Development of an Autonomous Humanoid Robot iSHA for Harmonized Human-Machine Environment

Kenji Suzuki, Riku Hikiji and Shuji Hashimoto

Humanoid Robotics Institute, Waseda University
17 Kikuicho, Shinjuku-ku, Tokyo, 162-0044, Japan
http://www.phys.waseda.ac.jp/shalab/
{kenji, riku, shuji}@shalab.phys.waseda.ac.jp

Our research aim is to build a harmonized human-machine environment where humans and machines can interact with each other in a natural, seamless and intuitive manner. Along the sketched scenario, we have been developing a humanoid robot "iSHA" (interactive Systems for Humanoid Agent), which is designed to behave like and interact with humans. We implemented an intelligent robotic architecture, which integrates goal-oriented subsystems by taking the flexibility and scalability of the system into consideration.

iSHA has an upper body resembling a human in shape and a mobile base with two wheels. The upper body with a head and two arms has 24 degrees-of-freedom. Two wheels equipped under the body provide safe and robust locomotion. Each eye equipped with a small CCD camera, small microphones embedded in the head, and touch sensory devices on the body respectively provide binocular vision, auditory and touch sensing ability to the robot.

1. Introduction

With the aid of the growth of multimedia technology, human-machine communication by means of speech, gestures and haptics has been implemented in various scenes of our lives. Nowadays, the improvements in technology have enabled us to develop human-like robots (for example, [1][2][3]). Advances in hardware and software computing have also enabled the sophisticated implementation such as motion control and dynamic planning strategy. Recently, human-cooperative robots have become more widespread all over the world. The research on a humanoid robot [4] has also been widely extended from the mechanical realization to the biological analysis of human beings.

We aim to build a harmonized environment, where people and machines can "live" together and interact with each other. The "harmonization" is defined as the naturalness and intuitiveness of communication. The machine is required to make its own decision according to the precise selection of communication channels. In this kind of communication, multi-modality is one of the key issues. It provides natural, seamless and intuitive communication between humans and machines. We thus focused on a robot that is a machine with mobility and high redundancy. It allows humans to interact with it in various ways. The machine also should be able to achieve a given task in various ways.

In recent years, pet-type robots have become commercially available. They can exhibit attractive and devoted behavior so that the audience is satisfied with the performance. These are examples of an advanced interface that has a substantial body with multi-modality. Such multi-modal human-machine interactions typified by non-verbal communication have been widely investigated.

So far, most robotic systems have been designed for achieving a particular task. In this conventional view of system design, one module or one function performs one task. On the contrary, a complicated system such as humanoid robot should be designed from an integrated point of view. For instance, some integrated systems for a humanoid robot have been reported [5][6][7]. These systems embed a mechanism through channeling all inputs into an integrated system in a competitive and cooperative manner.

Based on the above considerations, we have been developing a humanoid robot platform, which integrates a number of agents such as image and speech processing, and a haptic interface. We also propose a robotic architecture by taking a physically grounded approach into consideration. The architecture allows various types of behavior executed in parallel. The characteristics of the robotic design are 1) autonomous and self-subsistent ability 2) system plug-in and behavior plug-in architecture, and 3) human-like modalities.

In this paper, we present an overview of a developed humanoid robot and some experimental results. In particular, object tracking by a binocular vision system and human following by hand-shape interface with a cooperative body movement will be introduced. The developed robot has several distributed agents that can work independently. Each agent has channels of communication between human and machine in a multi-modal environment.

We first introduce the methodology to realize a harmonized human-machine environment. The robotics system and the distributed agent architecture are then described. Finally, conclusions and future work will be discussed.

2. Methodology

In the harmonized human-machine environment we sketched, the robot would behave in response to given
stimuli and its internal state in a real environment. The robot can continue to interact with the humans who do collaborative work and play together, even in situations where unexpected inputs, disturbances and interruptions occur.

As a first step in the realization of a harmonized human-machine environment, we developed a humanoid robot involving the above-mentioned characteristics. The robot’s behavior can reflect continuous inputs from a complicated external environment so that the robot seems to behave in a natural and intuitive manner. The robotic architecture involves two types of processing: physical and intelligent (logical). The multi-process and independently distributed modules provide adaptive and robust control to the robot.

