Immersion type virtual environment for human-robot interaction[*]

Tadashi Odashima
Bio-Mimetic Control Research Center
RIKEN
Nagoya, Japan
odashima@bmc.riken.go.jp

Masaki Onishi
Bio-Mimetic Control Research Center
RIKEN
Nagoya, Japan
onishi@bmc.riken.go.jp

Zhiwei Luo
Bio-Mimetic Control Research Center
RIKEN
Nagoya, Japan
luo@bmc.riken.go.jp

Shigeyuki Hosoe
Bio-Mimetic Control Research Center RIKEN,
Nagoya University
Nagoya, Japan
hosoe@nuem.nagoya-u.ac.jp

Abstract − *With the development of information science and robotic technology, it becomes more important to generate human interactive robots. The design platform for developing such robots should satisfy three basic conditions: (1) It can test safely the performance of the robot through the physical interaction with human, (2) Human subject can estimate subjectively the outside appearance of the robot, and (3) It can simulate the dynamic human interactive robot motion within real-time. This paper proposes our immersion type dynamic simulation platform. An application to estimate the robot's performance when performing cooperative object lifting task with human subject is chosen in order to show the effectiveness of our system. The analysis of the recorded data is useful to design the novel human interactive robots.*

Keywords: Immersion projection technology, 3D dynamic simulator, Subjective estimation, Human interactive robot.

1 Introduction

With the development of information science and robotic technology, it becomes more important to generate human interactive robots. One of the most desired specifications for the robot is to cooperate with human in general environment such as streets or homes. The final goal of our research is to construct such a robot which can interact physically with human. We call it a "human cooperative robot" in this paper. There are many different aspects between the traditional industrial robots and the human cooperative robot. One of

them is the basic requirement of the robots. It is required for the industrial robot to perform operations with high speed and high accuracy, whereas the high safety and affinity are more important for the cooperative robot. The criteria for the cooperative robot tend to be more subjective than that for the industrial robot. Another difference is the evaluation method. The situation where the cooperative robot is evaluated must also include human. This fact derives difficulties for evaluating the cooperative robot. The experiment using directly the real cooperative robot is very dangerous, especially, when the new algorithm or mechanical system is applied. In this sense, it is considered that computer simulation is the best way for the evaluation. However to build the precise human model in the simulator is much more difficult. Even if the model can be constructed, it is hard to evaluate the impressions received from the robot. It is also important for the human subject to feel the robot just as if the human subject and the robot exist in the same space.

Considering above aspects, novel design platform is necessary for developing the final human cooperative robot. The basic conditions for the environment are: 1) It is possible to interact with the robot safely. 2) The human subject can experience the robot with high reality for subjective evaluation, 3) It can simulate the dynamic robot motion within real-time. By now, several robot simulators are proposed [3, 4, 6]. In reference [6], a simulator satisfying conditions 1) and 3) is proposed. The simulator can also provide a reaction force from the virtual space by haptic interface, but the visual image is displayed just on a usual monitor as 2-dimensional image.
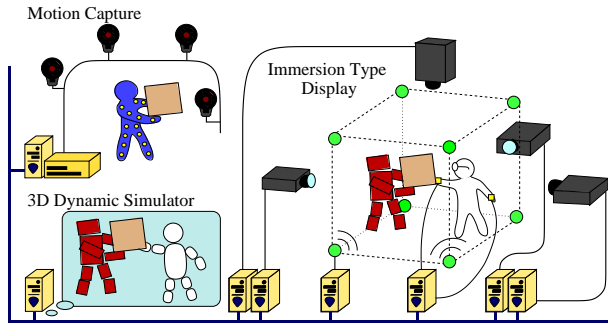
Figure 1: System overview

In this research, we construct an effective develop environment that satisfies all three conditions by combining the virtual reality techniques and 3-dimensional dynamic simulator[5]. In this paper, we express our system setup in detail and show the experimental results.

2   Outline of the system

As shown in Figure 1,the robot simulator which we propose consists of three parts. They are: (1) An immersion type display, (2) A 3-dimensional dynamic simulator, and (3) A motion capturing part, respectively. Each part consists of two or more PC computers, and all computers can communicate with each other via TCP/IP.

The immersion type display part presents human with audio-visual information and tactile sensation, which makes the human subject feel as if he/she is interacting with the virtual robot directly. For vision display, it is necessary to give the subject the feeling as if he/she is facing an actual robot. For audition, on the other hand, the location of the sound source such as the robot's footstep and voice should be set so that the subject can hear where the sound actually comes from. In addition, when a subject touches the robot, it is necessary to display the force, although it has not realized and is under research now.

