System Design to Evaluate Guidance by Robots Based on Immersive VR

Yoshiaki Mizuchi[1] and Tetsunari Inamura[1,2]

***Abstract*— Evaluating the ability of robots to guide humans is necessary to investigate strategies concerning more appropriate guidance, such as for recognizing situations or explaining with friendly expression. Using questionnaires is a standard approach for evaluating functions of robots depending on subjectivity and sensibility of humans. However, human behavior itself including unconscious acts cannot be observed with questionnaires. On the other hand, observing human behavior in real environments requires the huge cost of tests due to constructing various environments and recruiting research subjects. The use of cloud-based immersive virtual reality (VR) is a solution to reduce such costs. To this end, we propose a new system design to evaluate guidance by robots based on observed behavior of humans guided in an immersive VR environment.**

I. INTRODUCTION

In regard to robot intelligence, the ability of robots to guide humans is one important factor. For example, in case of teaching cooking procedure to an inexperienced user or asking for help from humans, robots must explain what to do to induce suitable acts of a user. Evaluating such ability of robots is required to investigate more superior strategies concerning recognition of a situation, generation of friendly expression, or natural language processing.

Use of questionnaire is a standard approach for evaluating functions of robots concerning subjectivity and sensibility of humans. Since the human behavior in daily-life activities is complex and ambiguous, the large number of samples are required to obtain reliable results. Furthermore, unconscious behavior such as the presence of superfluous acts cannot be observed from the result of questionnaires. Although human behavior itself must be observed, the huge cost of tests is inevitable to construct various environments and recruit research subjects.

One of the solutions to reduce the above cost is to use virtual reality (VR) technique. A new system design to evaluate guidance by robots based on observed behavior of humans guided in an immersive VR environment is proposed. The proposed system enables research subjects to log in to a VR avatar and to perform a task based on guidance by a virtual robot in a similar manner to a real environment. Results of demonstration carried out in RoboCup JapanOpen 2017 are shown to confirm the availability of the proposed evaluation system.

*This paper is based on results obtained from a project commissioned by the New Energy and Industrial Technology Development Organization (NEDO).

[1]Yoshiaki Mizuchi and Tetsunari Inamura are with National Institute of Informatics, 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo, Japan {mizuchi, inamura}@nii.ac.jp

[2]Tetsunari Inamura is with the Department of Informatics, SOKENDAI(The Graduate University for Advanced Studies), 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo, Japan

Fig. 1. Environment used for the test

II. EVALUATION SYSTEM FOR GUIDANCE BY ROBOTS

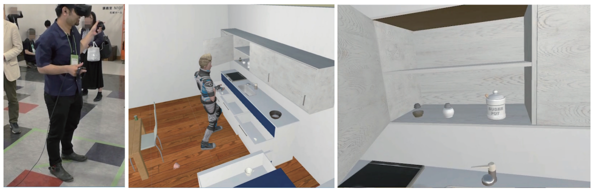
We developed a new version of the SIGVerse (Ver. 3) [1] that integrates Unity and ROS middleware to support a variety of VR devices and software resources created by the ROS community. Participants can log in to an avatar via VR interfaces such as head-mounted displays (HMDs), motion capture devices, and audio headsets. According to the input from such VR devices, behavior of the participant is reflected on the avatar by Unity scripts. Perceptual information such as perspective visual feedback is provided to the participant. Thus, the participant can interact with the virtual environment in a similar manner to a real environment. SIGVerse (Ver. 3) has a bridging mechanism between ROS and Unity. ROS-based software for virtual robot control can be reused in real robots without modification, and vice versa.

Utilizing the above functionalities, we realize evaluation tests that a participant is guided by a virtual robot while logging in to an avatar. Guidance by the robot is evaluated from observed behavior of the avatar and trajectories of virtual objects.

III. DEMONSTRATION USING PROPOSED SYSTEM

A. Condition of the experiment

Using the proposed system, we carried out a test in the RoboCup JapanOpen 2017. This test was done to evaluate the ability of a robot to generate natural and friendly language expressions to explain how to get to a destination and where a target object is. The robot has to generate natural language expression in accordance with the layout of the environment and positions of candidate objects. In this test,



(a) Participant (b) VR avatar (c) First-person viewpoint of the avatar

Fig. 2. Condition of a participant and a VR avatar

TABLE I
TIME NEEDED FOR GRASPING THE TARGET OBJECT (UNIT: SEC.)

