HUMAN-ROBOT COOPERATION SYSTEM USING SHARED CYBER SPACE THAT CONNECTS TO REAL WORLD

Development of SocioIntelliGenesis Simulator SIGVerse toward HRI

Tetsunari Inamura[1,2]
[1]*National Institute of Informatics, Tokyo, Japan*
[2]*The Graduate University for Advanced Studies, Hayama, Japan*

Keywords: Intelligent robots, Human-robot interaction.

Abstract: I focus on a synthetic research on elucidation of genesis of social intelligence – physical interaction between body and environment, social interaction between agents and role of evolution and so on –, with aiming to understand intelligence of humans and robots. For such an approach, we have set interdisciplinal discussions with wide viewpoint for various research field such as cognitive science, developmental psycology, brain science, evolutionary biology and robotics. In this approach, two interactions should be considered; physical interaction between agents and environments and social interaction between agents. However there is no integrated system with dynamics, perception, and social communication simulations. In this paper, I propose such a simulation environment called SIGVerse and potential to develop Human Robot Interaction systems that bridges real environment/users and cyber space based on the SIGVerse. As examples of the HRI systems based on SIGVerse, I introduce three applications.

1 INTRODUCTION

Understanding the mechanism of intelligence in human beings and animals is one of the most important approaches to developing intelligent robot systems. Since the mechanisms of such real-life intelligent systems are so complex, such as the physical interactions between agents and their environment and the social interactions between agents, comprehension and knowledge in many peripheral fields are required. To acquire a better understanding of human and robotic intelligence, I focus on a synthetic approach to research into the elucidation of the genesis of social intelligence, to cover aspects such as physical interactions between bodies and their environments, social interactions between agents, and the role of evolution. According to the concept, SIGVerse(Inamura, 2010) the simulateor platform to realize synthetic simulation experiments has been proposed.

The SIGVerse simulator can be accessed from any client computers by general public with easy interface. Everyone can send their own physical agents that have sensors, actuators and software modules for behavior decision. The agents automatically decide their behavior based on the software and acquired sensor information; then act in the virtual world with consideration of physical law, because the physical embodiment is recently regarded as important issue to develop intelligent robots and agents. Every agent can also make communicate with each other via voice and text modalities. Using such an environment, it is possible to hold interdisciplinary discussions from wide viewpoint covering various research fields, such as cognitive science, developmental psychology, brain science, evolutionary biology, and robotics.

In this paper, I propose an expansion usage of the SIGVerse simulator to connect real world and cyber world to evaluate human robot interaction experiments. Robotics research which uses expensive humanoid robots often costs much time and budget; however, many users can join human-robot interaction world using the SIGVerse simulator. Through several examples of interactive application, I show the feasibility of this system.

Inamura T..
HUMAN-ROBOT COOPERATION SYSTEM USING SHARED CYBER SPACE THAT CONNECTS TO REAL WORLD - Development of SocioIntelliGenesis Simulator SIGVerse toward HRI.
DOI: 10.5220/0003620904290434
In *Proceedings of 1st International Conference on Simulation and Modeling Methodologies, Technologies and Applications* (SIMULTECH-2011), pages 429-434
ISBN: 978-989-8425-78-2

2 INTEGRATED SIMULATION BETWEEN DYNAMICS, PERCEPTION AND COMMUNICATION

Many robot simulation systems are being developed around the world, to simulate the dynamics of robot systems. One well-known example is the Open Architecture Humanoid Robotics Platform (OpenHRP)(Kanehiro et al., 2002), which is a humanoid robot simulator developed by the Advanced Institute of Science and Technology (AIST), the University of Tokyo, and the Manufacturing Science and Technology Center (MSTC). This simulator has become popular not only in Japan but also abroad, to promote research into humanoid robot control. The latest version of the simulation, called OpenHRP3(Nakaoka et al., 3647), is currently being developed by AIST, the University of Tokyo, and General Robotix Inc. Another example is Webots(Michel, 2004), which is a commercial product. This system enables users to simulate multi-robot environments with dynamics calculations. The Player/Stage/Gazebo suite(Gerkey et al., 2003) is freeware and is also well-known. Microsoft also has released *Robotics Studio*(Jackson, 2007) to develop a software of autonomous agents from the similar background. These systems provide multi-agent environments with dynamics simulations, but there hasn't been much consideration of the simulation of sensor perceptions. Since the communication simulations between agents provided in these packages are just simple signal transfers, it is difficult to use them to simulate the effects of the physical conditions of the agents for dialogue-based communication abilities and qualities.

