Operating Humanoid Robots in Human Environments

Neo Ee Sian, Takeshi Sakaguchi and Kazuhito Yokoi
JRL, AIST,
Central 2, 1-1-1 Umezono, Tsukuba 305-8568, Japan.
Email:{rio.neo, sakaguchi.t, Kazuhito.Yokoi}
@aist.go.jp

Yoshihiro Kawai and Kenichi Maruyama
VVV, AIST,
Central 2, 1-1-1 Umezono, Tsukuba 305-8568, Japan.
Email:{y.kawai, k.maruyama}@aist.go.jp

Abstract— **Depending on the autonomous capability of the robot and the familiarity of the robot system with the task and environment, the level of human intervention differs. This paper introduces a methodology which allows a human operator to seamlessly switch between the continuous control of motion using an analog input device and the discrete behavior control cooperating with the robot using symbolic commands. Using the proposed methods, a human operator is able to operate humanoid robots with high flexibility by only using a simple operation interface. Successful experiments operating humanoid robot HRP-2 in executing everyday tasks proved the high reliability of the proposed operation system.**

I. Introduction

In recent years, advances in both mechanical systems and software architectures have contributed to the realization of some promising examples of humanoid robots executing tasks aimed for space missions, maintenance of industrial plants, home management services, human care services, industrial vehicle operations and cooperative works with human[1][2][3][4]. Nevertheless, the realization of reliable autonomous humanoid robots which can perform tasks without human supervision is still limited by the current level of perceptual capabilities, decision making technologies and motion control strategies. With their feet not fixed to the ground and the substantial increase in the dimension of the configuration space, the successful introduction of humanoid robots into human environments will rely on the development of reliable and practical systems integrating motion generation, perception, knowledge management and decision making technologies[5][6][7][8][9][10].

Human intervention is a must in creating safe and useful humanoid robot systems. Effective and reliable methodologies for operating humanoid robots are of great importance. Depending on the autonomous capability of the robot and the familiarity of the robot system with the task and environment, the level of human intervention differs.

This paper introduces a methodology which allows a human operator to seamlessly switch between the continuous control of motion using an analog input device and the discrete behavior control cooperating with the robot using symbolic commands.

II. Concepts in Operating Humanoid Robots

Let us recall how we act in the environment. When we are acting in an environment which we are familiar with, most of the actions we take are carried out with our attention focusing on some discrete information like where to go, which object to manipulate, which behavior to take, and etc.. Our motions are then generated subconsciously to satisfy the selected goal. When we come to a new environment or perform a new set of motions, the conscious attention we pay to the motions increases. As we are not familiar with the environment or the actions that we are performing, we are consciously attending to the continuous control of the direction that we look, the way we walk, and the way we control our body.

The same applies to operating humanoid robots. When operating a robot which is equipped with a high level of autonomy serving in a known environment, a small number of high level commands will be sufficient in achieving the intended tasks. However, when the system is used in a new environment, a human operator will have to command the robot where to look at, what to look for and how to act. And for commanding new motions, the human operator will have to continuously command the trajectory of the body of the robot.

A. Two Levels of Operation

Depending on the autonomous capability of the robot and the familiarity of the robot system with the task and environment, the operator will need to switch between the following two levels of operation:

- Motion Level Operation:
  Continuous control of the most important point of the motion for example the hand, the head, the waist, the leg, the direction of walking and etc..
  Motion level operation is related to motion generation and control function of the operation system. The human operator cooperates with the robot in generating and controlling the robot's motion by guiding the most important point of the movement. The human operator operates the motion of the robot continuously by using an analog input device which converts human motor action into physical signals to the robot. Whole-body motions are