The 3-dimensional dynamic simulator part controls a virtual robot to follow the given set of target angles while interacting with human subject. Here, it is necessary for the 3-dimensional dynamic simulator part to consider the external force applied to the robot from human as well as the gravity in its dynamic calculation. As a result of this calculation, the instruction torque is calculated based on the target angle and the torque added by the subject. The position and orientation of each link are then calculated in real time for every sampling period.

The robot's form takes as a humanoid-type, its target joint angle is measured from the motion of an actual human subject by using the motion capture system.
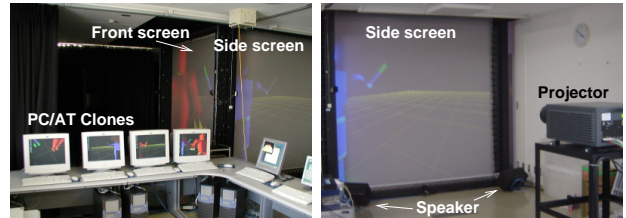

Figure 2: Backside of the PC-CAVE

3   Immersion type display

As shown in the Figure 2, the immersion type display part presents the human subjects with high presence of the virtual robot. The visual and auditory information playing an important part for perception are installed in this system. In addition, this part measures the motion of a human subject in order to reflect his/her motion into the virtual space.

3.1   Displaying visual information

Immersive Projection Technology (IPT) uses parallel image display system to show the 3-dimensional stereography. The displayed images are projected on the screens arranged as the wall surrounding the human subject. The high resolution of image and a large visual field angle provide a high feeling of immersion to human. Some types of IPT have been reported in the past. CAVE (CAVE Automatic Virtual Environment) [1] has four screens (front, right, left, and platform screen), and the basic projection structure of our immersion type display is the same. Immersion type display processes and renews the images of a virtual space by sensing the position and orientation of the human subject's head with magnetic sensor. The processed images are projected onto 4 screens simultaneously so as to be observed as the 3D image from the human subject.

The conventional IPT is constituted with the expensive supercomputer, whereas we use only four PC/AT-clone computers, which reduces the total cost of IPT and realizes the easy use without special skill. From our experiments, it is found that, although the calculation capability of a PC/AT-clone computer is inferior to an expensive supercomputer, it is possible to cover the lack of performance by reducing the complexity of the displayed contents. Moreover, considering the rapid increase of the calculation and/or drawing capability of PC/AT-clone computer and ease of maintenance, this setup can be used to realize our research objective. In the following paper, we simply call this projection system "PC-CAVE".

3.2   Displaying auditory information

The display of auditory information is realized using Huron of Lake Technology. Huron is the virtual reality technology that can generate the virtual sound space
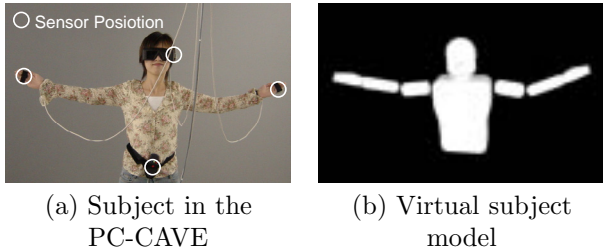
(a) Subject in the
PC-CAVE

(b) Virtual subject
model

Figure 3: Subject and virtual subject

15 D.O.F.

Figure 4: Virtual robot

with high reality by eight speakers array. Each speaker is set at every corner of the PC-CAVE structure. The applications of the Huron here are to play the sound from a virtual robot, for example, the utterance made by speech synthesis, the sound of footsteps and the sound with tumbling down.

3.3 Sensing subject's motion

It is necessary to send the subject's motion to the 3D dynamic simulator in order to realize the interaction between a human subject and the virtual robot. Therefore immersion type display performs not only presenting visual-audio information but also measuring the subject's motion simultaneously. To achieve accuracy simulation results, it is better to use sensors as many as possible for sensing the subject's position and orientation precisely. The increase of the sensors causes the enlargement of the system scale as well as the calculation load, which then leads to the serious problem in preventing the subject from receiving visual information with high presence. Therefore we limit the number of magnetic sensors measuring the position and orientation up to four. These sensors are attached on the subject's head, two wrists and body as shown in Figure.3(a). The visual image presented in immersion type display is refreshed based on the information comes from the sensor attached on head. The measured data from other sensors are sent to 3D dynamic simulator for controlling a virtual subject.

