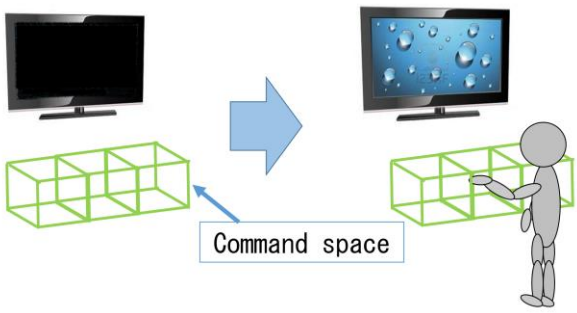

Figure 7. Conceptual image of the command space.

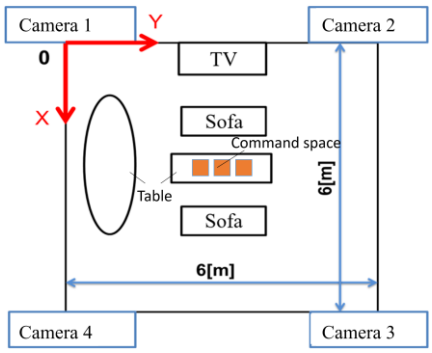

Figure 8. Experimental environments of the home appliance operation system.

detect the hand positioning gesture with reduced resolution. The process of detecting hand positioning is shown in Fig. 6. First, the image of $t=k$ (Fig. 6 (b)) is subtracted from the image of $t=k-1$ (Fig. 6 (a)). White pixels shown in Fig. 6 (e) are the result of frame subtraction. Second, frame subtraction is performed again between the images of $t=k$ (Fig. 6 (b)) and $t=k+1$ (Fig. 6 (c)). When the hand gesture is performed as in this case, white pixels are not extracted, as shown in Fig. 6 (f). Third, background subtraction is performed on the image $t=k+1$ (Fig. 6 (c)), and the result is shown in Fig. 6 (g). The white pixels shown in Fig. 6 (e) are still white as shown in Fig. 6 (g). Using the property described above, the region changed in the order of white pixels, black pixels, and white pixels is detected as the region where the hand positioning gesture has occurred. Finally, a 3D position for where the hand positioning gesture was performed is acquired based on the principle of stereo measurement. Therefore, each home appliance operation is performed according to the position of the hand.

B. *Device operation by hand positioning*

By associating each constructed command space with the operation of the home appliance, it is possible to operate the home appliance by a simple hand positioning gesture. Here, we use the concept of spatial memory, in which information is embedded in the space [10]. If the hand positioning gesture in the command space is performed, the specific command of the associated home appliance can be executed. A conceptual image of the command space in this study is shown in Fig. 7.

IV. EXPERIMENTS

A. *Experimental environments*

Figure 8 shows the experimental environments constructed based on a proposed intelligent room that includes four cameras to operate the home appliance. The cameras' angle of view was 55.8° × 43.3°. The raw resolution of the image was 640 × 480 pixels, and the reduced resolution was 80 × 60 pixels.

B. *Configuration of the command space*

As shown in Figs. 9 and 10, the command space defined in this study is arranged in a total of six pieces of two steps up and down arranged side by side. The command space is projected on a separate display, as shown in Fig. 11, to show the relative position intuitively. Here, the purple rectangle indicates the target home appliance.

C. *Verification of operational accuracy*

The command space was expanded on the table as shown in Fig. 8, and an accuracy verification experiment was performed. Command spaces were constructed 20 times by a hand-waving gesture to verify the detection accuracy under various experimental conditions. Then, the hand positioning gesture was performed 20 times in each constructed command space, and its recognition rate was evaluated. The situation of the experiment is shown in Fig. 12.

The recognition rate shown in Table 1 is the average recognition rate for each of the six command spaces. The total average recognition rate was 85 % at place A. Thus, reliable home appliance operation was possible with a hand

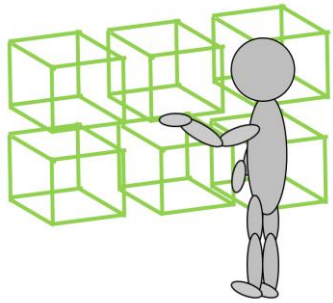

Figure 9. Arrangement of the command spaces.

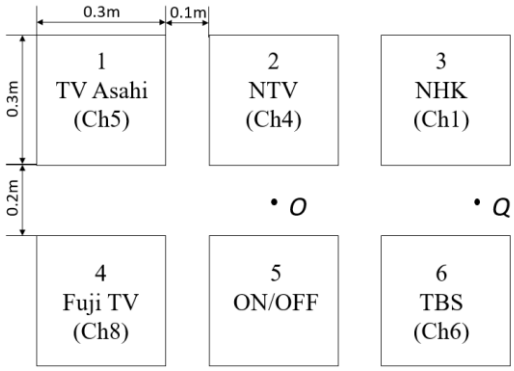

Figure 10. Front view of the command spaces.

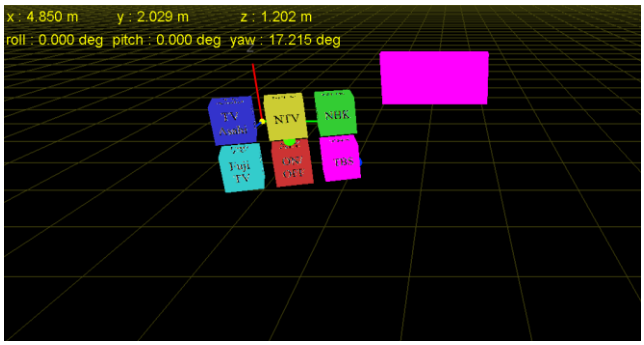

Figure 11. Command space display.

positioning gesture at each command space. The recognition rate in TV Asahi is relatively lower than that of other commands because it is farthest away from the center of the user's arm. Additionally, the recognition rate in the lower command space is higher than that in the upper command space because an elbow or other part may react (i.e., false detection) in the lower command space when hand positioning occurs on the upper side.