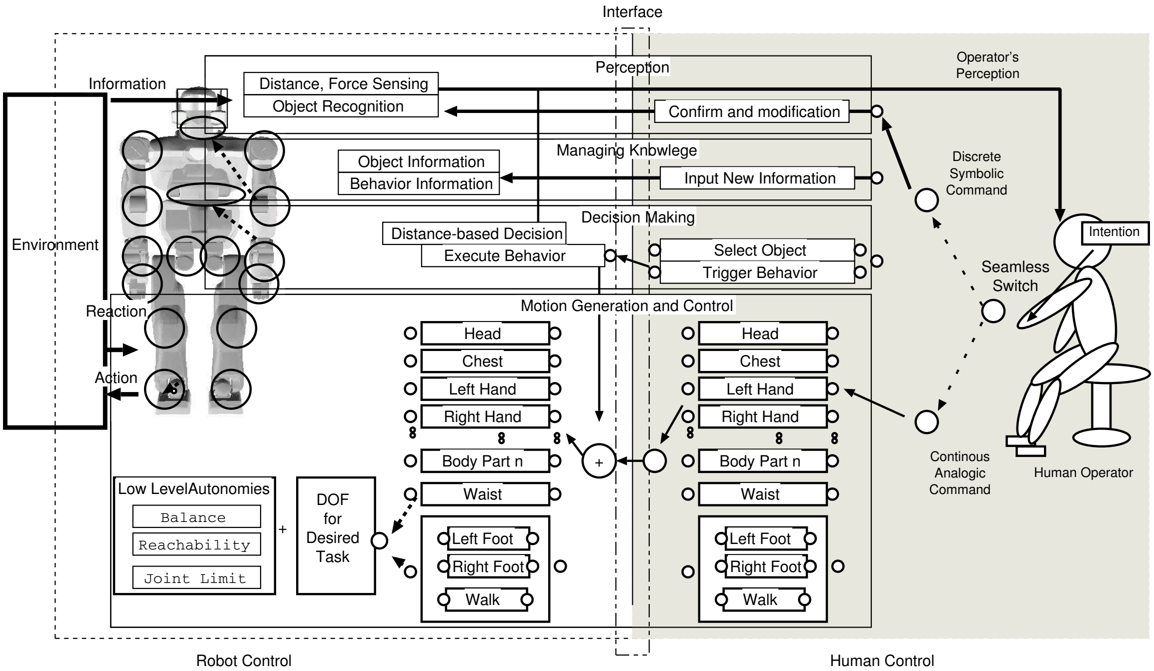

Fig. 1. Proposed System for Seamless Switching between Motion Level Operation and Behavior Level Operation

then generated integrating autonomous functions of the robot such as maintaining stability and extending reach. This method saves the operator from having to send commands to all joints of the robot. The operator can thus concentrate on executing commands only to the specific body parts without having to take care of the kinematical and dynamical constraints of the robot, such as reach limits and balance constraints.

- Behavior Level Operation:
  Discrete command such as "walk to A", "reach B using right hand", "pull C with left hand", "stand", "squat" and etc.
  Behavior level operation is done with both the operator and the robot cooperating in the perception, knowledge management and decision making functions. The operator selects objects, triggers actions, teaches new information to the robot and makes confirmation on the robot's perception using symbolic command interfaces such as a keyboard, a mouse or voice command. Shared Autonomy in behavior level operation is achieved with the cooperation of high level perception and cognition of human with the accuracy of low-level sensing capabilities of the robot. One manifestation of such kind of sharing is the human giving direction to where to look at and the robot locate the position of the target object using model-based vision. Another instance is the human indicating the target object to be grasped and picked up, whereupon the approach-grasp-and-pick operation would be carried out by the robot.

The operator will benefit if the operation system allows him to seamlessly move across both behavior level and motion level operations. Fig. 1 depicts our proposed system which enables the operator to do so. The seamless switching between the two levels of operation is realized by allowing the motion level command to be added to the command generated by behavior level operation.

III. WHOLE BODY MOTION GENERATION FRAMEWORK FOR HUMANOID ROBOTS

A humanoid robot can generally be modeled as a tree structure mechanism with five links attached to a free moving 6 degrees of freedom (DOF) body in space. We define body frame Σ_B as the frame fixed on the waist with linear and angular velocities ${}^W\boldsymbol{\xi}_B$. The leading superscript ^{W} indicates that the velocities are described using the world frame, which is the Cartesian frame fixed on the ground Σ_W.

