Development of RoboCup@Home Simulation towards Long-term Large Scale HRI

Tetsunari Inamura[1,2], Jeffrey Too Chuan Tan[1], Komei Sugiura[3], Takayuki Nagai[4], and Hiroyuki Okada[5]

[1] National Institute of Informatics, Japan
[2] The Graduate University for Advanced Studies, Japan
[3] National Institute of Information and Communication Technology, Japan
[4] The Univ. of Electro-communications, Japan
[5] Tmagawa Univ., Japan

Abstract. Research on high level human-robot interaction systems that aims skill acquisition, concept learning, modification of dialogue strategy and so on requires large-scaled experience database based on social and embodied interaction experiments. However, if we use real robot systems, costs for development of robots and performing many experiments will be too huge. If we choose virtual robot simulator, limitation arises on embodied interaction between virtual robots and real users. We thus propose an enhanced robot simulator that enables multiuser to connect to central simulation world, and enables users to join the virtual world through immersive user interface. As an example task, we propose an application to RoboCup@Home tasks. In this paper we explain configuration of our simulator platform and feasibility of the system in RoboCup@Home.

1 Introduction

An important task in the field of human-robot interaction (HRI) is elucidating the mechanisms of social and physical interactions and then embodying them into the design of robots. Completing this task requires the evaluation and modification of a hypothetical interaction model based on long-term and large-scale social interaction between people and robots. However, conventional HRI experiments are limited due to their laboratory setting, while large-scale experiments are quite costly in terms of people and time, especially if the aim of HRI is learning from demonstration or instruction such as was investigated by Sugiura et al. [1].

One of the purposes of the RoboCup@Home laps over the above goal. Typical tasks in the competition are designed on the basis of the assumption that robots must possess advanced social and embodied interaction functions. However, due to huge cost of developing robot hardware and executing experiments, benchmark tasks tend to focus on basic recognition functions and physical functions such as grasping, navigation, object recognition, and face recognition. The RoboCup@Home tasks are performed in a kitchen or living room environment,

S. Behnke et al. (Eds.): RoboCup 2013, LNAI 8371, pp. 672–680, 2014.

where a higher level of natural interaction is required. Such interaction includes dialogue management, intention understanding via gestures and eye gaze, and clarification of vague user instructions. Additionally, machine learning techniques for adapting to unknown environments and situations should be evaluated on the basis of past experience. These tasks shold be designed for RoboCup@Home not only from the viewpoint of competition but also from the viewpoint of academic interest. We believe that the RoboCup@Home competition should incorporate high-level intuitive and natural interaction tasks; however, the current rules requiring the use of real robots prevent the incorporation of such tasks.

Here, we present a RoboCup@Home simulation system with an immersive interface based on our SIGVerse simulation platform and discuss its feasibility. This platform was used in an official competition at the RoboCup Japan Open in 2013.

2 Current Mechanism of RoboCup@Home Competition

The current RoboCup@Home competition mechanism was designed on the basis of a benchmarking approach in which a set of functional abilities corresponding to the target robot technical abilities [2]. These functional abilities are weighted for use in the scoring of the yearly competitions. Table 1 shows the function weightings for the 2010 competition [2]. The cognition ability, which reflects the robot's high-level intuitive and natural interaction capability, was first introduced in 2010 and was given a weight of 13%, lower than the weighting for the two physical abilities, navigation and object manipulation. Cognition ability is mainly tested in the GPSR (General Purpose Service Robot) challenge. For the top five teams in 2012, the average achievement percentage for the GPSR challenge was only 8.2% (Table 2). This reflects the weakness of the cognition ability development at the latest RoboCup@Home competition.