The developed robot iSHA has a number of degrees-of-freedom in its body, especially around the head. Humans can make a variety of body expressions by using muscles. In place of the muscles, actuators (DC or AC) on joints conduct the behavior of the robot. Fig. 1 shows the overview of the developed humanoid robot.

iSHA can be divided into two body parts, an upper and a lower body. The upper body resembles a human in shape, while the lower body is a wheelchair. The upper body with a head and two arms has totally 24 DOFs; 8 for the head, 4 for the neck, and 6 for each arm. The lower body has two wheels, which are independently driven, that provide safe and robust locomotion to the robot. The total is thus 26 DOFs in its whole body. In particular, the eye structure has actuators independent of the head movement. This therefore helps the robot to achieve a fast object tracking.

The host computer (ART-Linux, Celeron 700MHz) that works to control the actuators with the hardware-scheduled real-time process is embedded in the backside of the robot. The images obtained by two small CCD cameras are transmitted to another embedded computer (Windows 2000, Pentium III 800MHz) that is engaged in the image processing with the image processing board (Hitachi IP5000), as well as the processing of the data from the sensory receptors of the microphone, the sound-sensing devices and tactile devices.

3. Robotics System Design

The characteristics of the developed humanoid robot are summarized as follows:

1) Autonomous and self-subsistent ability

Most individual creatures have an autonomous and self-subsistent ability. We focus on that as the fundamental character of system design. The developed humanoid robot does not need any power supplier from the external environment but can itself move and act with an embedded lead storage battery. Two included computer can then make the robot autonomous.

2) System plug-in and behavior plug-in

The developed robot has a substantial interface integrating a number of multi-modal components. In such a system, it is desirable that any subsystems can be added to the existing robotic system. As for the flexibility and scalability of instruments, we have chosen a network-based architecture inside the body. Modules such as a speech recognizer/synthesizer, an image processor, a behavior coordinator, a receptor of auditory sensing, and a touch sensing and haptic interface device have been developed as server applications. Moreover, the robot can be easily connected to the local area network via an embedded wireless connection. The robot, therefore, not only can perform autonomously but can also be handled by a remote control operation over a TCP/IP network connection. The proposed robot thus allows a system plug-in extension with the aid of a general Ethernet connection.

As for the robotic architecture, the concept of a behavior plug-in has been realized by a behavior-based architecture and a multi-process operating platform, ART-Linux, which is a real-time operating system designed for support development of large-scale real time processing software. In short, the real time processing performance is added to the Linux operating system. We can thus flexibly add any abilities or types of behavior to the developed robotic architecture. The hardware scheduling is guaranteed by the operating system. The details of software architecture are described in the next section.

3) Human-like modalities

iSHA has several channels of communication with the external environment. These are designed so as to enable humans to give stimuli to the robot in the same way as communication among humans. Each eye is equipped with a small CCD camera that provides binocular vision to the robot. The equipped stereo microphones provide an auditory sense and enable the robot to receive environmental sounds. The robot can thus execute simple sound localization.
We implemented touch-sensing devices consisting of a metal plate and a cushioning material on the front, back and left sides of the body. Additionally, the robot has a hand-shaped force interface to sense the human's intention by hand shaking that is one of the intuitive ways of communication. These allow us physical interaction with the robot through our hands or body.

4. Robotic Architecture

An architectural framework for sensing and reasoning processes should allow the robot to display goal-oriented behavior and should preserve the ability to respond to critical situations in a real-time environment.

Some key points to be considered in the design of a planning and control architecture are that the robotic architecture should be distributed, allow both reactive and deliberative reasoning, and involve a method for dealing with information from multiple sources.

The architecture developed for control of Shakey the Robot [8] is well-known as a centralized architecture. The robot operated by gathering all available sensory data and creating a unified representation of its environment. Although the centralized architecture has the advantage of enabling the robot to behave autonomously in a coherent fashion and with multiple goals, it is not appropriate for a real-time system in a dynamic and uncertain environment. In contrast to the centralized architecture, subsumption architecture [9] that employs priority-based arbitration is one of the representative instances of behavior-based architecture [10]. In the architecture as typically described, simple types of behavior are hierarchically organized so that more complex types of behavior emerge. A robot control system should be decomposed according to not the structure of the internal functions but the desired behavior of the machine in response to the external environment. Such behavior-based architecture for the supervision of mobile robots is recently in wide usage as intelligent robotics architecture. For example, an approach to build a sociable robot with the subsumption architecture has been reported [11]. It is, however, not a physically grounded architecture but a computational one based on intelligent processing.

4.1 Double-layered hierarchical architecture

In the developed robot, we adopted a double-layered structure as illustrated in Fig. 2. It should be noted that a double-layered hierarchical processing is implemented in order to clarify the stimulus difference.