4 3D dynamic simulator

The 3-dimensional dynamic simulator calculates the position and orientation of every part of the virtual robot. This dynamic simulator is coded by applying the programming library Vortex of CM-Labs Simulations Inc.. Vortex has functions to process the collision check between the links and the dynamic motion calculations of each link in real time. The calculated position and orientation of all links is transmitted to the immersion type display part using TCP/IP. Owing to the high effectiveness of the processing routine, the time delay for rendering on the immersion type display part is very small. The measurement data of two human motions are gathered into this simulator for controlling virtual
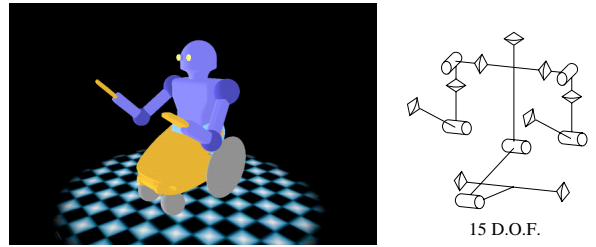
objects. One is the motion of the human subject in the immersion type display mentioned in **3.3**. This data is used for controlling the virtual subject. Another is the motion captured in motion capture system described in **5**, which is used for deciding the target joint angle of the virtual robot.

4.1 Virtual subject

A virtual model of the human subject in PC-CAVE is build in virtual space, so that he/she interacts with a virtual robot. In the virtual space, he/she can add force to virtual object by controlling the virtual subject. Since we limit the interaction on only the upper half of the subject's body, as shown in Figure.3(b), the virtual subject is constructed from a torso including head, upper arms, and forearms including hands. Each Link is constrained with next by free ball-socket joint. It is better to construct a detailed model copied human structure and to control it by the captured human motion in order to acquire accurate simulation results. As mentioned before, however, we can use only four magnetic sensor, the torso and forearms are controlled directory, and the upper arms are moved indirectly by the mechanical constrains. The length and size of every link is set based on the size of human subject.

4.2 Virtual robot

Since it is important for the robot interacting with human to have a high personal affinity design, the robot's form is better to be a humanoid. The main objective of our robot research here is not focused on the legged locomotion, but on the cooperation with human. As shown in Figure.4, the wheel is used for simplifying the locomotion mechanism, and only the upper half of the body is designed as humanoid. The virtual robot has totally 15 d.o.f., including wheels' rotation. Each arm part has 5 d.o.f., 3 at the shoulder, 1 at the elbow and 1 at the wrist. The neck, waist and knee has 1 d.o.f., respectively. The size of every part is designed based on the human parameters.

This simulator not only enables us to estimate the outside appearance of the robot subjectively but also makes it possible to evaluate the specification of the robot numerically. Using the real motor parameters for

