Journal of Computers Vol. 30 No. 1, 2019, pp. 174-182
doi:10.3966/199115992019023001017

Natural User Interface Design for Control of Robots by Hand Gestures

Chien-Hsing Chou*, Shu-Wei Lin

Department of Electrical Engineering, Tamkang University, New Taipei City 251, Taiwan
chchou@mail.tku.edu.tw; quierfake23@gmail.com

Received 9 September 2017; Revised 14 November 2017; Accepted 10 January 2018

Abstract. In this study, we design a human–robot interaction (HRI) interface for controlling robots by using hand gestures. The design employs the concept of the natural user interface, according to which intuitive hand gestures are used to operate a robot. This type of interaction can reduce the time needed by users to familiarize themselves with the control interface and improve user experience. The proposed HRI interface acquires hand gestures of users by using Kinect, and analyzes the gesture type and palm movement of users to operate the robot. The proposed operational method simulates in imaginings of telekinesis, directly connecting the human sense of space with robots and providing users and robots with a more intuitive interaction model. According to the results of field trials, users could rapidly familiarize themselves with the operational method for controlling the movement of a robot after 1 minute of oral instruction. In field trials, 94% of the participants expressed satisfaction with the design of the interface.

Keywords: hand gesture, human robot interaction, nature user interface, robot control

1 Introduction

The increasing popularity of robots has led to an increase in the number of researchers working on robotics [1-2]. The content of robotics research has traditionally included topics such as control, mechanical design, and application services. In recent years, an increasing number of researchers have investigated human–robot interaction (HRI) [3-6] to examine ways of designing a user-centered robot control model in order to enhance user experience of controlling robots. In HRI, a conventional operational interface is primarily composed of a keyboard, mouse, or joystick. However, in some applications, such as controlling high degrees-of-freedom machine arms and humanoid robots or controlling swarm robots, this conventional control interface is not the most suitable. Moreover, before users operate a robot, they must first become familiarized with a complicated operational flow and software environment, and they can operate it smoothly only after considerable practice. Because these training processes are often time consuming and tedious, several researchers have recently proposed the concept of natural user interface (NUI) to control robots [7-9]. Based on past experience and human thought models, researchers have designed more intuitive and reasonable interaction schemes to control robots. These schemes enable users to employ simple and intuitive methods to control robots, and they provide favorable user experiences.

To design a more intuitive HRI interface, Katić et al. [4] used Kinect to track the human skeleton and designed six types of hand gestures to control the movement of a four-wheeled robot. These six types of hand gestures can cause a robot to move forward, stop, turn right, turn left, slow down, and speed up. Although users can control robots through different hand gestures, there were no substantial connections between the gestures to their corresponding control commands. Moreover, during actual testing, the usability of this HRI interface was not assessed. Bandeira et al. [8] used Kinect to design four types of hand gestures for controlling a two-wheeled robot: move forward, move backward, turn right, and turn

* Corresponding Author

left. The hand gestures and robot controllers designed by them were consistent with natural instincts; for instance, the hand gesture to move a robot forward involved pushing both hands forward. The connection between the hand gesture and the “move forward” command was highly intuitive, and users could intuitively and rapidly understand the robot control methods.

Although several studies have been conducted on human–robot interaction [4-6], most such studies have focused on examining communication methods and constructing sensors in the system, and they lack discussions on HRI. Therefore, the aim of this study is to develop a human-centered interface to control robots. We use the movement of a three-wheeled robot as the control target to design a hand gesture-based NUI, and the effectiveness of the proposed NUI is investigated in actual testing.