A. Operational Points and Joint Utilization

Degrees of freedom necessary to realize the desired position and orientation of operational points during task executions are usually far less than the entire DOF of a humanoid robot. Here we divide the joints of the robot into Control joints and free joints. Control joints are the joints of the links of which the operational points are controlled in order to satisfy the desired tasks. Free joints are the joints of the free links which are not used for the desired tasks. With this categorization all joints of the robot can be described using

$$\dot{\boldsymbol{\theta}} = [\dot{\boldsymbol{\theta}}_{cl_1}^T \quad .... \quad \dot{\boldsymbol{\theta}}_{cl_n}^T \quad \dot{\boldsymbol{\theta}}_{fl}^T]^T, \tag{1}$$

where n denotes the number of operational points that are needed for the desired tasks, $\dot{\boldsymbol{\theta}}_{cl_i}$ denotes the vector for joint velocities of the respective control links and $\dot{\boldsymbol{\theta}}_{fl}$ denotes the vector for the joint velocities of free links. Please note that when n is equivalent to the total number of the links, $\dot{\boldsymbol{\theta}}_{fl}$ will become a null vector.

Fig.2 depicts the model of a humanoid robot and the joint utilization for the task of reaching using the left hand while standing, in which the operational points are both feet and the left hand. The target joint velocities for the respective control link from the body frame to the target operational frame, $\dot{\boldsymbol{\theta}}_{cl_i}^{trg}$, can be obtained by

$$\dot{\boldsymbol{\theta}}_{cl_i}^{trg} = J_{cl_i}^{\dagger} \left\{ {}^W\boldsymbol{\xi}_i^{ref} - \begin{pmatrix} E_3 & -{}^W\widehat{\boldsymbol{r}}_{B\to i} \\ 0 & E_3 \end{pmatrix} {}^W\boldsymbol{\xi}_B{}^{trg} \right\}. \quad (2)$$

where ${}^W\boldsymbol{\xi}_B^{ref}$ and ${}^W\boldsymbol{\xi}_i^{ref}$ denote the velocities of the body frame and desired velocities of the respective operational point frame. J_{cl_i} denotes the Jacobian matrixes calculated from the base frame to the respective operational point frame, ${}^W\boldsymbol{r}_{B\to i}$ denotes the position vectors from the base frame frame to the respective operational point frame, E_3 denotes a 3×3 identity matrix, and $\hat{}$ denotes an operator which translates a vector of 3×1 into a skew symmetric matrix 3×3 that is equivalent to an outer product.

B. Low Level Autonomies in Motion Generation

We have introduced the idea of integrating low level autonomies for safe operation of the robot in a sub-conscious fashion during motion generation which includes Balance Autonomy, Workspace Expansion Autonomy and Interactive Center of Mass(CoM) and Supporting Area Transition Autonomy[14].

Using a framework which generates whole body motions by controlling the total momentum of the robot [11], we calculate the velocities of the waist frame ${}^W\boldsymbol{\xi}_B^{trg}$ and the joint velocities of the free joints $\dot{\boldsymbol{\theta}}_{fl}^{trg}$ that realize both the reference velocities for the operational points ${}^W\boldsymbol{\xi}_i^{ref}$ and the reference momentum, ${}^W P^{ref}$, ${}^W L^{ref}$ as the least square solution by

$$\begin{bmatrix} {}^W\boldsymbol{\xi}_B^{trg} \\ \dot{\boldsymbol{\theta}}_{fl}^{trg} \end{bmatrix} = A^{\dagger} S \left\{ \begin{bmatrix} {}^W P^{ref} \\ {}^W L^{ref} \end{bmatrix} - \sum_{i=1}^{n} \begin{pmatrix} M_i \\ H_i \end{pmatrix} {}^W\boldsymbol{\xi}_i{}^{ref} \right\} + (E - A^{\dagger}A) \begin{bmatrix} {}^W\boldsymbol{\xi}_B{}^{ref} \\ \dot{\boldsymbol{\theta}}_{fl}^{ref} \end{bmatrix}. \quad (3)$$

