24

Development of an Android System Integrated with Sensor Networks

Takenobu Chikaraishi, Takashi Minato and Hiroshi Ishiguro
Osaka University
Japan

1. Introduction

In recent years, research and development has been conducted on humanoid robots (Kanda et al., 2004), (Kajita, 2002) that are capable of interacting with humans in daily life, because many robot tasks that need a human-like presence, namely communication with humans in daily life. A robot equipped with a humanlike presence enables natural human-robot communication, that is, a human can naturally communicate with the robot as if he were communicating with another human. One of the approaches to learning the principles of giving a human-like presence to a robot is to develop a robot that can sense and behave like a human. Developing such a robot will reveal the principle of human-like existence. In order to develop a robot that can naturally communicate with humans through communication channels (Daibo, 1993), we have to solve the following issues.

- A robot has to be provided with a very humanlike appearance and motion to interact socially with humans (Duffy, 2003).
- A robot has to be provided with perceptional functions to communicate with

humans through human communication channels. However, the degree to which a human-like nature is needed and how much perception is needed in order to realize natural human-robot communication have not yet been revealed. We have been conducting research on these issues to explore the principles for realizing natural human-robot communication using a robot with a very human-like appearance, which is called an "android" (Ishiguro, 2005). The android's human-like appearance must be sufficiently anthropomorphic to elicit a natural interpersonal reaction (Shimada et al., 2006). Thus studies on the android's human-like nature in appearance and motion have been conducted and they showed that a human-like presence is realized in short-interaction or interaction with a restricted context (Noma et al., 2006). However, the android's system is not yet sufficient to realize a human-like nature in a wider situation because the android's perceptional functions are not implemented or substituted by an experimenter in these studies. In order to clarify the perceptional functions required for natural communication, we try to realize human-like perception with an android.

The communication abilities of robots have been improved by achieving the same perception as humans using built-in sensors (e.g., (Okuno et al., 2001)). However, built-in sensors have many restrictions on range and resolution, and cause difficult issues for natural communication.

There is a technique for providing perception functions to the environment itself, that is, embedding sensors in environment (Morishita et al., 2003) (Mori et al., 2005) (Ishiguro, 1997). Such a system is quite different from a human perception system, but natural communications are possible if the robot pretends to perceive humans even though it actually does not.

In this research, to remove the restriction of built-in sensors and to achieve natural communications with humans, we propose a quite new robot sensor system. This system achieves perception and android action using as many kinds and numbers of sensors as possible including not only sensors embedded in the environment but also built-in sensors. This is called a "sensor network" in this paper. By adding various and efficient sensory system to the android with very human-like appearance, the world's first humanoid robot system can be realized, which can achieve a human-like nature integrating appearance, movement and perception.

The sensor networks give the robot a quite different perceptional modalities from a person's ones. However the modalities are not important for a human-like perception.

This means that the system needs to select the sensory information appropriate to the current situation suited for natural communication. For example, human finds where other person is by using vision or audio sensor (e.g., eyes or ear), but most important information for human-like interaction is the position (e.g., situation). Then, in order to facilitate dealing with the situation-dependent information, the system prepares various sorts of motion modules, where each module binds sensory information with an android's gesture.

This paper reports on the development of the communication system to integrate the sensors embedded in the environment and the android.

Fig. 1. Android Repliee Q2.