Communication among humans consists of two types: physical and intellectual interactions. The former is the communication under physical constraint and interaction in accordance with direct/indirect contact. The latter is informational interaction through intellectual ability of a high order with each other. We humans can do either of these interactions or a combination of both types of interaction. For instance, speech conversation is one of the most effective and intellectual interactions between humans but is not physical interaction. In contrast, the hand shaking has two aspects: we express a sign of goodwill with this motion and also gain force interaction according to the motion at the same time. These interactions are caused by different demands and processing. Physical interaction is composed mainly of immediate responses and simple types of behavior. Intellectual interaction is composed of intelligence and sophisticated types of behavior.

Most of the systems dealing with physical interaction, including computational compliance control, are carried out by the same method as intellectual interaction at the processing level. In contrast, we propose a hierarchical architecture that has two independent layers in order to clearly separate the physical and intellectual interactions. The types of behavior caused by the intellectual procedures of the system are constrained by the physical procedures. We assume that the mixture of processing can thus provide safety and credibility for communication to the machine.

There exist two layers: signal processing layer and computation layer. As for the signal processing layer, data from touch sensing devices are fed to the behavior coordinator. The cells of the output layer correspond by signals to the actuators.

As for the computation layer, sensing data are given to input cells from a sensing module installed in the robot (e.g., low-level robot sensor data, equipped stereo cameras, microphones and tactile sensors), and each cell has a unique source from an input channel. In the internal procedure in the computation layer, there are two types of behavior coordinator in which each one receives signals independently from all input cells. These coordinators are based on short-time and long-term memory. Each behavior coordinator is connected to and has influence on the others by activation and inhibition. A weighted sum of the input signals is fed to these coordinators. Each one corresponds to a style of behavior of the robot (e.g., dancing movements, binocular object tracking, response to tactile sensing). Specifically, a signal in the
input cells is transmitted to the behavior coordinator. The output of each coordinator is then multiplied by certain (fixed) weight parameters and transmitted to the output cells. Each cell has a unique connection to output channels such as actuators, sound and visual outputs. The operations through the connection constitute a linear combiner.

The signal to actuators from both layers is simply summed by the analog adder and is fed to each actuator. That is, the behavior of the robot depends upon the balance of the signals from both physical and intellectual processing layers.

In the present work, the connection between behavior coordinator and the weight parameters are predetermined and fixed. Consequently, the meaning of each type of behavior is predetermined by the system designer. The multi-goal tasks are achieved because each behavior coordinator produces an action independently.

5. Evaluation Experiments

In this section, we introduce some preliminary experimental results while considering the two styles interaction mentioned above: 1) Reaction to touch sensing, 2) Hand shaking, 3) Tracking and reaching an object by the binocular vision system, 4) Dancing movements according to a given tempo, 5) Reaction to auditory sensing and 6) Integrated types of behavior.

The robot allows humans around it to behave freely. The robot’s performance is designed for human intuitive understanding so that each type of behavior can be accepted easily by the companions who interact with the robot. Throughout these experiments, we evaluate the effectiveness of the proposed robotic architecture.

The robot performs an action in response to human stimuli in real-time. The types of behavior are chosen and carried out, depending upon the robot’s priorities.

In the following experiments, most actuators are operated by position control. By obtaining the angle of each joint with the embedded encoder, the control module provides each joint with the desired angle $\theta_d$. Conventional PID control is applied for each joint. The PID gain parameters are empirically chosen and fixed through experiments. The control module independently processes in the robot operating system by parallel computing.

5.1 Reaction to touch sensing

Reactive movement is one of the basic types of behavior implemented at the physical processing layer. Physical interaction with humans is the highest priority for the robot.

The developed simple touch-sensing devices are illustrated in Fig. 3(a). The device on the front or back side has three switches, while the one on the left side has one switch. By applying an external force to the plate, the robot can obtain a human’s intention by physical interaction. In this experiment, the robot moves in a direction so as to cancel the applied external force. The robot moves left, right, backward and forward according to the combination of switches.

For example, a push from the front side causes the robot to move backward. Although the robot is tracking an object and moving his arms to reach the object, physical interaction can still be possible. Therefore, the robot would respond immediately trying to continue tracking the object under the given constraint.