Table 1. Function weighting for 2010 competition [2]

Ability	Weight(%)
Navigation	22
Mapping	9
Person Recognition	12.5
Person Tracking	3
Object Recognition	7.5
Object Manipulation	14
Speech Recognition	15
Gesture Recognition	3.5
Cognition	13

Table 2. Average score for top 5 teams in 2012 competition

Test	Max. Score	Score Weight	Avg. Score	Percent Achvd (%)
RIPS	1000	7.69	721.8	72.18
FM	1000	7.69	195	19.5
WW	2000	15.38	655	32.75
CU	1000	7.69	366.6	36.66
OC	2000	15.38	1249.8	62.49
GPSR	2500	19.23	205	8.2
DC	1500	11.54	644	42.93
Res	2000	15.38	260	13
Total	13000	100	4297.2	33.06

The above analysis clearly shows that the latest RoboCup@Home competition did not sufficiently stimulate cognition ability development. Due to the requirement of using a real robot in the competition, only very basic and limited interaction was possible. Focusing on the balance between cognition ability and the physical functional abilities is far more important than focusing on high-level robot intelligence.

We have developed a new approach to addressing these limitations: robot simulation. Moving the robot in a simulated world reduces the complexity of the low-level hardware issues, making it easier to focus on the high-level cognition issues. However, maintaining a rich interaction medium requires that people in the real world participate naturally through embodiment and multimodal interaction. We thus use various techniques to immerse a real-world person into the 3D virtual world and thereby enable human interaction with the virtual robot. This human participation (without a physical presence) supports our higher aim of having many human participants interact with an easy-to-maintain virtual robot in much deeper, longer term, and larger scale HRI. We believe this will significantly facilitate high-level robot intelligence development through the RoboCup@Home competitions.

3 Platform for RoboCup@Home Simulation

3.1 SIGVerse: SocioIntelliGenesis Simulator

We developed a platform for RoboCup@Home simulation on the basis of our SIGVerse [3] system, which enables easy development of HRI experiments in which people and virtual robot agents can interact socially and physically. This system enables arbitrary users to join virtual HRI experiments through the Internet with log-on to the central virtual world. It has three basic simulation modules: dynamics, perception, and communication.

Dynamics Simulation. The Open Dynamics Engine (ODE)[1] is used for dynamic simulation of interactions between agents and objects. Basically, the motions of each agent and object are calculated by the dynamics engine, and the user controls the calculations to reduce simulation costs. A switch flag can be set for each object and agent to turn off the dynamics calculations if required.

Perception Simulation. The perception simulation module provides the senses of vision, sound, force, and touch. OpenGL is used for visual simulation; it provides each agent with a pixel map that is a visual image derived from the agent's viewpoint and field of view. A distance sensor is also emulated. A robot agent can get the distance vector as a single dimension if the laser range finder is selected. A distance image as a two-dimensional matrix is available if stereo vision is selected.

For the sense of touch, it is possible to acquire information on the force and torque between objects, which are calculated mainly by ODE.

[1] `http://www.ode.org`

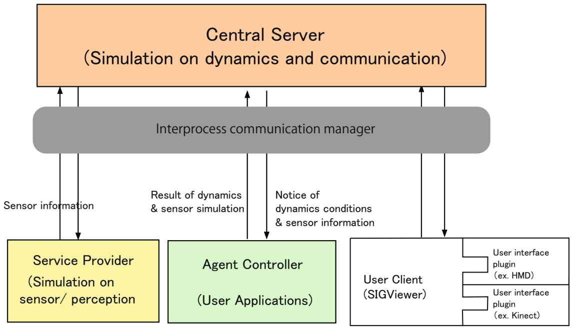

Fig. 1. Software configuration of SIGVerse

Communication Simulation. The sense of hearing is simulated by enabling every agent to communicate with audio data. The effect of sound volume is simulated by attenuating the sound in inverse relationship with the square of the distance under the assumption that the voice of an agent becomes more difficult to hear as the distance increases. It is possible to set the system so that only voices within a certain distance are audible.

With this system, not only can agents within the virtual environment communicate with each other but the virtual environment can interact with users in the real world.