Here,

$$A \equiv S \begin{pmatrix} M_B & M_{fl} \\ H_B & H_{fl} \end{pmatrix},$$

$$S \equiv \begin{bmatrix} \boldsymbol{e}_{S_1} & \dots & \boldsymbol{e}_{S_n} \end{bmatrix}^T,$$

$$\begin{pmatrix} M_i \\ H_i \end{pmatrix} \equiv \begin{pmatrix} M_{cl_i} \\ H_{cl_i} \end{pmatrix} J_{cl_i}^{-1},$$

$$\begin{pmatrix} M_B \\ H_B \end{pmatrix} \equiv \begin{pmatrix} \tilde{m}E_3 & -\tilde{m}{}^W\widehat{\boldsymbol{r}}_{B\to g} \\ 0 & \tilde{I} \end{pmatrix} - \sum_{i=1}^{n} \begin{pmatrix} M_i \\ H_i \end{pmatrix} \begin{pmatrix} E_3 & -{}^W\widehat{\boldsymbol{r}}_{B\to i} \\ 0 & E_3 \end{pmatrix},$$

where S denotes a $n\times 6$ matrix for the selection of the elements of the total linear and angular momentum for control, which consists of $\boldsymbol{e}_i$ denoting a 6×1 vector with parameter 1 for the activation of the selected i-th momentum and parameter 0 for the other elements of the vector. $A^{\dagger}$ is the pseudo-inverse of A and E is an identity matrix. ${}^W\boldsymbol{\xi}_B^{ref}$ denotes the adjustments of the velocities of the waist frame that can be made utilizing projection of the null space, depending on the selection of S. M_{cl_i}, H_{cl_i} denote the inertia matrix of which joint velocities of the respective control link affect the total linear and angular momentum of the robot. M_i, H_i denote the inertia matrix of which the velocities of the respective operational point affect the total linear and angular momentum of the robot. M_{fl}, H_{fl} denote the inertia matrix of which joint velocities of the free links affect the total linear and angular momentum of the robot. $\tilde{m}$ is the total mass of the robot, $\tilde{I}$ is the inertia tensor matrix around the center of mass(CoM), and $\boldsymbol{r}_{B\to g}$ is the vector from the origin of Σ_B to the CoM.

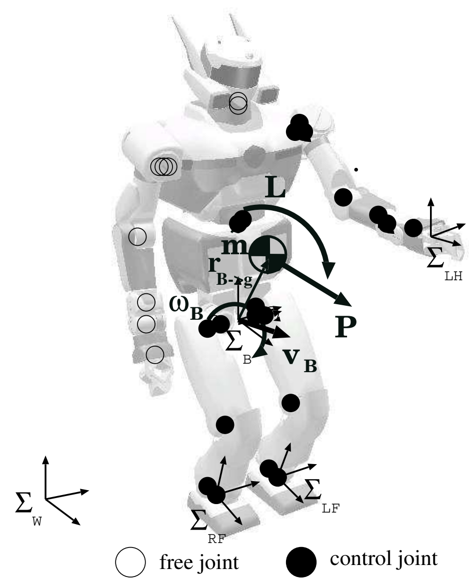

Fig. 2. Operational Points, Control Joints, Free Joints and Momentum During the Task of Reaching with Left Hand While Standing

Balance Autonomy is realized by controlling the position of CoM through manipulating the linear momentum P using

$$P^{ref} = \tilde{m}k(\boldsymbol{r}_{W\to g}^{ref} - \boldsymbol{r}_{W\to g}), \quad (4)$$

where $\boldsymbol{r}_{W\to g}^{ref}$ and $\boldsymbol{r}_{W\to g}$ denote the reference and estimated position for CoM and k denotes the gain of the control scheme. By controlling the CoM using Equation (4) so that the projection of CoM remains within the area of the support polygon will keep the static balance of the humanoid. Furthermore, by setting the reference value for angular momentum L^{ref} as zero or regulating the reference acceleration of the waist frame, the dynamic balance indicator ZMP can be controlled to remain within the support polygon. These values are controlled autonomously to allow the operator to only concentrate on manipulating the target points of the robot's body without having to take care of the robot's balance constraint.