Meanwhile, large-scale multi-agent systems are gaining attention from the social sciences field. One of the examples is GPGSiM(Ishiguro, 2007). In the field of language evolution, a system that has been proposed simulates language transmission between agents which is based on a repeatable learning model(Kirby and Hurford, 2002), However, such simulators do not consider the physical perception layer such as visual and auditory sensors, nor the physical communication layer such as limitations of communications based on the condition of each agent. The integration of dynamics, perception, and communication in the simulation world will play a great role in this social sciences field.

The proposed SIGVerse simulator(Inamura, 2010) combines dynamics, physical perceptions, and communications for a multi-agent system. One of the effective applications of this system is that machine learning experiments for real robots. Machine learning systems such as reinforcement learning for real robots often cost much time and trials; however it is easy to reduce the learning time to attend the SIGVerse simulator to make communicate with other expert robots which has already acquired a skill of the target problem. With the communication between two robots and fast physical experiments, the rookie robot can grow up within short time.

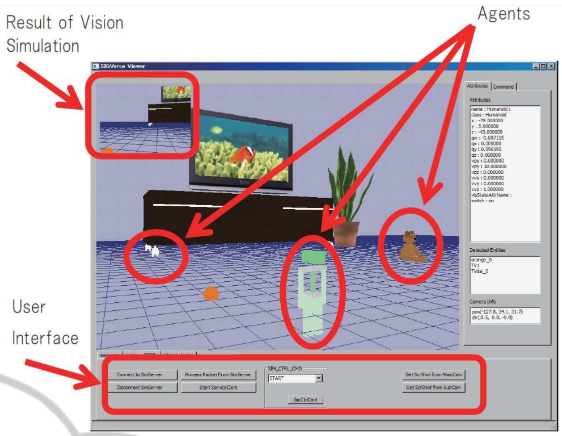

Figure 1: A screenshot of user interface.

This simulator consits of major three parts; Dynamics simulation, Perception simulation and Communication simulation.

2.1 Dynamics Simulation

The Open Dynamics Engine (ODE)(Smith, 2004) is used for dynamic simulations of interactions between agents and objects. Basically, the motions of each agent and object are calculated by the dynamics engine, but the user can control 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.

2.2 Perception Simulation

This system can provide the senses of vision, sound, force, and touch. OpenGL is used for visual simulations, to provide each agent with a pixel map that is a visual image derived from the viewpoint and field of view of that agent. In this case, the perception simulation has several levels that control the abstract level of perception. At a highly abstract perception level, the user is sent symbolized visual information, which comprises data such as the color, shape, size, and position of each object within sight, together with

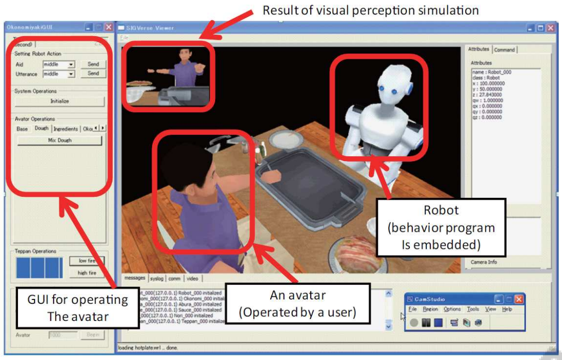

Figure 2: GUI interface for a cooperation task on SIGVerse.

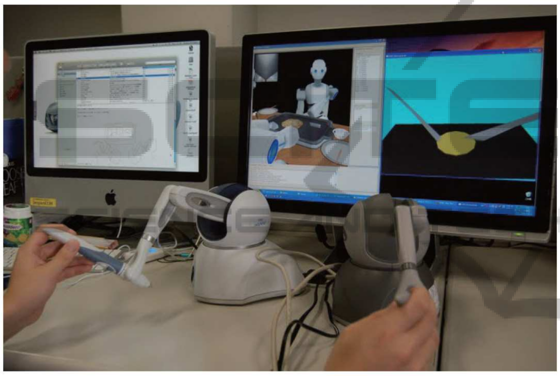

Figure 3: A haptic interface to handle an “okonomiyaki”.

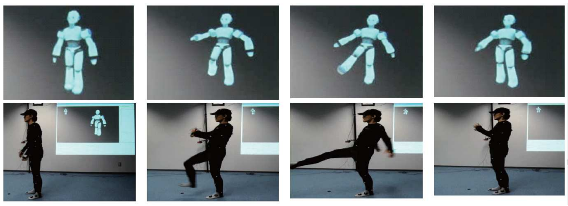

Figure 4: Realtime measurement and display of human motion pattern.

characteristic information on the name and ID number of each object. The visual perception processing also considers occlusions so that if an object is behind another object, the perception processing omits the hidden object from the list.