5.2 Hand shaking and tracking humans

We have developed a hand-shaped force interface as illustrated in Fig. 3(b). The interface consists of 2 DOFs with rotary variable resisters and 1 DOF with sliding variable resisters so as to sense the three directions of the human’s intention: push/pull, horizontal and vertical mo-

![Fig. 3 Equipped touch-sensing devices and hand-shaped force interface](image)

(a) grasping back (t=0.0, 1.0 and 1.5 [sec])

(b) pulling by the hand (t=0.0, 1.0 and 2.0 [sec])

![Fig. 4 Hand shaking and the control](image)
tions. The interface that is embedded in the right arm enables us to communicate with the robot by shaking hands.

In handshake communication, humans can express his/her mental intention in several ways: for example, holding kindly or strongly. During the handshake, a force emerges according to the difference between the motions of the hands. By applying a force to the other, one can lead the other. For the intuitive understanding, we set the following conditions. (see also [12])

1) When a human grasps the interface, the robot responds by grasping back with the thumb.
2) When a human applies a force to the interface, the robot behaves so as to cancel the applied force by utilizing the right arm (2 DOFs), wrist (2 DOFs) and wheels (2 DOFs).

Figure 4 shows examples of the handshake communication.

5.3 Tracking and reaching an object by binocular vision system

The robot tries to reach a recognized object through coordinated movement. The cooperative movement of each part, such as eyes, neck and wheels, is necessary for a robust and flexible object tracking. However, because the joint structure has redundancy, the specified path can be chosen under an arbitrary criterion.

Our experiment differs from the conventional binocular vision system [13] in utilizing the whole body including the eyes, neck, body and arms. In this experiment, each part is moved in cooperation with the others to track the object segmented by extracting color hue information through the binocular vision system. The tracking and algorithm are independently divided into two terms, head and arm movements.

5.3.1 Head movement: The head contains 8 DOFs as illustrated in Fig. 5. Each eye has 2 DOFs, horizontal (yaw) and vertical (pitch) axes. Seven of these are used for object tracking.

In this experiment, we use a small ball that has a specified color for extraction. The coordinates of the center of the image of the recognized object are calculated. The distance to the recognized object can also be calculated by the azimuth difference. The accuracy is, however, not sufficient to detect the object in 3D space for real use. Therefore, we use only 2D image coordinates of the object. As for the distance, the size of the object as the average of the object in the images obtained from both eyes is used as a substitute for the parallax. Based on these, the position of the object in 3D space can be estimated.

The velocity of response at each joint is different due to the difference in inertia for each part. Each eyeball moves faster than the neck, and the neck moves faster than the body. In general, in order to hold the object in the range of vision, not only eyeball motion but also the cooperative motions of the neck and body are necessary.

Before the object goes out of sight, the robot head should follow the object.

The patterns of gazing at the object are thus itemized as follows: When a target is located in the central part of the camera image and a short distance from the robot, eyeball motion is induced due to the fast tracking property. When the target is located peripherally around the camera image and a short distance from the robot, neck motion results. When the target is located at a great distance from the robot, body motion is finally caused. In addition, when the target is located on the left/right side of each camera image, the priority is given to the left/right eye, respectively. The torque applied to each part is tuned so that it reflects the above characteristics.
5.3.2 Arm movement: So far, many studies on the kinematics of a robot have been reported [14][15][16]. As for arm control, we adopted a method of the combination of primitive motions. By using the estimated 3D position from the binocular vision system, the robot changes its arm posture incrementally so that the robot hand reaches the object.

Unlike deriving a locus to the desired position, the robot takes a posture for minimizing the mean-square error to the target incrementally at the local coordinates. This algorithm does not aim to acquire an optimal pathway; however, it contains adaptive and robust features. Even if a joint is broken or disabled, the robot can continue to assume a posture using other joints.

A robot achieves the reaching task with 4+1 DOFs joints as illustrated in Fig. 5 (joints (8)-(10), and (12)). The yaw axis of the lower arm (joint (11)) is used only as a conditioning direction for the palm of the hand. The position of the hand at time \( t \) can be estimated by the given kinematics. Each joint can be moved with the basis shifting, and \( \Delta \theta \) is a bipolar step function at each joint.

The yaw axis of the lower arm (joint (11)) is used only as a conditioning direction for the palm of the hand. The position of the hand at time \( t \) can be estimated by the given kinematics. Each joint can be moved with the basis shifting defined as the minim shifting \( \Delta \theta \). The prospective posture at time \( t+1 \) can then be obtained by the product of the minim shifting \( \Delta \theta \) and the arbitrary derivative gain \( k_i \). By repeating this operation, the robot arm approaches the object.