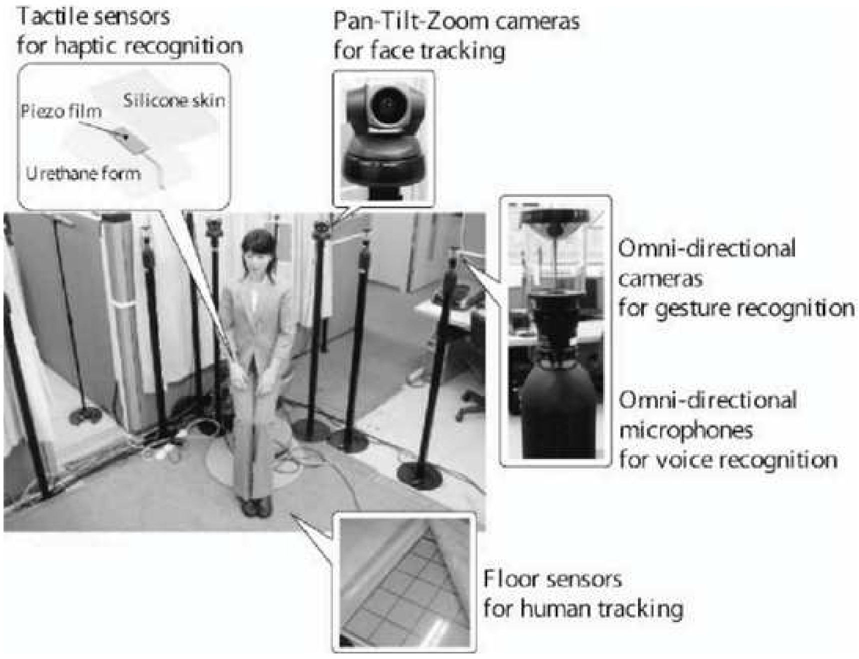


Fig. 2. Hardware composition of the developed system.

In this paper, to show that the system we have developed can achieve natural human-like behavior by an android, an experiment of evaluating impressions is conducted. We focus on a situation in which the android is sitting quietly. This situation has been implemented and evaluated in the previous studies (Shimada et al., 2006)(Noma et al., 2006). But these studies did not use sensory information. In order to make the android behavior more human-like, reactive behaviors (e.g., when a person appears, the android turns its head to the person) must be necessary. In the experiment, subjects compare two kinds of the android motion. One is a motion where the android shows reactions based on sensory information, the other is where the android does not depend on sensory information and randomly selects motions. In this paper, first we explain the hardware composition and software architecture of the system we have developed in chapter 2. Next, we report the result of the above-mentioned verification experiment in chapter 3.

2. An android system integrated with sensor networks

2.1 Hardware composition

The sensors used this system are illustrated in Figure 1. In the android's environment, omni-directional cameras, microphones, pan-tilt cameras and floor tactile sensors are arranged. Tactile skin sensors are built into the main body of the android. These sensors are connected to sensor processing computers, and sensor information is transferred between mutually connected computers through TCP/IP(Figure 3.).

One circle in Figure 3. denotes one sensor (the number of skin sensors built into the android body is 42, but only 4 circles are illustrated because all information is read by four reading circuits). In the following, first the specifications of the android are described, then the specifications of the each sensor and sensor information processing are given.

2.1.1 Android Repliee Q2

Figure 2. shows the Repliee Q2 android used in this system. To achieve a very human-like appearance, it was made by using a plaster cast of a real human. The skin material is made from silicon rubber for medical treatments, so both the appearance and the feeling of touch is close to that of a human. Movable parts all have 42 degrees of freedom, including 3 on the eyes, 1 on the eye lid, 1 on the cheek, 7 on the mouth, 3 on the neck, 9 on both arms, 2 on both hands and 4 on the waist. The android can make various shapes for the mouth and face with 13 degrees of freedom on the part of the head, thus giving it very rich expression. The android has skin sensors only in its body. Under the silicon rubber skin or clothes above the knee, 42 highly sensitive tactile skin sensors are built using PVDF film.

Each degree of freedom is driven by a pneumatic actuator. The reactions of the actuators are very natural due to air dampers against external forces without any special kind of control. This achieves a much safer interaction with humans than other actuators such as oil pressure actuators and DC servomotors. Many degrees of freedom allow the android to express the unconscious movements of humans, including movements of the shoulders and chest caused by breathing. The pneumatic actuators give it very smooth human-like motion. However, it has no ability to stand up and move. These actuators also make no large noise when driving. The pneumatic compressor, which is the source of power, generates a large noise, but it can be located far from android's main body. Because of this, the drive noise of the actuators does not disturb the communications.

The servo controller for the pneumatic actuators is located outside of the android's main body, and the android is controlled by receiving order values for each degree of freedom over a serial connection from a computer. Transmission over serial connection is executed at the speed of 20Hz. The speech is achieved by playing a recoded voice from flat speakers built into the android's chest. One computer controls the pneumatic actuators and plays the voice. The detailed specifications of the android are described in (Ishiguro, 2007)