For the sense of touch, it is possible to acquire force and torque information between objects that has been calculated mainly by ODE.

2.3 Communication Simulation

In simulating the sense of hearing, every agents can make communications with audio data. I also enabled the effect of the volume of sound attenuating in inverse relationship with the square of distance, based on a setting of a condition that the voice emitted by an agent becomes more difficult to hear with distance. It is also possible to set the system so that only voices within a certain threshold distance are acquired.

Figure 5: Embodied interaction system for a okonomiyaki task on SIGVerse.

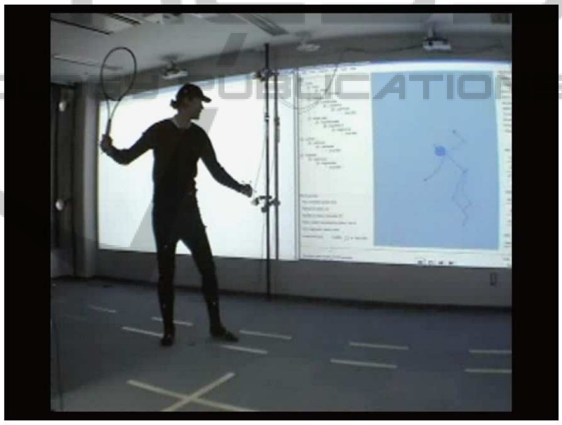

Figure 6: Experiment on motion coaching application using the SIGVerse.

With this system, not only can agents within the virtual environment communicate with each other, it is also possible to provide a function that enables interactions between the virtual environment and users in the real world. An example of the display of the virtual environment is shown in Figure 2.

2.4 Simulator Configuration

This simulator has a server/client format, with dynamics calculations being mainly done on a central server system. Bodies that use perception and perform actions are called ”agents”, and robot and human avatars are available as agents. The previously described perception and communication functions can be enabled by using dedicated C++ APIs to define the actions of agents. Some of the APIs that can be used are listed in Table 1. In the future, I plan to extend the programming beyond just C++ to include interpreter languages such as Python. The avatars do not just behave as programmed—they can also act on the basis of instructions given to them by operators in real time.

To simulate perception, it is necessary to spread the load so the system is configured to enable calculations not just by the server system but also by individually installed perception simulation servers. More specifically, the module that provides a pixel map of an image to simulate the sense of sight is operated by the perception simulation server, not the central server.

The configuration of the SIGVerse software is shown in Figure 3.

3 EXAMPLE OF SIGVERSE USE

A feature of SIGVerse is the way in which dynamic calculations, perception simulations, and communication simulations can be done simultaneously. In this section, I describe an example of humans and robots working in partnership to execute a task, and another example of multi-agent system, as examples of applications that fully utilize all three of these functions.

3.1 Use as Evaluation of Human-machine Cooperation

The objectives of the developers who use this simulation are to determine how to develop the intelligence of a robot that can execute a task in partnership with a human, and how to implement efficient cooperative behavior. The developers created decision and action modules for the robot while adopting various different models and hypotheses, and have confirmed their performance on the simulator. During the simulation, cooperation is required between a real-life human and a robot, which cannot be implemented otherwise without purchasing and developing a life-size humanoid robot. In this simulation, the operator who is in partnership with the robot manipulates an avatar in a virtual environment to reproduce cooperative actions between a user and a humanoid robot. An intelligence module created by the developers uses virtual equivalents of the senses of sight and hearing to comprehend the situation within that space and recognize the state of the user, performs dynamic calculations to control arms, and also simulates communications between the avatar and the robot. Expanding on this kind of usage example will not only further research into simple human-machine cooperation, it will also enable the construction of a research and teaching system with a competitive base for applications such as Robocup(Kitano et al., 1998).

Taking the above application as an example, I implemented a situation in which a human being and a robot cooperate in the task of ”cooking okonomiyaki” in SIGVerse (”okonomiyaki” is a popular cook-at-the-table food in Japan, like a thick pancake). Examples of the screens during the execution of this application are shown in Figure 2. The GUI that the operator can use has buttons such as ”flip the pancake”, ”oil the hotplate”, ”apply sauce”, and ”adjust the heat”.