2 Robot Control Interface with Hand Gesture

In this study, a human-centered interface to control a three-wheeled robot is proposed. We design a hand gesture-based NUI to control the movement of the robot. When a user makes a fist, the robot will cease all actions and wait for a new instruction. When the user opens her/his palm and moves it, the robot moves its position corresponding to the movement of the user’s palm. Fig. 1 shows the relationship between the palm position and the robot position. The idea is simple and intuitive—if the user moves her/his hand forward, the robot moves forward to new position along the y-axis. If the user moves her/his hand backward, the robot moves back to a new position along the y-axis. Fig. 2 shows an example to illustrate the proposed HRI interface. When the user moves her/his palm from the palm-starting position $\overrightarrow{P_{p\_0}}$ to the palm-ending position $\overrightarrow{P_{p\_e}}$, a palm-moving vector $\overrightarrow{V_{p\_m}}$ is calculated by subtracting $\overrightarrow{P_{h\_e}}$ and $\overrightarrow{P_{p\_0}}$. A robot-moving vector $\overrightarrow{V_{r\_m}}$ is then calculated by multiplying the palm-moving vector $\overrightarrow{P_{p\_m}}$ with a scale parameter K, where K is the movement-scaling parameter set by the user. Finally, the robot moves until it reach the target position $\overrightarrow{P_{r\_t}}$.