IV. DESIGNING BEHAVIORS FOR OPERATING HUMANOID ROBOTS

We have constructed behaviors to facilitate online operations for humanoid robots under the following categories:

- Perceptual Behaviors:
  Behaviors involving the sensing and recognition of environment through visual, auditory, haptics sensing and etc..
- Motion Behaviors:
  - Object Oriented Motion Behaviors:
    Behaviors involving body movements which are object oriented, such as reaching, touching, tracking, pulling, holding and etc..
  - Non-Object Oriented Motion Behaviors:
    Body movements which are not object oriented such as squatting, standing and etc..

Motion Behaviors for humanoid robots are often difficult to design due to the multi-dimensionality of the mechanism. One way of deciding the behaviors is being carried out experimentally by using the online operation system we have developed [14]. The operator first operate the robot to perform the task both using simulator and on the real robot to check the feasibility of the motion, and find the suitable standing position, grasping point, and approach motions relative to the object for manipulation. These informations are utilized to design object-orientated behavior programs, and used as parameters for motion generation during behavior executions.

We have constructed several behaviors with the integration of online motion generation framework, 3D visual recognition functions and human-robot interaction. For tasks such as reaching to an object, vision functions are used to estimate the distance between the object and the robot, the system will judge whether the object is within the reachable space of the current supporting polygon and judge the necessity of walking. The walking distance is calculated by the target standing position predefined according to the task and the walking command is issued. During behavior level operations, the triggered behavior will generate reference velocities for the required operational points or control joints for every control loop used in Equations (2) and (3).

V. IMPLEMENTATION ON HUMANOID ROBOT HRP-2

The proposed system is implemented to operate humanoid robot HRP-2[12]. Humanoid HRP-2 is 1540 [mm] tall and weights 58 [kg] including batteries. It has 30 DOF. In order to construct a supervisory operation system for HRP-2, we have replaced the original black and white three video cameras of HRP-2 No.10 with four color video cameras (Fig. 3). Three of them are mounted with a narrow-angle (33.1 [deg.] x 25.0 [deg.]) lens for increasing accuracy of visual recognition and one is mounted with a wide-angle (93.6 [deg.] x 70.8 [deg.]) lens for teleoperation. We have also developed a new hand named JRL-hand for HRP-2. The JRL-hand is a one DOF parallel gripper. It allows humanoid HRP-2 to grasp a wider variety of objects, from a thin board to a 75 [mm] thick one and a cylindrical object 85 [mm] in diameter. The maximum grip strength is 14 [N] at the tip of the finger.

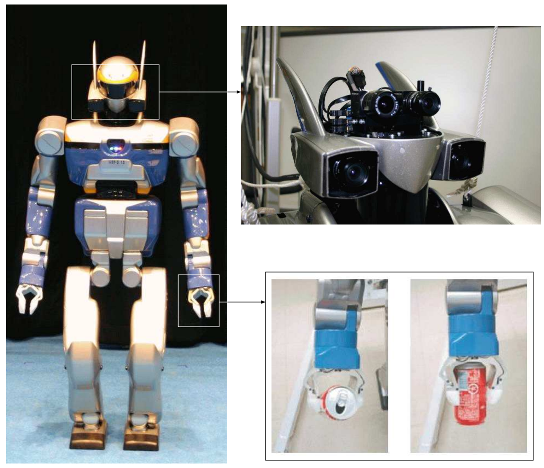

Fig. 3. HRP-2 No. 10 with Modified Camera System and Hands

A. Operation Interface

Fig. 4 depicts the operation interface. The operation interface is composed of a Graphical User Interface(GUI), two 3-DOF joysticks, a keyboard and a mouse. The GUI is composed of Command Input Box, Command Dialogue Box, Online Camera View, Online Robot Simulator, Pop-up Recognition Results, and Console Output.

For continuous motion level operations, the operator selects the operational points by pressing the joysticks' buttons allocated for the respective mode and input linear and angular velocity command of the operational points by manipulating the joysticks' levers. For walking operation control, the operator controls the walking direction, foot step distance, and the distance between the feet to generate walking patterns in real time. The operator can simultaneously control the position and orientation of the head and both hands during walking[14].

For behavior level operations, the operator gives discrete symbolic commands using keyboard and mouse while monitoring the robot's condition and the robot's perception of the environment using the GUI.