Furthermore, providing immersive interface to the users is very important to conduct realistic phycophisical experiments throught the simulator. Fig.3 shows an example in which the user can operate the cooking devices with haptic interface *PHANToM Omni* to manipulate the “okonomiyaki”.

The objective of the task is to cooperate in cooking the okonomiyaki as fast as possible, without burning it. The operator basically uses the GUI to propel the work forward, but the robot continuously judges the current situation and, if it considers it can do something in parallel with the work that the operator is doing, asks the operator questions such as ”Should I oil the hotplate now?” or ”Should I turn the heat down?” It then executes those jobs while viewing the operator’s responses. Figure 2 shows a scene in which the avatar in the virtual environment is about to flip the pancake based on the operator’s instructions with a help of robot agent.

I performed experiments on two cases: one in which the operator performed all of the steps through the GUI, and one in which the robot did suitable parts of the operator’s work instead. In the first case in which the operator did all of the work, the task required three minutes 14 seconds before it was finished, but in the second case involving cooperation, the task took one minute 58 seconds to complete. In this manner, it is possible to make effective use of this system as a tool for quantitatively evaluating the human-machine cooperation systems.

3.2 Introduction of Immersive Interaction Space for the SIGVerse

Above applications used an usual display interface such as web browser worked on personal computer. However, if the aim of the application is to treat natural and real motion patterns of whole body that should connect real world and cyber world, interface devices

Table 1: Lists of available API functions.

setJointAngle (*arg*1, *arg*2)	Set the angle of joint *arg*1 be *arg*2
setJointTorque (*arg*1, *arg*2)	Set the torque of joint *arg*1 be *arg*2
getPosition	Get 3D position of the target object
getRawSound	Get audio information cast by other agents
sendRawSound	Utter speech as sound information
sendText(*text*, *distance*)	Cast a text message *text* to agents who are existing within a distance *distance*
captureView	Get pixel map as a visual sensor from eyesight of an agent
detectEntities(*arg*)	List up all of the agents which is seen by an agent *arg*
getObj(*arg*)	Get ID number of an object *arg*
getObjAttribution(*arg*1, *arg*2)	Get attribution value for attribute *arg*2 of an aobject *arg*1

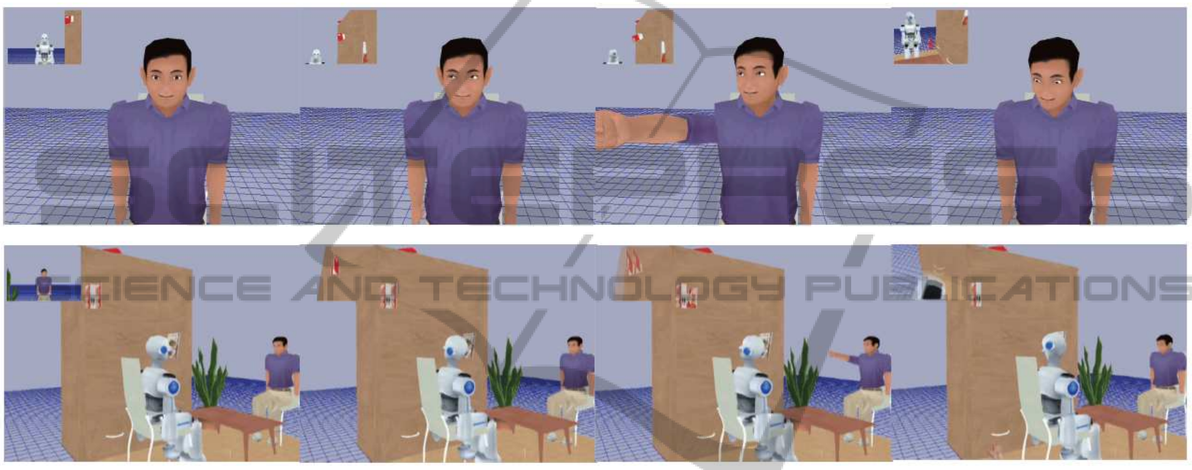

Figure 7: Experiment on joint attention between human and robot. The robot observes direction of the user's eyes. Using joint attention, the robot could avoid the falling object from a shelf.

would be a bottleneck. To solve this problem, an immersive interaction system named SD-GRIS(Kwon and Inamura, 2010) was developed and integrated to the SIGVerse system. The SD-GRIS can project all-around view of SIGVerse world to surrounding displays. Additionally, an optical motion capturing system is installed in the display space. User can interact with virtual agents with gestures and such as intuitive instruction system with mobile robot agents(Kwon and Inamura, 2010) as shown in Fig.5, that is previous Okonomiyaki application.