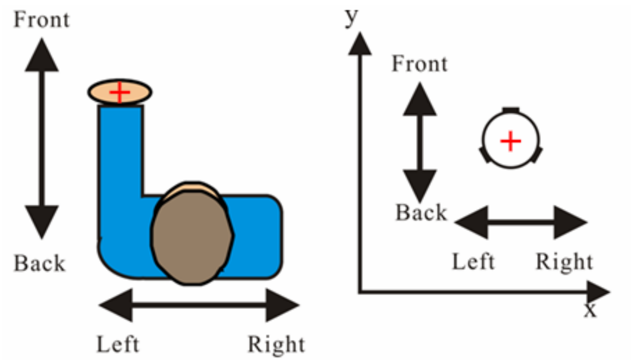


Fig. 1. Relationship between palm position and robot position

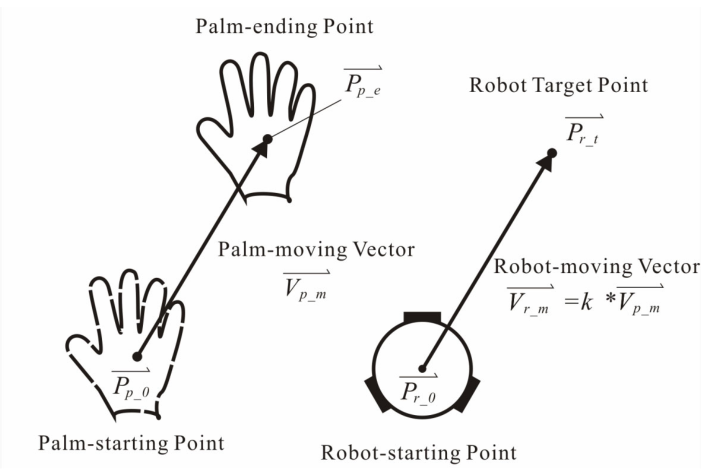


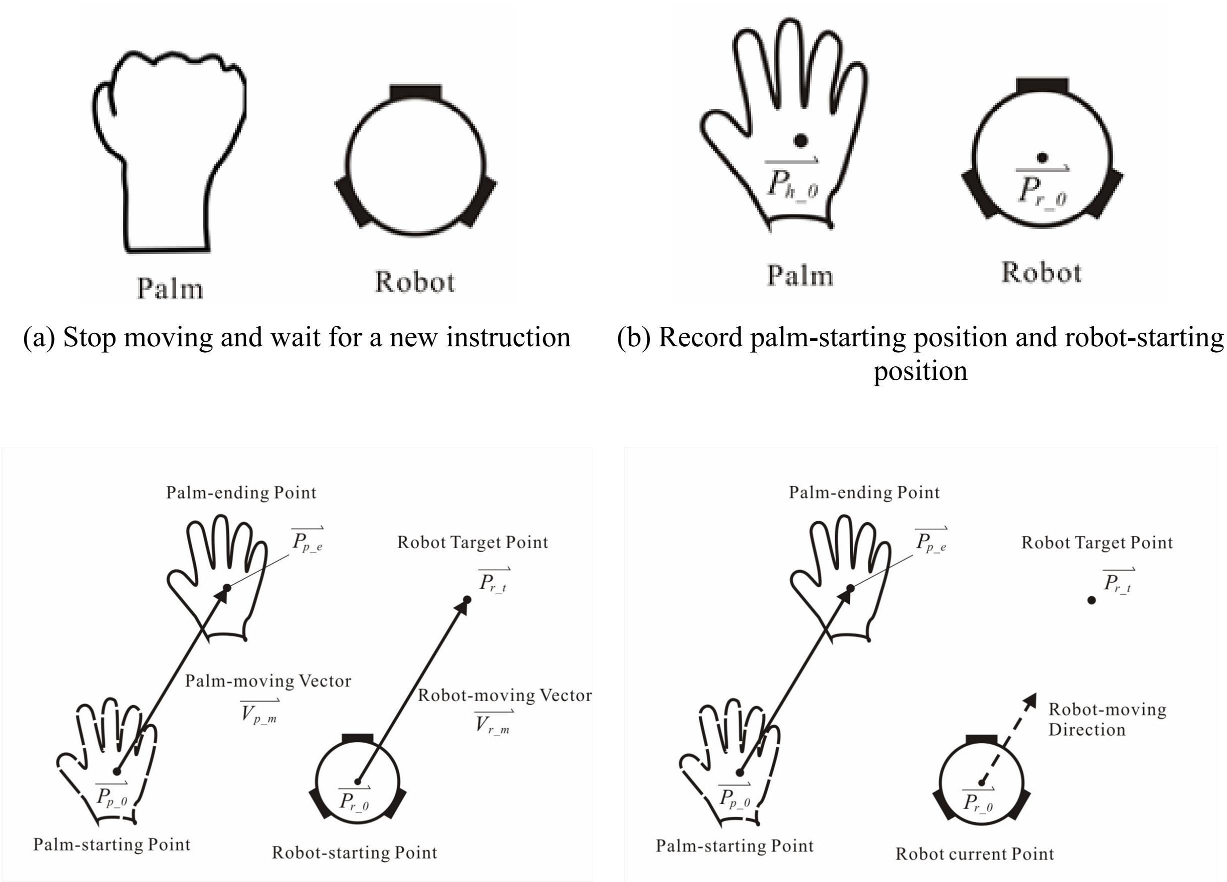
Fig. 2. Relationship between palm displacement and robot movement

In summary, the proposed HRI interface applies fist closing and fist opening to stop and start the robot movement, respectively; it applies the displacement of the user's palm in the horizontal plane to control robot movement and decide its target position. Given that the hand gestures and robot controllers designed in our research are consistent with natural instincts, the proposed HRI interface is not only simple but also easy to imagine. The control process of the proposed HRI interface is given as follows:

Step 1: Make a fist, and the robot stops moving and waits for a new instruction (Fig. 3(a)).

Step 2: Open the fist and record the palm-starting position $\overrightarrow{P_{p\_0}}$ and robot-starting position $\overrightarrow{P_{r\_0}}$ (Fig. 3(b)).

Step 3: If the user's palm is moving, calculate the palm-moving vector $\overrightarrow{P_{p\_m}}$ and then calculate the corresponding target position of the robot $\overrightarrow{P_{r\_t}}$ (Fig. 3(c)).



(a) Stop moving and wait for a new instruction

(b) Record palm-starting position and robot-starting position

(c) Calculate palm-moving vector and corresponding target position of robot

(d) Calculate moving direction of robot by applying a path-planning algorithm

Fig. 3. The control process of the proposed HRI interface

Step 4: Calculate the moving direction of the robot by applying a path-planning algorithm, and control the robot to approach the target position $\overrightarrow{P_{r\_t}}$ (Fig 3(d)).

Step 5: If the user make a fist, stop moving and return to Step 1. Otherwise, control the robot until it reaches the target position.

One thing should be emphasized here, if the user moves her/his palm to a new position $\overrightarrow{P'_{p\_t}}$ before the robot reaches the previous target position, as shown in Fig. 4, the new target position ($\overrightarrow{P'_{r\_t}}$) and moving direction the robot are recalculated, and the robot then moves toward the new target position.