The seamless switching between motion and behavior level operations is made possible with the reference velocities for the operational points being calculated as the addition of the reference of both motion and behavior level operations within the real-time control loop as

$$\boldsymbol{\xi}_i^{ref} = \boldsymbol{\xi}_i^{ref_{motion}} + \boldsymbol{\xi}_i^{ref_{behavior}}. \quad (5)$$

Here $\boldsymbol{\xi}_i^{ref_{motion}}$ denotes the desired velocities of motion level operations input using joysticks, $\boldsymbol{\xi}_i^{ref_{behavior}}$ denotes the desired velocities generated by behaviors, and $\boldsymbol{\xi}_i^{ref}$ denotes the unified desired velocities of the operational point which is used in Equations (2) and (3).

B. Software System

The overview of the software system is shown in Fig. 6. The distributed server system consists of the following servers:

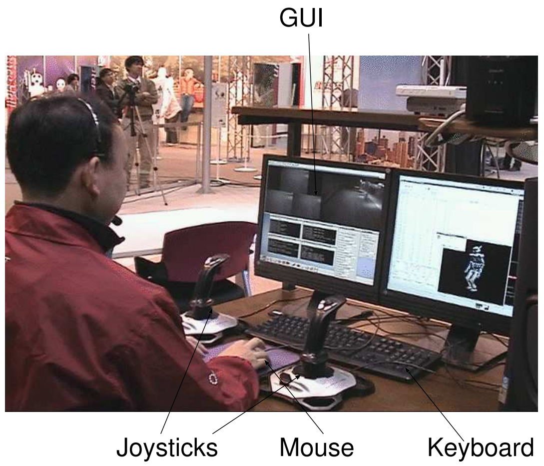

Fig. 4. Hardware System for HRP-2 Operation

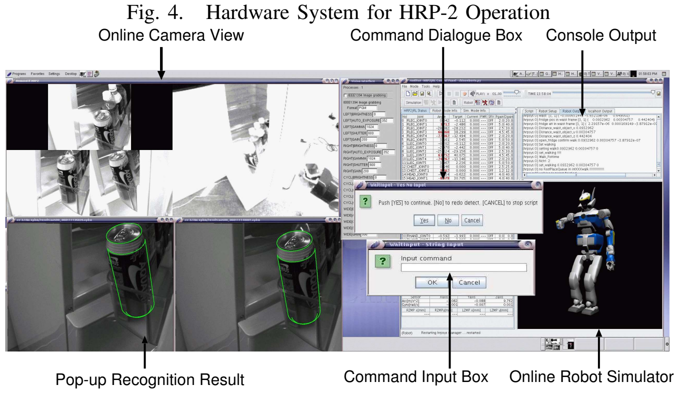

Fig. 5. GUI of the Proposed System

- Input Command Server provides mouse and keyboard parameters for discrete behavior command and interprets the conditions of the buttons and the lever of the joystick devices, and registers them as parameters for motion level operations.
- Vision Function Servers accesses to four cameras on-board the robot and provides raw image stream, stereo vision information and object recognition results implemented using the Volumetric Versatile Vision system[13] which has a segment-based object recognition engine with an expandable library of objects.
- Whole Body Motion Generator receives command from the Input Command Server to generate whole body motions for continuous motion control and behavior control. It sends commands to the Vision Function Server for the triggering of recognition process and receives information on recognized objects from Vision Function Server and information on robot's condition from the IO Board of the robot.

All the servers are implemented using CORBA with the Whole Body Motion Generator and Vision Function Server implemented on two different CPUs on board the humanoid robot HRP-2. The Whole Body Motion Generator is implemented on a real-time operating system, ARTLinux. Motor commands to the I/O board are sent every 5 [msec].