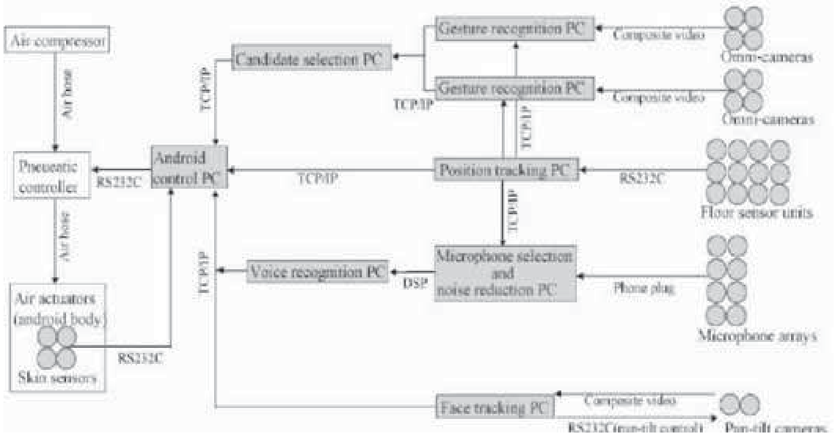


Fig. 3. Network architecture of the developed system.

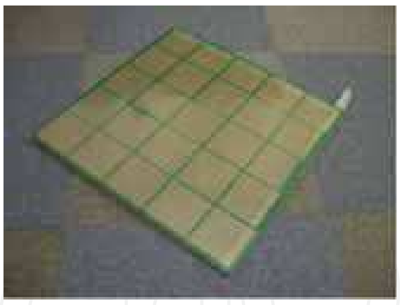


Fig. 4. Floor sensor unit

2.1.2 Floor tactile sensor

Main purpose A floor tactile sensor unit is installed in front of the android in order to detect and track the position where a human stands.

Hardware We adopted floor sensor units as depicted in Figure 4. (Vstone VS-SS-SF55) for the floor sensor. The floor sensor consists of pressure-sensitive sensors that are 5 mm apart in a lattice form in order to detect and track human positions. Spreading this sensor over the floor surface allows it to detect the human position. Each floor sensor unit has a size of 500 x 500 x15 mm and a sensitivity of 200-250 g/cm^2. It has a resolution of 100 mm x100 mm, each unit can detect weight positions of 5 x 5. That is sufficiently efficient resolution for detecting a human position because adult feet generally have a size of about 200-300 x100150 mm. In this system, 16 (4x 4) parts of floor sensor units (totally 2 mx 2 m) are installed on the floor surface beneath carpet in front of the android. Each floor sensor unit is connected to a computer through a serial connection, and the transmission speed is about 37 Hz.

Data processing When a human is walking, the sensor signal is temporally and spatially discrete. The signal pattern is also greatly different when one foot is on the ground and when two feet are on the ground. Therefore that causes a many-to-one association problem when the human position is required from the floor sensor data. Murakita et al. proposed a method based on the Markov Chain Monte Carlo Method to track plural humans solving this intrinsic association problem of the floor sensor (Murakita et al., 2004). In this system, we implement this method and track several people using the floor sensor. Human detection is executed at the speed of about 17 Hz.

2.1.3 Omni-directional cameras

Main purpose We installed 8 omni-directional cameras around the front space of the android (Figure 1.), in order to recognize human gestures in addition to human tracking.

Hardware We adopted a VS-C42N-TR (Vstone Co. Ltd.) omni-directional camera. The imaging range is about 15 degrees to the upper side from the horizontal plane of all azimuths and about 60 degrees downside from the horizontal plane of all azimuths. The number of effective pixels is 768x 494. Eight these cameras are installed at a level of 1.5 m from the floor surface. Four cameras are connected to one computer through a capture device and two computers process the image data from all of the omni-directional cameras.

Data processing A human tracking and gesture recognition method is proposed with an omni-directional camera network using a triangulation of omni-directional cameras and omni-directional appearance-based model (Nishimura et al., 2001) (Ishiguro & Nishimura, 2001). This method allows it to recognize human gestures independent from the object direction from the cameras. However, when moving object(s) exist in the image, the system may not determine a recognition region in the image, because this method employs subtraction images to detect human positions and determine the recognition region of the image. Therefore, we adjust the floor sensor information for this system. The human region of the image is estimated from the information of human position obtained from the floor sensor, and the gesture recognition is executed. The candidates for gestures recognized by the two computers and information equivalent to the reliability of the recognition are transferred to other computer. The computer selects the one candidate with the greatest reliability and this is the final result of the gesture recognition.