As an application of human-robot interaction using the immersive system is robotic motion coaching system(Okuno and Inamura, 2010), in which a robot trains human beings to be able to perform good sports motions even though the subjects are beginners. A trainer, that is a virtual humanoid robot, performs a target motion. A human as a subject imitates the performed motion; the virtual robot evaluate the human's motion as trainer. If the human's performance was far from the target motion, the robot modify the next performance according to the human's error to let the human's performance be biased to the target motion. Currently, a optical motion capturing system is used for the training application; however, simple and convenient device such as Kinect was connected to the system, training could be applied for general public who are connecting to the system from all over the world. An outlook of the motion coaching experiment is shown in Fig.6. This is another potential of the proposed system.

3.3 Joint Attention

Another significant function is to simulate direction of eyes of each agent. One of the important element in human-robot interaction is to recognize and control direction of eyes to establish joint attention for natural intaraction. Each agent in this system has a propety of eyes' direction. If a user ware eye tracker with HMD device, the information would be sent to the simulation system; the avator's direction of eyes is controlled by the real information. Since the surrounding scene image is displayed on the HMD, the user can behave if the user was standing just in front of the robot in the virtual SIGVerse world. Fig.7 shows a sequence of experiment on establishment of joint attention between virtual robot and user.

4 CONCLUSIONS

I have proposed the concept of a simulation platform in which dynamic simulations of bodies, simulations of senses, and simulations of social communications are integrated into the same system, as an approach to the interdisciplinary research necessary for comprehending the mechanism of intelligence in human beings and robots, and reported on the implementation of a prototype system named the SIGVerse(Inamura, 2010).

In this paper, an expansion usage of the SIGVerse was introduced to promote researches on human-robot interaction. Usual human-robot interaction experiments tend to work on only one user and one robot. However, it is important to perform wide and long term experiments with a lot of users and multiple robots to discuss social intelligence. Additionally, not only simple interface such as web browser, but also natural and rich interface such as whole body motion and eyesight are important to performe natural human-robot interaction. Through several functions and applications, I showed the feasibility of applying the SIGVerse to this grand challenge for human-robot interaction.

REFERENCES

Gerkey, B., Vaughan, R. T., and Howard, A. (2003). The player/stage project: Tools for multi-robot and distributed sensor systems. In *Proceedings of the 11th International Conference on Advanced Robotics*, page 317323.

Inamura, T. (2010). Simulator platform that enables social interaction simulation sigverse: Sociointelligenesis simulator. In *Proc. of IEEE/SICE International Symposium on System Integration*, page 212217.

Ishiguro, K. (2007). Trade liberalization and bureaupluralism in japan: Two-level game analysis. technical report 53. Technical report, Kobe University Economic Review.

Jackson, J. (2007). Microsoft robotics studio: A technical introduction. *IEEE Robotics & Automation Magazine*, 14(4):82–87.

Kanehiro, F., Fujiwara, K., Kajita, S., Yokoi, K., Kaneko, K., Hirukawa, H., Nakamura, Y., and Yamane, K. (2002). Open architecture humanoid robotics platform. In *Proceedings of the 2002 IEEE International Conference on Robotics & Automation*, page 2430.

Kirby, S. and Hurford, J. (2002). *Simulating the Evolution of Language*, chapter The emergence of linguistic structure: An overview of the iterated learning model., page 121148. Springer Verlag.

Kitano, H., Asada, M., Kuniyoshi, Y., Noda, I., Osawai, E., and Matsubara, H. (1998). *Lecture Notes in Computer Science*, volume 1395, chapter Robocup: A challenge problem for ai and robotics., page 119. Springer.

Kwon, O. and Inamura, T. (2010). Surrounding display and gesture based robot interaction space to enhance user perception for teleoperated robots. In *Intl Conf. on Advanced Mechatronics*, page 277282.

Michel, O. (2004). Webots: Professional mobile robot simulation. *International Journal of Advanced Robotic Systems*, 1(1):394.

Nakaoka, S., Hattori, S., Kanehiro, F., Kajita, S., and Hirukawa, H. (36413647). Constraint-based dynamics simulator for humanoid robots with shock absorbing mechanisms. In *Proceedings of the 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*.

Okuno, K. and Inamura, T. (2010). A research on motion coaching for human beings by a robotics system that uses emphatic motion and verbal attention. In *Humanoids 2010 Workshop on Humanoid Robots Learning from Human Interaction*.

Smith, R. (2004). Open dynamics engine v0.5 user guide. http://www.ode.org/.