The arm posture at time \( t \), \( u(t) \), is described as:

\[
u(t) = f(\theta_1(t), \theta_2(t), \theta_3(t), \theta_4(t))
\]  

The posture at time \( t+1 \) is delivered as:

\[
u(t+1) = \left[ \begin{array}{l} \theta_1(t) + k_1 \delta_1 \cdot \Delta \theta_1 \\ \theta_2(t) + k_2 \delta_2 \cdot \Delta \theta_2 \\ \theta_3(t) + k_3 \delta_3 \cdot \Delta \theta_3 \\ \theta_4(t) + k_4 \delta_4 \cdot \Delta \theta_4 \\ \end{array} \right]
\]

where \( k_1, k_2, \ldots, k_4 \) denote a gain constant to each basis shifting, and \( \delta \) is a bipolar step function at each joint. The prospective posture is determined in order to minimize the following evaluation function \( E \), aiming at moving the hand position close to the estimated position \( \hat{u}(t) \) of the recognized object.

\[
E = (\hat{u}(t) - u(t+1))^2
\]

With the use of a motion simulator system, the proposed algorithm is proved to reach the vicinity of the object.

The object position is not estimated precisely and the pathway obtained by the incremental reaching method is not proved to achieve the task in all cases. However, in the cases in which the target is located in front of or near the robot, the reaching movement can be successfully achieved.

Figure 6 shows camera images from both eyes, the result of segmentation of a recognized object and the data flow of the object recognition.

5.4 Dancing movements according to a given tempo

The robot can perform a dance movement according to a given tempo. The tempo is defined as the rate of speed, motion or activity. For example, we can extract the tempo and beats from music and sound. We consider that the tempo is one of the parameters by which the robot and the external environment are synchronized.

The pre-defined gestures are very simple; the robot dances by swinging his arms and head. By giving the size and frequency of motion of the arms and wagging of the head, the robot can dance with his whole body. The movement of both arms and head is synchronized with the given tempo. The position is given by a sinusoidal pattern with a frequency that corresponds to the tempo and an amplitude that corresponds to the robot’s intention (currently a fixed value). We implemented the above dance movement with three joints for the neck and four joints for each arm as a behavior coordinator. The movement rule is described as follows.

\[
x_i(t) = A_i \sin(\theta_i + \omega t / \Theta_i)
\]

\( A_i \) denotes the amplitude, and \( \omega \) denotes frequency, which corresponds to a given tempo of the movement at joint \( i \). \( x_i(t) \) represents the angle. \( \Theta_i \) represents the predetermined range of movement at each joint \( i \). The robot can perform a coordinated action by hardware scheduling with the body of 13 DOFs.

We have implemented a tempo tracker that enables us to provide a tempo to the robot by handclaps. The tempo of the robot \( T_{robot} \) is determined according to the obtained tempo with the equipped stereo microphone. By tuning the timing, the robot can thus synchronize the given periodical signals.

5.4 Reaction to auditory sensing

The three microphones embedded in the head are used for sound localization. Moreover, the robot can receive some voice commands. One more microphone and stereo speakers are attached for speech recognition and synthesis. In the present system, the robot utters a voice command before executing the ordered task such as stop, start, forward, backward, left, right, tracking a ball and dancing.

5.5 Integrated types of behavior

Some integrated types of behavior are illustrated in Fig. 7. The above-mentioned types of behavior appear in parallel or simultaneously, not in series. For example, when the robot failed to track the object by the head, humans can help the robot to find it by turning it using the hand, e.g.) we can use handclaps to draw its attention. The important advantage of multi-modal interaction is such an integrated type of behavior. The robot should be able to achieve a given task in various ways. The proposed system and architecture allow a number of types of behavior by means of several channels of communication such as vision, auditory and haptics. These channels make the robot behave more sophisticated and flexible.
6. Discussion

In this paper, we introduced the specifications and architecture of the autonomous robot iSHA for harmonized human-machine environment. The experimental results showed that the developed robot provides various channels of communication between humans and the robot. As for the robotic architecture, we proposed a double-layered structure for physical and intellectual interactions. The physical layer is placed as the lowest one and has the highest priority. The communication ability of the robot with its external environment depends upon the physical constraint. The intellectual processing by computation is dependent upon and limited by that.