Therefore, three computers are used in all for gesture recognition. In this system, gesture results including "bowing" and "indicating a direction" can be recognized. The recognition is processed at a speed of 20 Hz. When several people exist in the recognition area, each person's gesture can be recognized.

2.1.4 Microphones

Main purpose In order to recognize human voices, eight omni-directional microphones are installed in the same places as the omni-directional cameras.

Hardware One omni-directional microphone consists of four built-in condenser microphones in the bottom part of the omni-directional cameras. One omni-directional microphone is achieved by synthesizing four signals from four microphones. All of the microphones are connected in parallel to one computer, and the computer extracts voice data for the voice recognition program. All of the voice data are transferred to another computer that recognizes the speech. Two computers in all are used for voice recognition.

Data processing For voice recognition, a voice signal with as little noise as possible is needed. For that, the system must select the nearest microphone to the human as the sound source based on human position information obtained from the floor sensor. Next, background noise must be filtered out of the signal obtained from the selected microphone. An existing voice recognition method (Microsoft Speech SDK) is applied to the voice signal thus obtained. Recognition grammar can be changed according to the situation.

2.1.5 Pan-tilt cameras

Main purpose The system must know the human face position and direction in order to know the direction for the android to talk to. Therefore, two cameras that can pan, tilt and zoom are put on both sides of the android.

Hardware We used EVI-D100 (Sony Co. Ltd.) pan-tilt cameras. These cameras can turn across a field of view of ± 100 degrees in the horizontal direction (maximum 300°/sec.), and ± 25 degrees in the vertical direction (maximum 125°/sec.). The two pan-tilt cameras are connected to a computer by serial connection, and the computer can control the pan, tilt and zoom functions of the camera. These cameras are connected to the same computer through a capture board, for image processing and controlling the image direction by the computer.

Data processing We implemented a face detection function based on a histogram of the edge image and extracted a skin color image (e.g. (Adachi et al., 2001)). Cameras are controlled by pan, tilt and zoom functions to keep the detected face region in the center of the sight and to keep the same size image. Therefore the face position in three-dimensional space is known from the position of the pan, tilt and zoom functions. The face direction can be detected from three candidates (front, right and left). This process is executed at a speed of 20 Hz.

2.1.6 Skin sensors

Main purpose The skin sensors detect contact by humans. Only these sensors are installed in the body of the android, unlike other sensors, because communication by contact needs to use the body. In all, 42 skin sensors are installed beneath the skin above the knees.

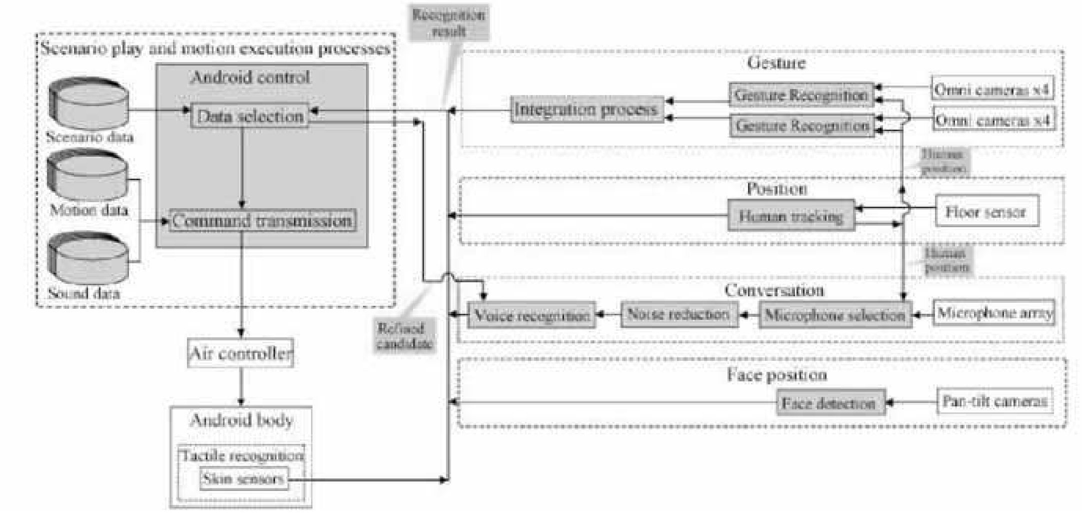


Fig. 5. Software architecture of the system.