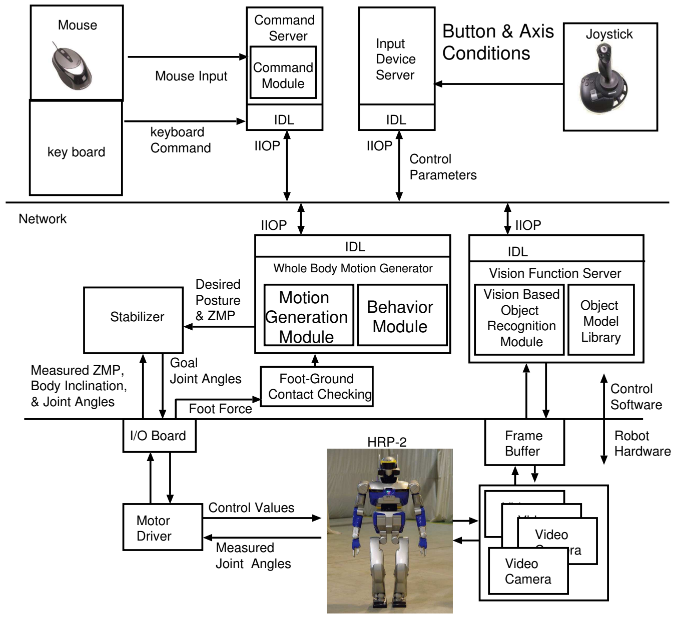

Fig. 6. Software System Architecture of the Unified Operation System

VI. Experiments

The effectiveness of the proposed system is experimentally confirmed by operating humanoid robot HRP-2 No.10 to perform various tasks including those conducted at the Prototype Robot Exhibition held at EXPO 2005 Aichi.

A total of 20 experiments have been conducted during the 11-days exhibition held at EXPO 2005 under the following conditions:

- the geometric models of a 300[ml] can drink with 65 [mm] diameter, a table with 0.75 [m] width X 1.5 [m] depth X 0.7 [m] height, and a trash box with a 120 [mm] X 280 [mm] rectangular hole for vision recognition are known
- the locations of the table and the trash box are known
- the location of the 300ml can drink as well as the knowledge about other objects in the environment is unknown

The operator was given the following tasks:

- locating a 300ml can drink on the table
- take the 300ml can drink
- throw the 300ml can drink into the trash box
- take an unknown object on the floor and put the object on the table

All of the 20 experiments were carried out without fail. Fig. 7 depicts the image of the recognition results of objects during behavior operations and the snapshots of the experiments. Fig. 7(a) depicts the scene where the operator uses behavior level command "Take Can" to take a 300ml can drink after locating the position of the can. Fig. 7(b) depicts the motion generated for behavior level command "Throw Can" after locating the position of the trash box. Fig. 7(c) depicts the scene where the operator operates the robot's right hand using joysticks to put the unknown bag on the table after approaching to the table using behavior level command "Approach Table".