(a) Real human movements

(b) Results of motion capture

(c) 3D dynamic simulation

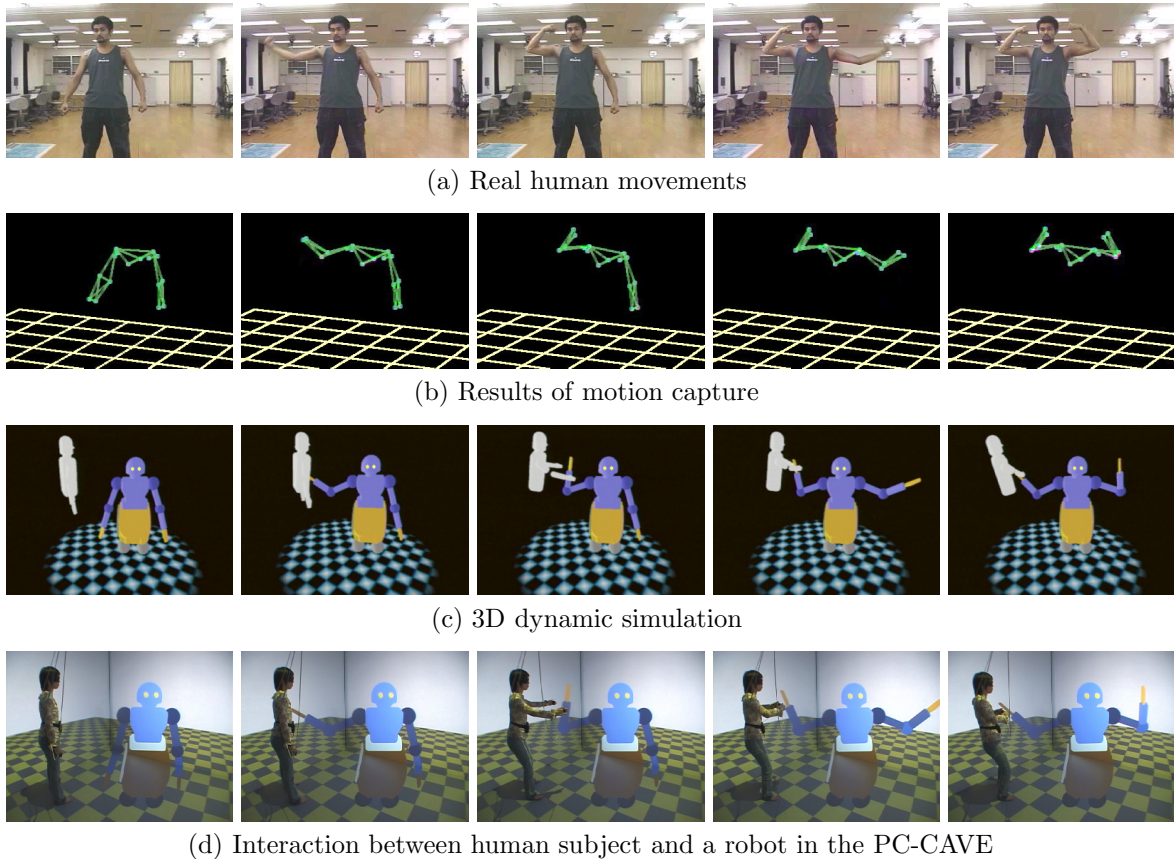(d) Interaction between human subject and a robot in the PC-CAVE

Figure 5: Experimental scene

controlling the virtual robot motion may help the selection of actuators for designing the real robot. In addition, many kinds of information about the robot can be measured easily in the simulator. For example, the force and torque at any contact point is measured by setting virtual sensor function, and the visual data from any point is obtained by setting the camera function. These functions are useful for the later analysis of the robot motion. The simulation results about the measured data are presented in **6**.

The joint angles of the virtual robot are controlled by PD control, and the desired angle is decided from the motion data captured in the motion capture system. The simulation result consists of the data of position, orientation and index label of all links except the links constructing the virtual subject, and is sent to the immersion type display for presenting 3D image.

5   Motion capture

Considering the affinity of human interactive robot, we selected the humanoid design for our robot and built the motion capture part for getting the human motion in order to acquire the desired angle for the virtual robot.

In our system, the capture data is used to control the virtual robot in the simulator. Here, VICON with six high-speed infrared cameras is used to observe the markers attached on the human body with the sampling period of 120Hz. From the precise point of view, it is desired to attach markers on the whole body of the subject, however, since the biped walk is not easy, in this research, we focus only on the motion of upper body part, especially the movement of both arms. In order to capture the motion of both arms, 16 markers are attached on specific points between the shoulder and hand of both arms. These makers cover totally the motion up to 10 d.o.f., that is, 3 d.o.f. of each shoulder, 1 d.o.f. of the elbow, and 1 d.o.f. of the wrist, respectively. The measured joint angles are sent to the 3-dimensional dynamic simulator part explained in **4** via TCP/IP and are used as the target joint angles of the virtual robot.

6   Experimental results

The experiments using the proposed robot simulator have been performed to show the effectiveness of our research.

6.1   Interaction in PC-CAVE

Figure 5 shows the results of an experiment using our system. Figure 5 (a) shows the motion of a human subject who was attached with 16 markers, and then 10
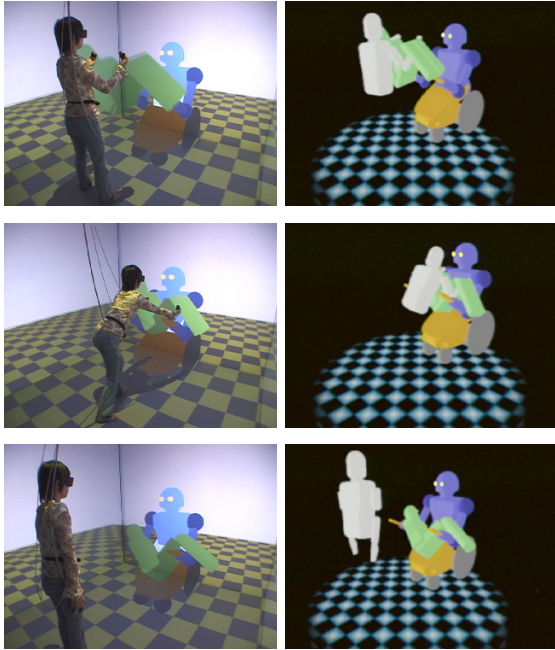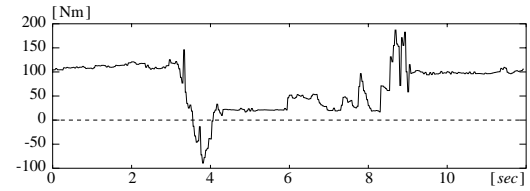
Figure 6: Example of whole body manipulation