Hardware The skin sensors are made of PVDF film. The PVDF film is sandwiched by urethane rubber and silicon rubber skin. Because each sensor element has a size of 10 X10 mm, 42 sensors are not sufficient to cover the whole surface of the android's body. However, the skin sensor can be activated even when a body part with no skin sensor is touched because the distortion of the urethane layer can easily spread on the android body when the android is touched.

Data processing A value equivalent to the contact force can be obtained by temporal integration of the output value, because the PVDF film outputs the value proportional to the transformation speed. The low distribution density of the skin sensor restricts the ability to distinguish contact behavior accompanied with movement on the skin surface (e.g., "stroking"). The android detects which part of the body is touched and how much external force is applied. Thus, contact behavior including "tapping the shoulder" "touching arms" can be recognized. The detection of contact is executed by the computer for the android's motion control. Data processing is executed at a speed of 20 Hz.

2.2 Software architecture

The flow of the recognition and determination of behavior is illustrated in Figure 5. Eight computers (Pentium IV 3GHz) are used for sensor data processing and motion control. Sensor data processing consists mainly of five parts: gesture recognition, human tracking, voice recognition, face detection and contact detection. Each data processing is explained in chapter 2.1. Each processing part is not independent, as gesture recognition and voice recognition depend on the result of human tracking. Applying the results of other recognition helps increasing recognition accuracy. Each sensor processing outputs the following information.

- Gesture recognition: Gestures of each person and human IDs.
- Human position tracking: Positions of each person and human IDs
- Voice recognition: Recognized word sequence.
- Face detection: Face position in three dimensional space and face direction.

• Contact detection: Touched body part and strength of external force. The android's motion is treated by module units gathering time-series command values for each degree of freedom and voice data. One motion module is from a few seconds to about 30 seconds (e.g. "bowing" and "waving a hand"). In this paper, we call packages of motion module and a diverging condition process "events". In an event, first the motion is executed, then the following motion is selected based on the result of sensor processing. Additionally we make a script that describes the steps of the procedure of the events. We call this a "scenario". An event includes a module to execute, sensor data processing to identify the situation after the module executes and the following event ID selected by that recognition result. If voice recognition is selected as sensor processing in the event, the recognition grammar must be specified.

When the scenario is pointed out in the android control computer, a control program is sequentially executed. The procedure is as described below.

1. Obtain the information of an event.
2. Execute the motion described and play the voice sound file.
3. Obtain the result of the recognition from the sensor processing that is pointed to. If voice recognition is pointed to, grammar ID is transferred to the voice recognition module. According to the result of the recognition, following ID of the ID will be decided according to diverging condition.
4. Back to 1.

The event limits the candidates of recognition and the proper sensor for the recognition. That increases the recognition accuracy of each sensor data processing part. The scenario is described in XML for non trained programmers. An example of a scenario is depicted in Figure 6. In the figure, a part marked by "<event>" and "</event>" means one event. For example, the "eventno="1"" means the following. The motion module subscribed by "G001.dat" is executed and the scenario goes to the "eventno="2"" with no sensor information. The "eventno="4"" means the following. The motion G004.dat is executed and the voice recognition program is executed based on "grammar_G004.xml". In the case that the result of voice recognition is "YES", the scenario goes to the "eventno="5"". In the case that the result is "NO" it goes to the "eventno="8"". In the case that the voice recognition failed, it goes to "eventno 16".

3. Natural idling motion and the evaluation

The ReplieeQ2 android realized a human-like presence in short-term and restricted situations (Shimada et al., 2006)(Noma et al., 2006). However, it is difficult to realize a human-like presence in various situations without a variety of reactions for changing surroundings. In this section, in order to prove the effectiveness of the system, experiments are conducted.

```
<?xml          versions.0"?>
<scenario>
          <event eventno="l" sensor="UNUSE" motion="G001.dat">
                    <resu I tXUNUSE nextevent ="2'7></resu 11 X/event > <event
          eventno="2" sensor="UNUSE" motion="G002.dat">
                    <resu I tXUNUSE nextevent = "3"/X/resu 11 X/event > <event
          eventno="3" sensor="UNUSE" motion="G003.dat">
                    <resu I tXUNUSE nextevent="4VX/resultX/event>
          <event eventno-"4" sensor-"VOICE" mot ion-"G004.dat" grammar-"grammar_G004.xml "> <resultXYES nextevent="5"/>
          <N0 next event = "8"/> <ERR0R nextevent="16"/> </resu 11 X/event > <event eventno="5" sensor="UNUSE"
          motion="G005.dat">
```

Fig. 6 An example of scenario.