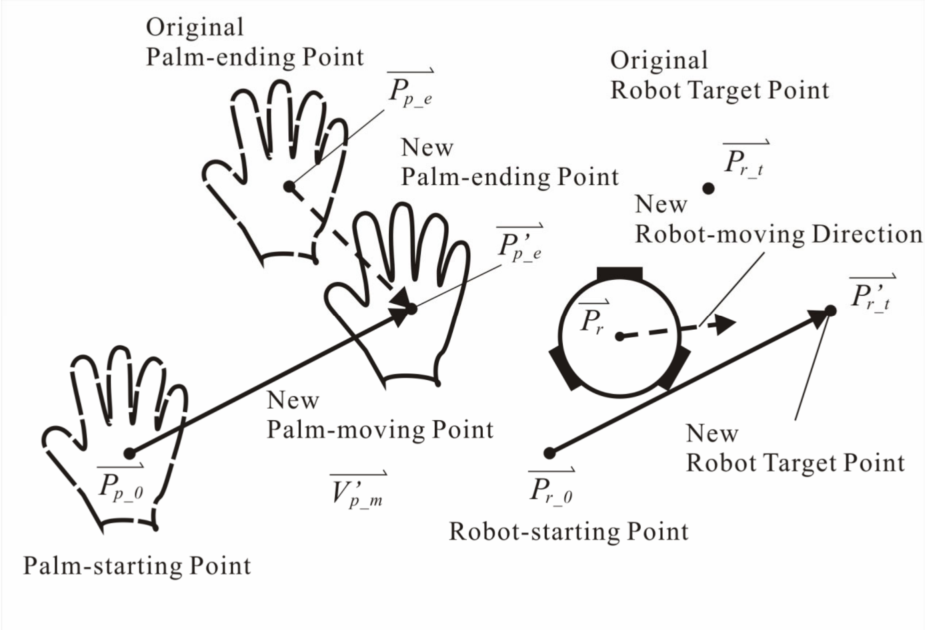


Fig. 4. Robot moves toward new target position if user moves her/his palm to a new position

3 Hardware of Robot Control System

The robot control system is divided in two parts, as shown in Fig. 5, namely the input subsystem and the output subsystem. The input subsystem includes a Microsoft Kinect (faced toward user) to detect the position of the user's palm and to extract an RGB image of the user's palm. A personal computer (PC1) analyzes these pieces of information to track the movement of the user's palm and identify whether the user's palm is opening in real time. According to the control process in Section 2, the operational instructions of the robot are then sent to the output subsystem to control the robot through Wi-Fi communication. The output subsystem uses another Microsoft Kinect (perpendicular to the ground) to extract depth information from the ground direction. Another personal computer (PC2) analyzes this depth information to locate the robot position and control its movement through Bluetooth communication.

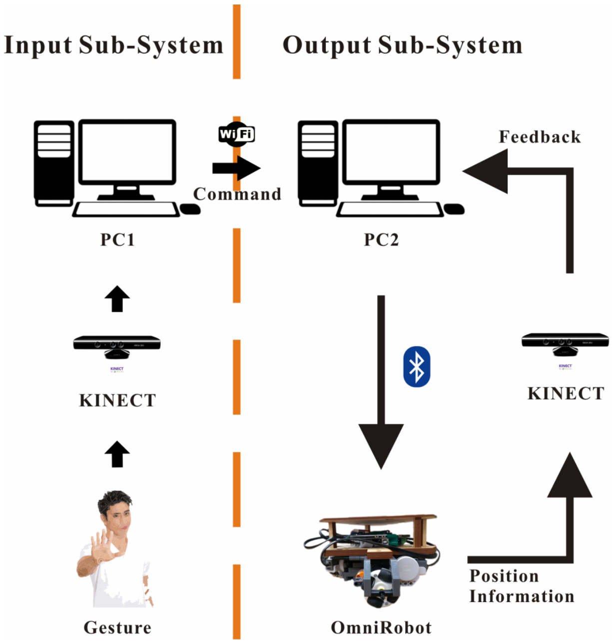


Fig. 5. Hardware architecture of the robot control system

According the operational instructions from PC1 and the position of the robot, PC2 calculates the moving direction of robot by applying a path-planning algorithm and controls the robot to approach the target position. Fig. 6 shows the scenario of controlling the robot in the present study.