(a) Torque of the shoulder joint

(b) Torque of the elbow joint
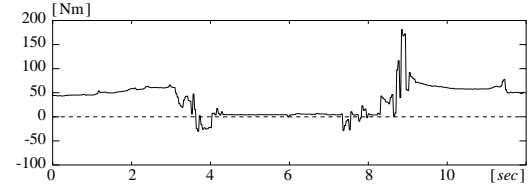
Figure 7: Torque of the each joint

joint angles of the upper half body of the subject are calculated from these markers' position data. (b) shows the calculated result of motion capturing by the VI-CON. The set of calculated angles are sent to the 3D dynamic simulator as a set of target angles. The virtual robot shown in (c) basically follows the target angles by PD control in 3D dynamic simulator, and the virtual subject (in (c) left) follows the human subject's motion measured in the immersion type display. The stereoscopic glasses are set on the side of the video camera when the images in (d) are recorded so as to keep the consistency of these images as seen from the human subject.

The subject in the immersion type display can recognize the real scale of the virtual robot and add forces by pushing or pulling some parts of the virtual robot. In Figure 5(d), the human subject tries to stop the swing of the virtual robot arm. It is clear that the added force works as the external force, and the force makes difference between the posture of robot arm and that of captured motion shown in (b). The robot falls down when a large force is added, and a tumbling sound is generated from the falling point by sound effect.

Figure 6 shows the experiment of a cooperative task to carry the human-formed object from the human subject to the virtual robot. Such kind of cooperation is considered as an essential task for the robot in the human care and welfare areas. The subject can get an impression against the robot through this task, and the subjective evaluation is important for designing the real robot's shape and motions. Of course, numerical data can also

be monitored in this simulator. Recently, our research team has also studied on the Whole Body Manipulation (WBM)[2] for human-robot cooperation. The numerical data derived from our system is useful for analyzing the algorithm of the WBM and for estimating the efficiency of the WBM.

6.2 Sensor simulator

In virtual space created by the 3D dynamic simulator, we can monitor easily any value by only implementing virtual sensory functions, even if it is difficult to monitor the values practically in real space. Currently, three types of virtual sensory functions have been developed. They are torque sensor, force sensor and cameras.

The reaction force at any collision point and/or the force and torque added at specific constraint is calculated in the 3D dynamic simulator for deciding motion of every object. The virtual force/torque sensor has functions to get these values.

Figure 7 shows the time series of the joint torque during the cooperative task mentioned in previous section (Figure.6). The weight of the human formed object is set to 47.5[kg] in this experiment. According to this result, a peak appears at the moment when the robot receive the object.

Responses of the force sensor function is shown in Figure.8. The human subject hits the virtual robot hand three times, twice from up side, and once from below. Each graph illustrates each direction corresponding to each axis in a coordinate system fixed on ground. Considering the upper direction is a positive on y axis, the hitting direction is clear according to (b).

The subject cannot feel reaction force from virtual space. The development of the force display device is our future work. The monitored force on the body of the virtual subject can be used as the target value for the device.
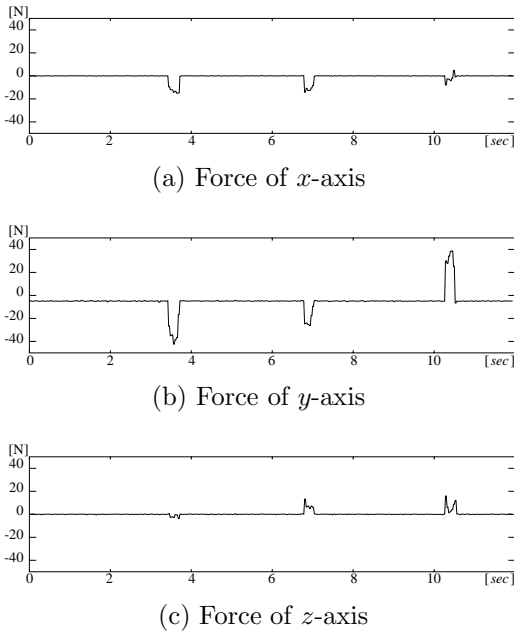
(a) Force of x-axis