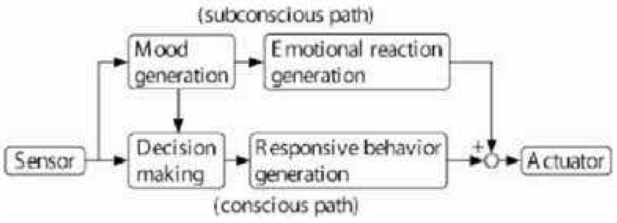


Fig. 7. A control of the android based on a mental state.

The experiments show that the human-like nature is improved by the information of a sensor network. Noma et al. evaluated the human-like nature of an android without sensor information, but sensor information could improve its human-like nature. Therefore, we prepared a situation where subjects come close to an idling android, and evaluated the impressions of the android's behavior in that situation. Two impressions of human-like presence were compared. One is the impression where the android's reactions are generated based on sensor information, and the other is the impression of idling motion without sensor information, i.e., the same as the Noma's study. The idling motion is a motion where the android stays quietly sitting and stares around or stares at something according to the surrounding situation.

This is an interaction where the active communication between a human and the robot is not established, as mentioned in chapter 1. However, it is important to achieve a natural interaction in such a situation for human-like presence for androids in our daily life. A quietly sitting person often shows a behavior in which he/she turns to a direction where some small sound is heard, but it is very difficult to achieve this behavior based on the information of a microphone. In contrast, the floor sensor allows it to detect the approach of a human and implement a motion so that the android acts as if it hears a slight sound from the human. Thus, it is necessary to observe the movement of a human's position in the distance. Before establishing active communications, it is difficult to limit the sensors and sensor data processing. A system integrated with sensor networks would improve the idling motion of the android.

3.1 Basic concept of generating idling behavior

The idling behavior to be achieved here does not contain reactions that are strongly expected, in contrast to the reaction of android in active human-robot communications. However, behaviors that are completely selected in a random manner would not give a human-like nature to the android because a quietly sitting human would show some reaction behaviors toward a change in the surroundings. Thus, a certain control model is required. For subjects to feel the human-like nature of an android, it is necessary to design the android's behavior so subjects will recognize the android's intention or mental state. For that, existing studies show that to construct mental states for the robot. For example, Miwa et al. construct a mental model integrating a conscious model (Miwa et al., 2003). This model allows a robot to clearly select a stimulus (sensor input) among several stimuli, but not randomly. Adding a reason to the selection of a stimulus among several stimuli enables a robot to show a more human-like reaction. Breazeal mentioned that the emotional expression of a robot makes interaction with humans more complex and richer (Breazeal,2003).

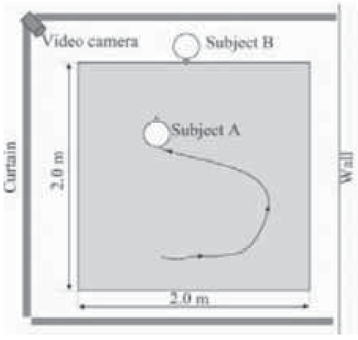


Fig. 8. Observation of a waiting person.

In this paper, we construct a motion generation model changing unconscious physical reactions according to android's mental state (Figure 7.). At first, the mental state is decided based on sensor information. The mental state is a state of emotion (e.g., fear and interest). The android choose a reaction motion based on the sensor information and the mental state. This corresponds to a conscious generation of human motion. In addition, emotional reaction is generated according to the mental state. This corresponds to unconscious emotional reaction. The emotional reactions here include frequency of blinking and change of the eye movements. Among the mental states, a neutral state is defined. The emotional reactions are expressed by difference from the neutral state (e.g., frequency of blinking, speed of the eye movement). When the mental state is neutral, one motion module is randomly selected among several motion modules prepared. As far as the mental state is neutral, this procedure will be repeated. If the mental state changes, a reaction according to the mental state is added.