In addition, human-like modality is the other characteristic of the developed robot. Following a particular object with the binocular vision system and sound localization by an embedded stereo microphone are implemented. Moreover, we have introduced haptic and tactile interfaces for communication between humans and the robot, which allow humans to perform embodied interaction. At present, although binary sensors are implemented for touch sensing, humans can handle the robot with their hands through physical interaction. The physical interaction layer plays a role in gathering the reflex actions of the robot. The developed robot thus provides various ways of sophisticated communication by combining physical and intellectual interactions.

The other focus is the robustness. The robotic architecture allows multi-process tasks, and the control modules are implemented independently. In addition, any instruments can be connected to each other with the network protocol. Therefore, the robot has the robustness of both the system and of behavior. Even if lost connections or machine problems occur in one of these instruments, the robot tries to behave in a possible alternative way. The proposed method of reaching an object is an example. Because the robot assumes a posture by the defined basis shifting of each joint, it can continue to follow the target even in the case where some joints are disabled, or the motion is disturbed by an external force.

In such a complicated robotic system, the scalability of the system is also an important discussed attribute. We installed a system plug-in architecture preserving the advantages of behavior-based architecture. The instruments included in the robotic system are connected with the common TCP/IP network protocol. For example, when we install a system module into the existing robotic system, we simply put it in the network to exchange data via the network with/without cable. The constructed network system inside the robot can accept the connection from any optional devices such as the speech processing module, the image processing module, and a wireless connection from an external computer. For example, a robot simulator that runs on an external laptop computer can handle the robot over a wireless LAN connection.

As for collaborative work with humans, dancing movements are an example. The robot can dance according to a given tempo. The torque applied to each actuator in phase with a given tempo represents the synergic effect of the robot. By matching the phase between the external signals and the robot’s internal cycle, a rhythmic synchronicity would take place. We consider that a variety of input signals should influence the robot movements. The robot would synchronize its motion with the extracted periodic signal.

7. Conclusions and Future Work

In the present work, the developed robot requires numbers of empirically predetermined parameters. Not only fundamental parameters such as the PID gain, but also parameters for intellectual interaction such as the range of reaching its object and the priority of the types of behavior. Tuning the parameters by learning is considered as one of the necessary abilities.

As for the processing manner, we classified two types of procedures, physical and intelligent processing. In the proposed architecture, each behavior coordinator received the sum of weighted stimuli from input channels. However, when one type of behavior is classified into a procedure, the role is then consequently established. The difficulty is how to associate a behavior type with the type of processing. Self-recognition and self-evaluation are the future research topics. The robot must build an awareness of itself.

So far we have investigated the communication of
KANSEI that is regarded to include some terms such as feeling and sensibility, between human and machine. In communications between humans, the role of KANSEI information is as important as logical information [17]. When particular stimuli are given, humans can perceive and understand them. We notice that such information is often processed not logically but unconsciously and involuntarily. Such characteristics, or rather types of behavior, might be regarded as KANSEI. Humans have the ability to understand things by intuition. In other words, aside from logic, humans adopt information processing based on KANSEI. Consequently, a robot that interacts with humans should deal with KANSEI information processing. This is one of the requirements for the harmonized human-machine environment.

The term “artificial life” has become one of the key issues in robotics research. For instance, applications of evolutionary robotics (for example, [18][19][20]) have often been conducted. We have been investigating an autonomous mobile robot that has a self-sustaining ability and can acquire its survival strategy in the real environment where the robot is living [21]. With the improvement of a large-capacity battery and a high-power inverter, the autonomous and self-subsistent ability can be accrued to the robot. These characteristics are also essential factors in order for the robot to behave in coexistence with humans.

In this paper, we sketched the direction for building a harmonized human-machine environment for a humanoid robot. The environment is not limited only to humanoids but can also be extended to general machines. In the near future, human-machine interaction will enter the next stage, human-humanoid interaction. Such human-cooperative robots will become an “emotion activator”, which is regarded as a metaphor meaning that the robot can not only imitate human motions and gestures but also can stimulate humans with its movements and behavior.

Acknowledgements

The authors would like to acknowledge Dr. Takanishi, A. and Dr. Takanobu, H. for the mechanical design of the humanoid robot.

The presented work was partly supported by the JSPS Research for the Future Program in the Area of KANSEI (Intuitive & Affective) Human Interfaces, and by a grant-in-aid from the Japanese Ministry of Education, Science, Sports and Culture (A-11305021). This work was also done as a part of the “Domestic Study for the International Joint Research Program into Intelligent Manufacturing System” commissioned by IMS Promotion Center.

References

[17] Hashimoto, S., "Humanoid Robot for Kansei Communication," Proceedings of the Second Inter-