There have been several studies [4][5] in which human agents act in a virtual social world based on the Second Life framework. However, it is difficult to implement dynamics and perception simulation.

3.2 Configuration of Simulator Software

SIGVerse is a client/server system consisting of a Linux server and a Windows client application. Dynamics calculations are mainly performed on a central server system. The behaviors of the robot and human avatars are controlled by an 'agent controller,' which is a dynamic link library on the Linux SIGVerse server. Autonomous actions and sensing functions of the agents are written using C++ APIs. Since the software libraries for ROS (Robot Operating System)[2] and OpenRTM[3] are also available, source code compatibility is supported. The avatars' behaviors are programmed using the APIs; they can also be controlled by user operation through the user interface in real time. Each user connects to the server system through the Windows client system. The configuration of the SIGVerse software is shown in Fig. 1.

[2] http://www.ros.org/wiki/
[3] http://www.openrtm.org/

4 User Interface for RoboCup@Home Simulation

4.1 Frequent Functions in RoboCup@Home Tasks

The tasks in the RoboCup@Home competition, as defined in the rule-book [6], are *Follow Me, Clean Up, Enhanced Who is Who,* and so on. In the *Follow Me* task, the robot has to recognize the facial image of the user, track the image, and follow the walking user. In the *Clean Up* task, the robot has to grasp a piece of trash targeted by the user and place it in a receptacle. In the *Enhanced Who is Who* task, the robot has to bring a drink in a cup or bottle to the user and hand it to the user. The implementation of these tasks requires the development of many basic functions such as receiving the user's instructions via speech and gestures, understanding the meaning of the instructions by speech recognition, image processing of pictures captured with a camera, grasping an object using end effectors, navigation by wheels, and dialogue management. These basic functions can be simulated in the SIGVerse world.

4.2 Immersive Interface for RoboCup@Home Simulation

We developed an inexpensive, flexible, and immersive interface for SIGVerse to enable social and embodied interaction with the virtual agents that frequently appear in RoboCup@Home tasks.

Projection of SIGVerse World to HMD. The interface uses a head mounted display (HMD) (eMagin Z800 3DVisor) that has a motion detector for the pan and tilt directions. Detected motion is transferred to the SIGVerse system for use in controlling the user's avatar. The HMD can display the sequence of images captured by the avatar's eye in accordance with the head motion. An overview of a user and the scene image projected on the HMD is shown in Fig. 4.2.

Control of Avatar Body Motion. Microsoft's Kinect controller can be connected to the latest SIGVerse client terminal. The motion pattern measured with the Kinect controller is transferred to the server for use in controlling the motion of the avatar's body. This enables virtual pointing gestures. In addition, the avatar can be made to grasp virtual objects by displaying a grasping motion in front of the Kinect controller. Since finger motion is difficult to measure with the Kinect controller, the system provides a grasp/release command for controlling avatar grasping. An image of avatar control using the Kinect controller is shown in Fig. 4.2.

These immersive interfaces can be connected to the SIGVerse client system as plug-in modules. There are also text-based chat and speech recognition and synthesis functions provided as standard user interfaces. Future developments in the RoboCup@Home simulation competition should lead to additional proposals from the participants for immersive interfaces such as a joystick controller and a motion capture system. Therefore, the SIGVerse client has a plug-in system

Fig. 2. A user joining a virtual world through HMD

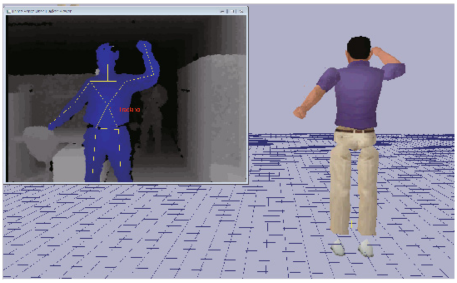

Fig. 3. Whole body control of an avatar by Kinect sensor

that accepts arbitrary interface systems such as users' original speech recognition systems, sound source detection systems, face recognition systems, and haptic interfaces.