The android's motion modules, mental states, emotional reactions, and idling scenario are constructed by observations of real humans and reflection reports. The motions of human without reasonable tasks should depend on individual characters. It is thinkable that extract common motions from a large number of people's motion to obtain motions independent from the characters. However, common motions are unlikely to be able to extract common motions if the motions depend on personal character. In this study, we did not research common motions of human. It is possible that results could be changed by subjects' character, but at the very least we can verify the effectiveness of the sensor network based on the models we made.

3.2 Observation of quietly sitting humans

To observe quietly sitting humans, an environment was prepared as illustrated in Figure 8. Observation was conducted in a place enclosed by curtains or walls. We took a video of the motions of a quietly sitting person and the motions of a person in his/her surrounding area by a video camera. An experimenter instructed the subject to sit in a chair and stay sitting for 20 minutes. This is subject B. Subject B knew that this experiment was a study on the motion of human-like nature in advance. The chair looks like the chair in which the ReplieeQ2 was sitting (Figure 2.). The experimenter instructed another subject to act freely for 20 minutes looking at subject B within the 2.0 m X 2.0 m area, marked by the black line Figure 8.. This was subject A. Subject A also knew that this experiment was a study on the motion of human-like nature in advance. This region of 2.0 m X 2.0 m was the same size as the floor sensor area installed in front of the android. We used two male students in age of 20's as the subject A, and two students in 20's as the subject B.

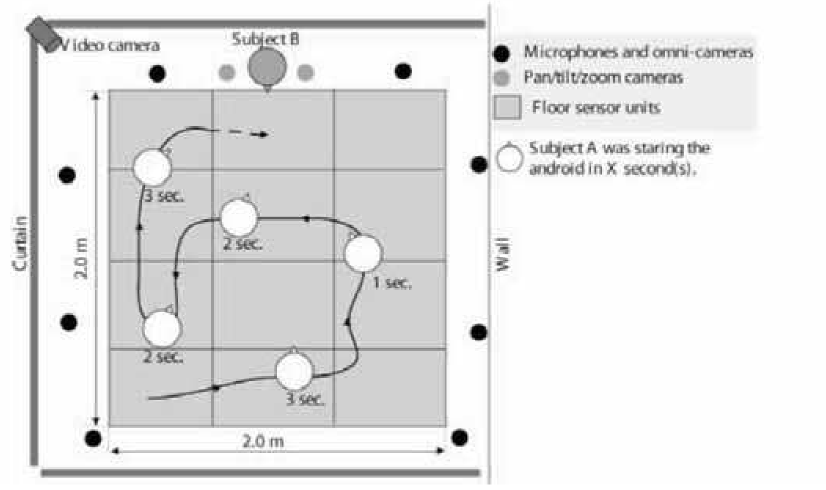


Fig. 9. Example of a behavior of the subject A.

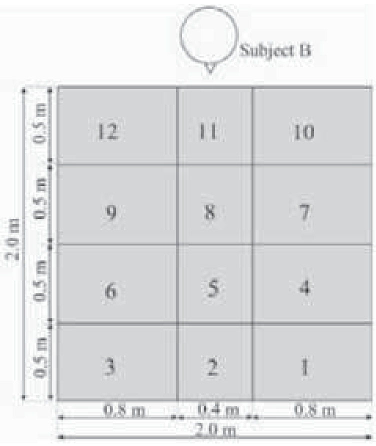


Fig. 10. ID of the position of the subjects A. Subject B was later replaced with the android in the experiment.

After the observation, subjects B gave their impressions of subjects A in order to construct a mental state model. The subjects gave as many impressions of things in these 20 minutes as they could remember.

3.3 Construction of basic motion modules

The behavior of subject A is illustrated in Figure 9. We observed the behavior in which subject B walked around subject A, sometimes stopped in a few seconds and looked at the subject A. In order to define basic motion modules, we extracted common motions frequently emerge in the motions of two subjects B. As a result of this observation, the six common motions below were obtained. These motions were defined as basic motion modules for the android, because they are frequently seen in quietly sitting humans.

- Look toward the direction subject A looks.
- Look toward subject A a few times with only eye movements and as few body movements as possible.
- Look at subject A a few times.
- Look toward the front so as not to see subject A.
- Look down to the ground.
- Keep looking at subject A (following subject A).