D. Verification of the time required for operation

We conducted an additional comparative experiment to determine the operation time required to recognize the hand waving and hand positioning gestures in six command spaces. The operation time here is the time spent on hand positioning or hand waving all of the six command spaces. We measured 10 times for each gesture in order to reduce the random error. The experimental results are shown in Table 2. The average time required for operation was 22.692 s by hand positioning gesture and 25.884 s by hand waving. The time required for hand positioning gesture is 12.33 % faster than that for hand waving. Moreover, the time required for hand waving was over 30 s, since the user became tired. Otherwise, the time to operate with the hand positioning was almost 25 s or less. However, it was necessary to wait for 3–4 s after positioning the hand because it did not output when the detection is stabilized on the system. Decreasing the waiting time will be one of the most important goals of future work.

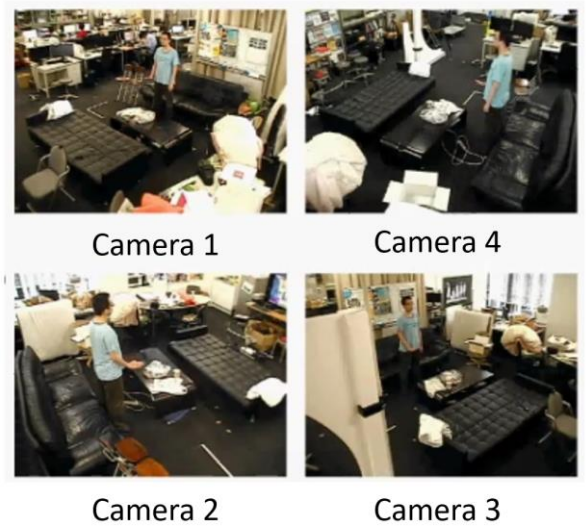

Figure 12. Experimental situation.

Table 1. Recognition rate for each position

	A
TV Asahi	65%
NTV	85%
NHK	80%
Fuji TV	90%
ON/OFF	95%
TBS	95%
Average	85%

Table 2. Time required for operation

Times	Positioning [s]	Waving [s]
1	21.69	20.01
2	18.86	26.60
3	26.31	22.43
4	28.54	23.30
5	22.28	26.07
6	22.62	32.88
7	15.01	30.00
8	27.50	31.70
9	22.46	26.29
10	21.65	19.56
Average	22.692	25.884

V. CONCLUSION

In this paper, we introduced a pseudo-relative coordinate system in order to improve the operation of a home appliance system that works by a user's hand waving and hand positioning without space constraints on the construction of the command space. Future work related to this paper will involve improving the recognition rate of hand waving and hand positioning from the image data for constructing the command space and operating the home appliance. In addition, we also need to improve the detection accuracy of the command space farthest from the arm.

REFERENCES

[1] C. Hsieh, D. Liou and D. Lee, "A Real Time Hand Gesture Recognition System using Motion History Image," 2010 2nd International Conference on Signal Processing Systems, pp. 394-398, July, 2010.

[2] J. Cho, D. Park and Y. Kim, "A method of remote control for home appliance using free hand gesture," 2012 IEEE International Conference on Consumer Electronics (ICCE), pp. 293-294, Jan., 2012.

[3] E. Peshkova, M. Hitz and B. Kaufmann, "Natural Interaction Techniques for an Unmanned Aerial Vehicle System," IEEE Pervasive Computing, vol. 16, no. 1, pp. 34–42, 2017.

[4] G. Plouffe and A. M. Cretu, "Static and Dynamic Hand Gesture Recognition in Depth Data Using Dynamic Time Warping," Trans. On Instrumentation and Measurement, vol. 65, no. 2, pp. 305–316, 2016.

[5] T. Nagayasu, H. Asano, M. Takahashi and K. Umeda, “Improvement of an Intelligent Room that Detects Hand Waving Motion for Operation of Home Appliances,” Proc. SICE Annual Conference, pp. 821–826, 2011.
[6] K. Terabayashi, H. Asano, T. Nagayasu, T. Orimo, M. Ohta, T. Oiwa and K. Umeda, “Detection of Small-Waving Hand by Distributed Camera System,” Proc. 9th International Conference on Networked Sensing Systems (INSS2012), pp. 1–5, 2012.
[7] T. Kano, T. Kawamura, H. Asano, T. Nagayasu and K. Umeda, “Hand Waving in Command Spaces: A Framework for Operating Home Appliances,” Advanced Robotics, vol. 32, no. 18, pp. 999–1006, 2018.
[8] S. Yan, Y. Ji and K. Umeda, “A System for Home Appliance Operation by Hand Waving in a User-definable Command Space,” Fourteenth International Conference on Quality Control by Artificial Vision, vol. 11172, no. 14, July 16, 2019.
[9] H. Nakayama, N. Yabuki, H. Inoue, Y.Sumi and T. Tsukutani, “A control system for electrical appliances using eye-gaze input,” 2012 International Symposium on Intelligent Signal Processing and Communications Systems, pp. 410-413, Nov., 2012.
[10] M. Niitsuma, H. Kobayashi and A. Shiraishi, “Enhancement of Spatial Memory for Applying to Sequential Activity,” Journal of Advanced Sciences, vol. 9, no. 1, pp. 121–137, 2012.