4.3 Limitations of the RoboCup@Home Simulator Interface

In the *Follow Me* task, the robot has to track the user's facial image and follow the walking user; however, the SIGVerse user interface is fixed on a client computer. Therefore, the user has to interact in a limited area covered by fixed sensors, such as the Kinect sensors. Although SIGVerse can be connected with motion capture systems, such systems are difficult to use as an interface for the RoboCup@Home competition.

Due to the huge computational cost of simulation physical grasping due to the need to consider the friction coefficient of soft and flexible material, the proposed system does not currently simulate grasping. Instead, it uses a binary status: grasping/not grasping. Likewise, the handing over of an object from a robot agent to a user avatar is only reflected as a change of status. The user does not feel the reaction force.

We do not aim to implement all of the RoboCup@Home simulation tasks in the SIGVerse simulator. Instead, we aim to implement advanced interaction scenarios that require high-level agent functionality such as decision making, machine learning from conversation, and inference of future events. Examples of such scenarios are presented in the next section.

5 Implementation of RoboCup@Home Simulation Tasks

5.1 Clean Up Task

A simulated version of the Clean Up task introduced in the 2011 RoboCup@Home competition [6] is practical (Figure 5.1). This task tests the robot's abilities to detect, recognize, and manipulate objects, to navigate, and to systematically search and explore. The robot is directed to explore a room and determine

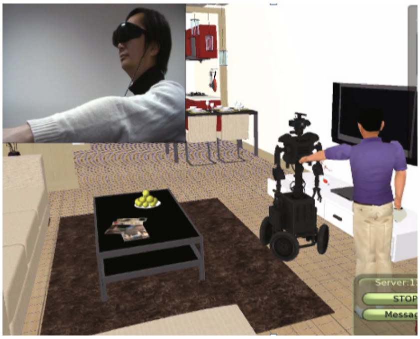

Fig. 4. Screen-shot of Clean Up task

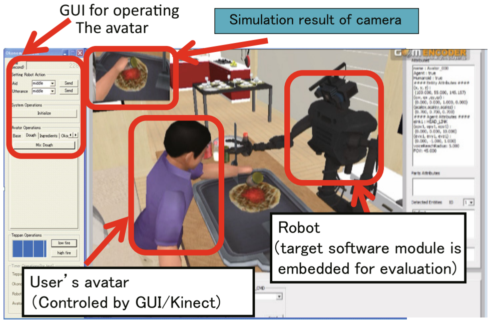

Fig. 5. Screen-shot of Cooperative Cooking task

whether known and unknown objects are items to be discarded. In this simulation, techniques similar to those used by real robots can be simulated on a virtual robot, for instance, the use of computer vision (e.g., OpenCV) to perform image processing for object detection and recognition.

Another function that can be simulated is using natural language and gesture instruction to recognize which object is being referred to by the user. A user might ask the robot to move and/or manipulate an object by saying something like, "Please bring that dish to the dining table" while pointing to the dish. If the pointing and/or speech are vague, the robot should be able to ask appropriate questions to remove the uncertainty. Such dialogue management is a high-level interaction function inherent in high-level HRI.

5.2 Cooperative Cooking Task (Future Candidate)

Tasks not described in the rulebook can also be simulated in the SIGVerse system and fit within the scope of the RoboCup@Home competition. One such task is a Cooperative Cooking task, which is a very high-level HRI task [7]. The task requires recognition of human behavior, real-time planning in accordance with recipe, dialogue management, and so on. Here, we describe an implementation of the Cooperative Cooking task as a future representative task. Since typical cooking tasks are cutting foodstuffs, grasping dishes and cups, operating microwaves, and so on, observation of such behaviors can focus on upper-body motions, and it is easily accomplished with a Kinect controller. Figure 5.1 shows a screen shot of this task.