Condition	Team A	Team B	Team C
1	70.1	57.5	89.7
2	64.0	80.5	54.2
3	80.0	(300.0)	(300.0)
4	68.5	88.9	(300.0)
5	(300.0)	57.4	162.5
6	118.3	96.9	(300.0)
7	34.9	155.3	36.1
8	62.3	253.6	136.2
9	51.0	45.2	(300.0)
10	82.1	157.9	116.0
11	75.6	75.3	209.3
12	73.9	54.7	63.0
Total	**1080.7**	**1423.1**	**2067.1**

communication is done only through text messages and voice utterances. The robot cannot use any gestures or visual information.

An example VR environment for the test is shown in Fig. 1. We constructed four environments which consist of five rooms, a corridor, and an entrance room. Seventeen kinds of objects were used as graspable candidate objects. More than two objects of the same kind were placed in each environment. Some objects were placed in the inside of cabinets or drawers; therefore, participants have to open the door to find a target object. The robot itself does not exist in the environment because communication is carried out using only text messages and voice utterances.

The condition of a participant and an avatar are shown in Fig. 2. Participants wore an HMD (Oculus Rift CV1) to log in to the avatar. They controlled the avatar by using a handheld device (Oculus Touch). We asked six participants to take part in this task. Each participant was guided under six different conditions for twice each team.

B. Result of the experiment

Guide messages of team A included a route to a target room, furniture on which the target object placed, what the target object is, and positional relation with the similar object placed nearby the target object. Although guide messages of team B included similar kinds of information as team A, positional relation was not included. Guide messages of team C included only object ID. Additionally, many grammatical mistakes were included.

Time needed for grasping the target object is listed in Table I. Time in parentheses denotes cases that the participant could not take the target object within the time limit. Team A succeeded tasks under eleven conditions except condition 5. Although team B succeeded tasks under eleven conditions except condition 3, total time is 300 seconds longer than that of team A. Team C failed tasks under four conditions (3, 4, 6, 7). Total time of team C is longer than that of the other teams.

TABLE II
FREQUENCIES THAT PARTICIPANTS GRASPED INCORRECT OBJECTS

Condition	Team A	Team B	Team C
1	0	0	0
2	0	0	0
3	0	4	0
4	0	1	1
5	0	0	3
6	0	1	3
7	0	0	0
8	0	1	2
9	0	0	4
10	0	3	3
11	0	0	5
12	0	1	4
Total	**0**	**11**	**25**

Frequencies that participants grasped incorrect objects are listed in Table II. As well as the total time, total frequency of team A is the smallest, and that of team C is the largest. With respect to team B and team C, incorrect objects were grasped by participants in half and more of conditions; therefore, these results show that their guide messages were ambiguous. Thus, the ability to guide humans are evaluated statistically, thanks to the proposed system.

IV. CONCLUSION

A cloud-based immersive VR system for evaluating the ability of robots to guide humans is proposed. Although using questionnaire is a standard approach for evaluating such abilities concerning subjectivity of human, human behavior itself such as unconscious acts is not an evaluation factor. The cost for constructing various environments and recruiting research subjects is a critical problem to observe the human behavior. The proposed evaluation system allows to reduce the above costs and to carry out iterative evaluation. The results of demonstration in RoboCup JapanOpen confirm that the guidance by robots can be evaluated statistically based on human behavior in a VR environment.

Not only evaluating cognitive functions of robots but storing multimodal experience data in daily-life activities is available with the proposed system. The cloud-based immersive VR (SIGVerse[1]) is open source software; therefore, experimental environments and multimodal experience data can be shared and reused. This system has the potential to shorten development and evaluation cycles of intelligent robots or AI systems.

REFERENCES

[1] T. Inamura and Y. Mizuchi, "Competition design to evaluate cognitive functions in human-robot interaction based on immersive VR," RoboCup International Symposium, 2017.

[1] http://www.sigverse.org/wiki/en/