(b) Force of y-axis

(c) Force of z-axis

Figure 8: Input data of force sensor

(a) Image of left eye

(b) Image of right eye

(c) Edge image

(d) Edge image

Figure 9: Input image of virtual visual sensors

Rendering the visual image from specific point in virtual space corresponds to set the camera in real space. The images input from the virtual robot's left and right eyes are shown in Figure 9 (b) and (c), respectively. We understand that the virtual robot is looking at the human formed object, the virtual human subject and its hands. Since the virtual subject image means the presence of the human subject, the virtual robot can detect the human subject in real space and makes an action to the human subject, which means the bilateral interaction could be realized. In order to detect the human subject, the virtual robot can select two ways. One is to get the information about the virtual subject directly from 3D dynamic simulator, and another is to process the image. The later is more realistic situation for applying to the real robot. Figure 9(d) and (e) show the edge detection as an example of image processing.

7 Conclusions

In this paper, we express the 3-dimensional dynamic simulation environment for evaluating the human cooperative robot which can interact physically with human. This environment consists of three parts, the immersion type display part, the 3-dimensional dynamic simulator part, and the motion capturing part. The subject can interact with a virtual robot, which follows a human motion, using advanced virtual reality technology. In this experiment, we confirmed the performance of the proposed system by monitoring the virtual robot motion under an external force from the subject. In addition, we developed the virtual sensor functions. These are a force sensor, torque sensor, and cameras.
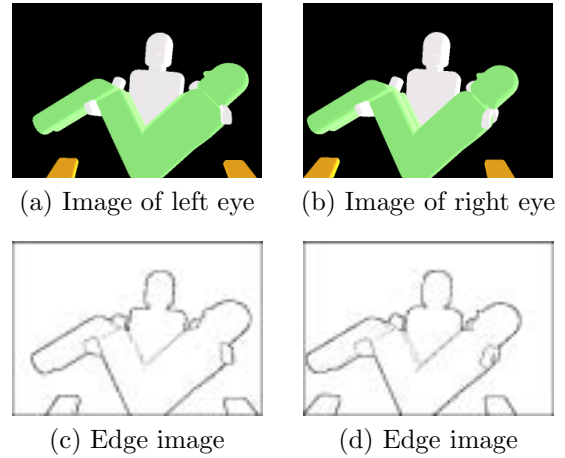
Currently, the development of force display device which represents reaction forces from virtual space to the human subject is not finished. It is our future research subject. Moreover, the measurement of joint torque of human motion capturing part is also under working. The torques are measured by implementing multi channels of EMG, 8 force plates as well as multi channels of accelerometer set in the motion capture part. It is our future subject for the virtual robot to mimic the skillful and dynamic movement of human subjects.

References

[1] C. Cruz-Neira, et al., "Surround-Screen Projection-Based Virtual Reality: The Design and Implementation of the CAVE," *Proc. SIGGRAPH* 93, pp.135–142, 1993.

[2] F. Asano, et al., "Dynamic Modeling and Control for Whole Body Manipulation," Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems(IROS), 2003, to appear.

[3] F. Gravez, et al., "Dynamic Simulation of a Humanoid Robot with Four DOFs Torso," *Proc. IEEE International Conference on Robotics and Automation (ICRA2002)*, vol.1, pp.511–516, May 2002.

[4] F. Kanehiro, et al., "Open Architecture Humanoid Robotics Platform," *Proc. IEEE International Conference on Robotics and Automation (ICRA2002)*, vol.1, pp.24–30, May 2002.

[5] M. Onishi, et al., "Development of PC-based 3D Dynamic Human Interactive Robot Simulator," *Proc. IEEE International Symposium Computational Intelligence in Robotics and Automation (CIRA)*, 2003, to appear.

[6] O. Khatib, et al., "Robotics and Interactive Simulation," *Communication of the ACM*, vol.45, no.3, pp.46–51, March 2002.