The evaluation target is an effective human-robot interaction strategy. For example, if the user is not taking a cooking action, the robot should ask the user to do something complementary to what the robot is doing. If the user is taking a cooking action, the robot should search for, find, and take a suitable complementary action. Implementation of the required action selection algorithm would make the system more effective.

We performed preliminary experiments corresponding to two cases: one in which the user performed all of the actions through the interface, and one in which the robot performed complementary tasks as well. In the first case, the task took 194 [sec], and in the second case it took 58 [sec]. The task completion time would be an evaluation criterion in the RoboCup@Home competition.

6 Conclusion

Social and embodied interaction and adaptation and learning methods based on big data will be important issues in the next stage of human-robot interaction. Implementation of a simulation platform on which many people can easily participate in social interaction experiments will be a breakthrough in promoting human-robot interaction research.

We have developed an immersive interface for our SIGVerse simulation platform. Effective tasks for the RoboCup@Home competition can be developed by using a multi-user connection environment on the Internet. The interaction is currently one-to-one; however, it is easy to develop many-to-many human-agent interaction scenarios based on the multi-user connection system. Use of this simulation platform should lead to various contributions such as an autonomous judge system and machine learning competition for intelligent behavior through long-term large-scale interaction between robots and humans.

The proposed multi-user connection system has limitations, so its application is limited to simulating interactions that use kinematic motion, eye direction, and speech. Nevertheless, the tasks still have a moderate level of complexity.

The first official RoboCup@Home simulation competition was held in RoboCup Japan Open in May 2013. Since our team, the eR@sers, took first prize at the RoboCup@Home competitions in 2008 and 2010, our proposed use of simulation for the competition was accepted by the Japanese RoboCup@Home community. In the Japan Open, Cleaan Up task was chosen as a competition task and 5 teams participated. Since autonomous judge system was introduced at the competition, scoring procedure progressed smoothly.

The SIGVerse client application can be downloaded at http://www.sigverse.org. The source code for a sample application is available at[4]. The tutorial has more than 15 sample programs such as 1) controller for mobile robots, humanoid robots and human avatar, 2) sensor emulators such as distance sensor, and 3) usage of Kinect and head mounted display for immersive virtual reality system.

References

1. Sugiura, K., Iwahashi, N., Kawai, H., Nakamura, S.: Situated spoken dialogue with robots using active learning. Adv. Robotics 25(17), 2207–2232 (2011)
2. van der Zant, T., Iocchi, L.: Robocup @home: Adaptive benchmarking of robot bodies and minds. In: Mutlu, B., Bartneck, C., Ham, J., Evers, V., Kanda, T. (eds.) ICSR 2011. LNCS (LNAI), vol. 7072, pp. 214–225. Springer, Heidelberg (2011)

[4] http://www.sociointelligenesis.org/SIGVerse/index.php?Tutorial

3. Inamura, T., et al.: Simulator platform that enables social interaction simulation –SIGVerse: SocioIntelliGenesis simulator–. In: IEEE/SICE Int'l Symp. on System Integration, pp. 212–217 (2010)
4. van der Kapri, A., Ullrich, S., Brandherm, B., Prendinger, H.: Global lab: An interaction, simulation, and experimentation platform based on "second life" and "opensimulator". In: Proc. Pacific-Rim Symp. on Image and Video Technology (2009)
5. Johansson, M., Verhagen, H.: Massively multiple online role playing games as normative multiagent systems. In: Normative Multi-Agent Systems. Dagstuhl Seminar Proceedings, vol. 09121 (2009)
6. Robocup @home rules & regulations, Version, Revision 286:288 (2012), http://www.ais.uni-bonn.de/~holz/2012_rulebook.pdf
7. Bollini, M., Tellex, S., Thompson, T., Roy, N., Rus, D.: Interpreting and executing recipes with a cooking robot. In: Int'l Symp. on Experimental Robotics (2012)