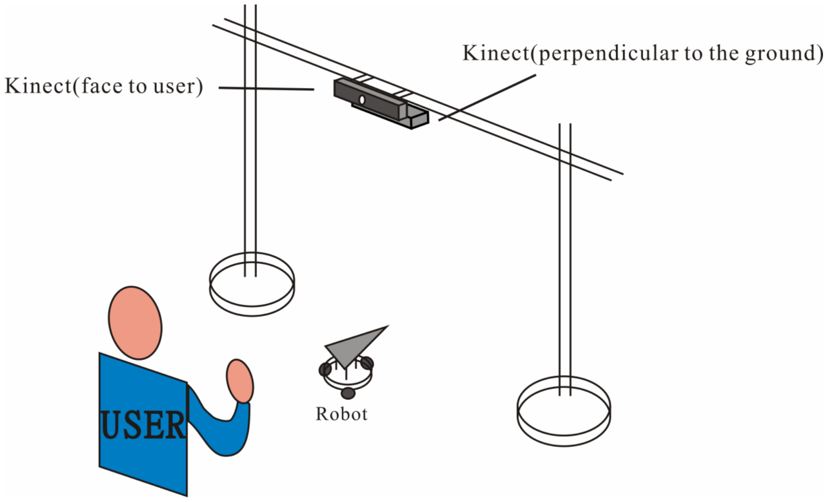


Fig. 6. Scenario of controlling the robot

Because the control method of a three-wheeled omni robot is more intuitive than that of a two-wheeled robot, in this study, we employ a three-wheeled omni robot for controlling (Fig. 7). The three-wheeled omni robot is equipped with an Arduino Uno as its embedded system. The embedded system was assigned two functions: (a) control three servo motors, (b) communicate with PC2 through Bluetooth communication. Fig. 7 shows an image of the three-wheeled omni robot.

Fig. 7. Three-wheeled omni robot

4 Experiment Design

To demonstrate the usability of the proposed HRI interface, an experimental test was conducted. Seventeen subjects participated in this experimental test. Each subject learned how to control the robot after 1 min of oral instruction and then controlled the robot continuously to arrive to seven different positions. Fig. 8 shows the seven positions in this experimental test.

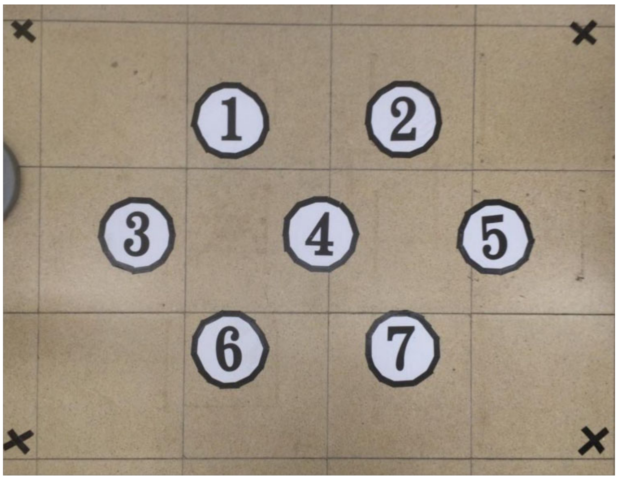


Fig. 8. Seven positions in the experimental test

In this experimental test, we focused on four test targets. The following discussions introduce the methodology for testing the four test targets:

Spending time to become familiar with the operational method. In the experimental test, each subject was asked to complete the trial three times. We recorded the time spent on the three trials and compared it with the time spent by a skilled user. Our expectation was that a subject may spend less time to finish a new trial in subsequent trials, and the time spent would be close to that spent by a skilled user.