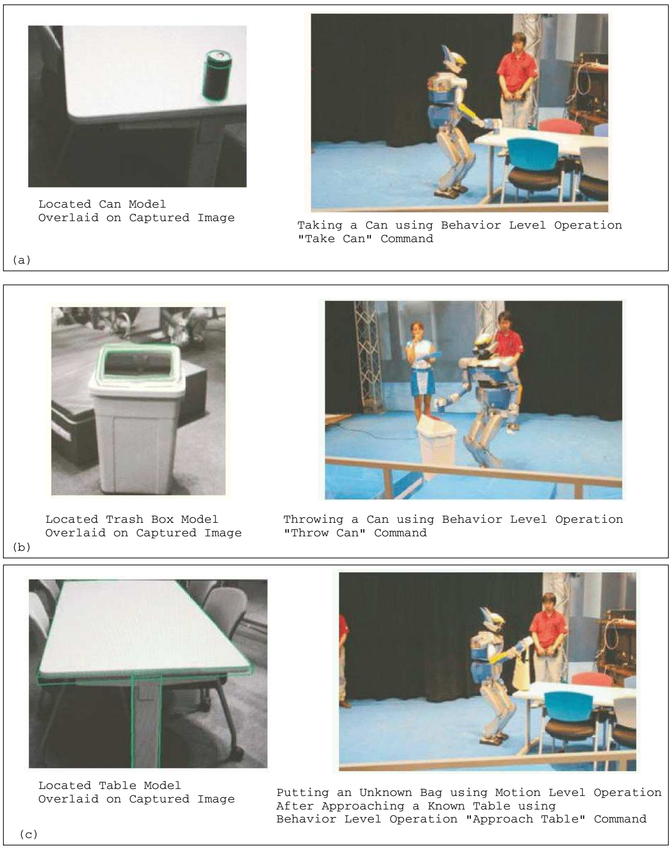

Fig. 7. Snapshots of the Demonstration during EXPO 2005 Aichi

The accuracy of the object recognition and the motion of humanoid HRP- 2 have proven to be high enough to achieve the tasks. For example, HRP-2 located the trash box from about 2 [m] far away, walked forward about 1.4 [m], and threw the can with 65 [mm] diameter into 120 [mm] by 280 [mm] rectangular hole of the trash box. The total accuracy of the location and the motion was less than 40 [mm] in this experiment.

Other experiments utilizing the proposed system were also carried out with HRP-2 No. 10 accomplishing manipulation tasks such as picking up an unknown trash on the floor using its right hand, manipulating a chair using its left hand, manipulating a can on the floor using its right foot and taking out a can from the fridge with its right hand holding the fridge door and the left hand taking out the can(Fig. 8).

VII. Conclusions

This paper presented a methodology for constructing an operation system for humanoid robots which allows the operator to seamlessly switch between continuous motion level operation and discrete behavior level operation. With these two levels of operation, the operation of humanoid robots is made possible with high robustness in human environments. Successful experiments operating humanoid robot HRP-2 with the proposed system in executing object manipulation tasks including those conducted at the EXPO 2005 Aichi confirmed the effectiveness of the proposed system.

Fig. 8. Various Tasks Realized using the Proposed Operation System

References

[1] Robert Ambrose. Humanoids Designed to do Work. Proc. IEEE-RAS Int. Conf. on Humanoid Robots, pp. 173-180,2001.
[2] K. Yokoi et al., Humanoid Robot Applications in HRP, Internataional Journal of Humanoid Robotics, Vol. 1, No. 3, pp. 410–428, Sep, 2004.
[3] K. Okada et al., Humanoid Motion Generation System on HRP2-JSK for Daily Life Environment. In: Proc Int Conf on Mechatronics and Automation 2005.
[4] http://asimo.honda.com/docs/News/newsarticle_0048.asp
[5] R. Dillmann, Teaching and learning of robot tasks via observation of human performance. Robotics and Autonomous Systems, vol 47, Issues 2-3, pp.109-116, 2004.
[6] R. Brooks et al., Sensing and Manipulating built-for-human environments. Vol. 1, No.1, Int. Journal of Humanoid Robotics, 2004.
[7] S. Schaal et al., Computational approaches to motor learning by immitation. Phil. Trans. Royal Society London, Series B, Biological Sciences, pp.358:537–547,2003.
[8] G. Metta et al., The RobotCub Project an open framework for ressaerch in embodied cognition, in the Proceedings 2005 IEEE-RAS Int. Conf. Humanoid Robots, 2005.
[9] S.Kagami et al. Humanoid Arm Motion Planning Using Stereo Vision and RRT Search. in Journal of Robotics and Mechatronics, Vol.15, No.2, pp.200-207, 2003.
[10] L. Sentis et al., A Whole-Body Control Framework for Humanoids Operating in Human Environments, Proc. of the IEEE Int. Conf. in Robotics and Automation, 2006.
[11] S. Kajita et al., *Resolved Momentum Control: Humanoid Motion Planning based on the Linear and Angular momentum*, Proc. IEEE Int. Conf. Intelligent Robots and Systems, 2003.
[12] K. Kaneko et al., Humanoid Robot HRP-2, Proc. IEEE Int. Conf. Robotics and Automation, pp.1085–1090, 2004.
[13] Y. Sumi et al.. 3D Object Recognition in Cluttered Environments by Segment-Based Stereo Vision. Int. Journal of Computer Vision, 46, 1, pp.5–23, 2002.
[14] E. S. Neo et al.. A Framework for Remote Execution of Whole Body Motions for Humanoid Robots. Proc. IEEE-RAS/RSJ Int. Conf. Humanoid Robots, 2004.