Efficiency of completing mission trial. For the sake of comparison with the proposed HRI interface, a skilled user used a graphical user interface (GUI) for controlling the robot to complete the same trial. Fig. 9 shows the GUI tested in the experiment. The operating area was limited to inside the outer circle, and the black small square in the upper-right part of the figure represents the position of mouse clicking. When a user clicked on the operating area, the robot moved toward the new direction. By comparing the time spent, we could evaluate the efficiencies of the two aforementioned user interfaces.

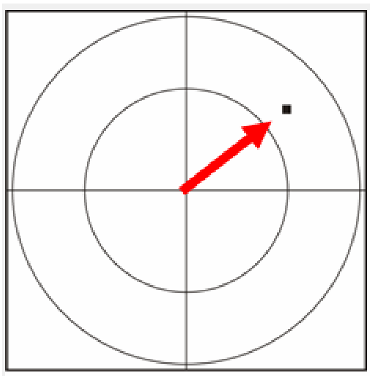


Fig. 9. Testing a graphical user interface to control the robot

Frequency of error occurrence. If an error occurred during the trial execution, we recorded the underlying reason for the error and its frequency.

Evaluation of user experience. After a subject finished the experimental test, we collected her/his comments and experiences about the proposed HRI interface.

5 Experiment Results and Discussions

We invited 17 subjects to conduct the experimental test; among them, one subject continuously misread the numbers during mission execution, leading to inaccuracy in the experiment. Consequently, the 16 sets of valid data listed in Table 1 were used.

Table 1. Spending time of 16 subjects for three mission trials

Subject ID	Trial 1 (s)	Trial 2 (s)	Trial 3 (s)
1	136	96	95
2	106	74	76
3	183	186	147
4	169	177	150
5	158	158	68
6	143	64	64
7	89	71	63
8	77	76	93
9	113	86	88
10	132	87	78
11	277	149	101
12	93	92	66
13	93	90	79
14	94	93	71
15	156	120	85
16	129	77	68
Average Time (s)	134	106	87

According to our field trial conducted in advance, a skilled user spent an approximately 55 s on average to complete a mission trial. To compare the operational time of the subjects with that of a skilled user, we plotted the time spent by the subjects and that of a skilled user in Fig. 10. The chart shows that after three trials, 63% of subjects became skilled in controlling the robot (<80 s); however, three of the subjects did not become skilled (>100 s). Although several subjects exhibited slower operational speeds in the experiment, most subjects demonstrated substantial improvement during the experimental process. For example, subject 11 spent 277s in the trial 1 but 149s and 101s in trials 2 and 3, respectively. Therefore, the subjects could rapidly familiarize themselves with the proposed HRI interface.

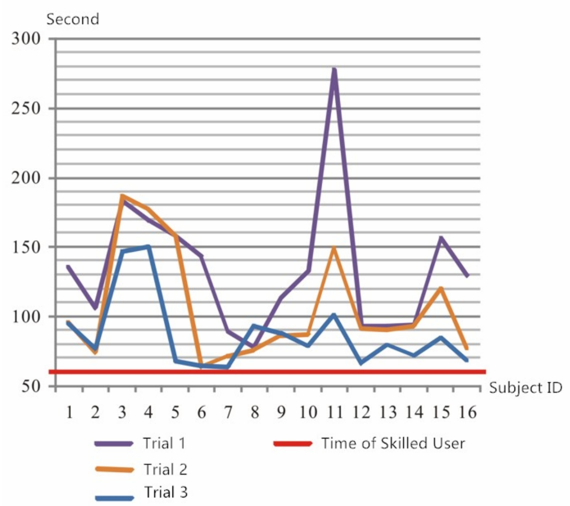


Fig. 10. Time spent by 16 subjects and a skilled user

Another comparison between the GUI and our user interface is summarized in Table 2. According to the test results, both robot operation interfaces had nearly equivalent efficacy, demonstrating the efficiency of the user interface designed in this paper. In addition, the most common errors that occurred during mission execution by the subjects were as follows:

Table 2. Spending time of GUI and our user interface for 5 mission trials

	GUI (s)	Our User Interface (s)
Trial 1	55	53.6
Trial 2	49.8	53.3
Trial 3	56.9	57.2
Trial 4	51.2	55.8
Trial 5	48.3	56.2
Average Time (s)	52.2	55.2

Exceeding the image detection range. This error included moving overly close to the Kinect or exceeding the detection boundaries of Kinect. This error could be improved by the addition of prompts that can be provided when the subject's palms leave the detection boundaries or are too close to Kinect.
Misunderstanding the operational methods. In this study, the movement of a user's palm represents a vector rather than a direction. Occasionally, the subjects misunderstood the operational method and controlled the robot by outputting an incorrect direction. After the subjects were reminded of the operational methods, the operational process of controlling the robot improved greatly.

After completing three trials, 94% of the subjects provided positive feedback based on their user experiences, as follows: "There is a feeling that it follows your wishes," "It has a very science-fiction feel," and "It feels as if you really are interacting with the robot." Based on the aforementioned evaluations, we believe that the proposed HRI interface provides exceptional user experience. However, 31% of the subjects felt hand fatigue, and during continuous arm extension, a few subjects experienced discomfort in their arms. Therefore, the proposed HRI interface is not suitable for long-term operation. Finally, regardless of whether operation was smooth, 35% of subjects suggested that the accuracy of palm recognition and movement detection could be enhanced, which we plan to work on in the future.

6 Conclusion

In this study, we designed an HRI interface for controlling the motions of a robot. The design employed the NUI concept, wherein intuitive hand gestures were used to operate a robot. The proposed operational method simulates in imaginings of telekinesis, directly connecting the human sense of space with robots and providing users and robots with a more intuitive interaction model. This type of interaction can shorten the time required by users to familiarize themselves with the control interface, and it can provide favorable user experiences. According to field trials, most of the study participants expressed satisfaction with the designed interface.

Acknowledgements

This work was supported by the Ministry of Science and Technology, Taiwan, R.O.C., under the Grant MOST 106-2221-E-032-047.

References

[1] J. Li, L. Zhang, The teleoperation system of service robot based on cloud services, Journal of Computers 28(2)(2017) 231-245.

[2] T.-F. Wu, H.-C. Huang, P.-S. Tsai, N.-T. Hu, Z.-Q. Yang, The tracking control design of adaptive fuzzy CMAC for an omni-directional mobile robot, Journal of Computers 28(1)(2017) 247-260.

[3] O.M.-I.E. Sucar, S.H. Aviles, C. Miranda-Palma, From HCI to HRI - usability inspection in multimodal human-robot interactions, in: Proc. IEEE International Workshop on Robot and Human Interactive Communication, 2003.

[4] R. Kimmel, N. Kiryati, A.M. Bruckstein, Distance maps and weighted distance transforms, Journal of Mathematical Imaging

and Vision, Topology and Geometry in Computer Vision 6(2-3)(1996) 223-233.

[5] S.E. Ghobadi, O.E. Loepprich, F. Ahmadov, J. Bernshausen, K. Hartmann, O. Loffeld, Real time hand based robot control using multimodal images, IAENG International Journal of Computer Science 35(4)(2008) 500-505.

[6] Z. Ren, J. Yuan, Z. Zhang, Robust hand gesture recognition based on finger-earth mover's distance with a commodity depth camera, in: Proc. the 19th ACM International Conference on Multimedia, 2011.

[7] D. Katić, P. Radulović, S. Spasojević, Ž. Đurović, A. Rodić, Advanced gesture and pose recognition algorithms using computational intelligence and microsoft KINECT sensor, in: A. Rodić, D. Pisla, H. Bleuler (Eds.), New Trends in Medical and Service Robots. Mechanisms and Machine Science, vol. 20, Springer, 2014, pp. 193-207.

[8] G.M. Bandeira, M. Carmo, B. Ximenes, J. Kelner, Using gesture-based interfaces to control robots, in: Proc. International Conference on Human-Computer Interaction, 2015.

[9] D. Norman, Natural user interfaces are not natural, Interactions 17(3)(2010) 6-10.