3.4 Construction of mental state model

The answers of the impression of subject B show that the impressions depend on the position of subject A. We thus classified the mental states based on the behavior and position of subject A. Table 1 denotes these classes. This is the mental state model. The positions of subject A are classified in Figure 10. The first column of Table 1. denotes the position IDs of subject A illustrated in Figure 10. The range "1 to 6" in this column means any number from 1 to 6. The second column denotes the behavior of subject A in the position indicated by the first column. The third column denotes the feeling of subject B for the behavior of subject A. The fourth column denotes the emotion of subject B, which is the origin of the mental state. The emotions are defined based on the emotional model of Plutchik (Plutchik, 1962). Each emotion is defined by the reasons below.

- Disgust: In the impression questionnaire, subject B reports "hate to be stared at and want it to stop" and "feeling of pressure" . We assume the disgust is caused because it is an emotion meant to remove a harmful thing.
- Acceptance: In the impression questionnaire, they report "do not care so much because subject A does not pay attention to me" . We assume the acceptance is caused because it is an emotion meant to accept the current situation without feeling comfort or discomfort.
- Fear: In the impression questionnaire, they report "feels painful to be stared at". We assume the fear is caused because it is an emotion under a situation of pain, destruction or surprise.
- Apprehension: Fear is caused because they report a feeling of discomfort to be stared at or to be in a very close place. They report that subject A has something to say. The anticipation is caused because it is an emotion under a situation of looking for the purpose. Therefore we assume the intermediate between two states of Pluichik's model, the apprehension, is caused.

- Interest: Surprise is caused because they report "caring about the direction that subject A is facing" in the impression questionnaire. It is an emotion under a situation where the other person's purpose is not known to be beneficial or not, or whether it is pleasant or not. Acceptance is caused because they tried to accept subject A's action. Therefore, we assume the intermediate between two states of Pluichik's model, interest, is caused. The mental state of a quietly sitting person is modeled based on the position of the surrounding person and his/her behavior. Employing the sensor networks we developed, this position and behaviors can be obtained more robustly than that use systems using built-in sensors on the robot body.

3.5 Construction of sub-scenarios of each mental state

The sub-scenarios for each mental state are defined based on the observation of subject B. According to the mental state, the frequency of blinking, speed of blinking and speed of eye movements changed as emotional reactions. Each value in the basic motion module is defined below.

- Frequency of blinking (BF) is once every 2-3 sec
- Speed of blinking (BF) is 0.3-0.5sec from eyes closing to opening
- Speed of eye movement is about 15 deg/sec

Sub-scenarios are explained below. The emotional reactions in each mental state are expressed by abbreviations. For example, when the case of the frequency of blinking is twice as high as the basic motion, it is described as "BFx 2".

Sub-scenario of acceptance Look at a person. BF x1, BV x1, EV x1.

Sub-scenario of fear Look at a person. BF x 3, BV x1, EV x 2.

Sub-scenario of disgust Look at a person when he/she is near (i.e., 2 and 5 in Figure 10.). Look down at the ground when a person is far (i.e., 8 and 11). BF x1, BV x 2, EV x1. **Sub-scenario of interest** Glance at a person when he/she is in a far place (i.e., 7, 9, 10 and 12). Look forward of the direction where a person looks when he/she is near (i.e., 1-6) (Two patterns: right and left). BFx 3, BVx 1, EVx 2.

Sub-scenario of apprehension Glance at a person when he/she is walking while watching the android. Look down at the ground when a person is walking in a far place (i.e., 7-12) without watching the android or when a person is looking somewhere at a far place (i.e., 8 and 11). BF x 3, BV x1, EV x 2.

3.6 Recognition of a person's behavior and decision of android's behavior

Position ID of A	Behavior of A	Feeling of B towards A	Mood of B
2,5 8 11	Stares at B Stares at B Stares at B	Unpleasant Unpleasant + Oppressive Very unpleasant + Oppressive	Disgust
1 to 6	Walks around without staring at B	Not so concerned about the action of A	Acceptance
1, 3, 4, 6	Stares at B	Harassed	Fear
7 to 12 8,11 1 to 12	Walks around without staring at B Looks at somewhere Walks around with staring at B	Anxiety Anxiety Anxiety	Apprehension
1 to7,9, 10,12	Looks at somewhere	Worry about where A is looking	Interest

Table 1. Classification of the feelings of subjects B toward subjects A. Subject B was later replaced with the android in the experiment.

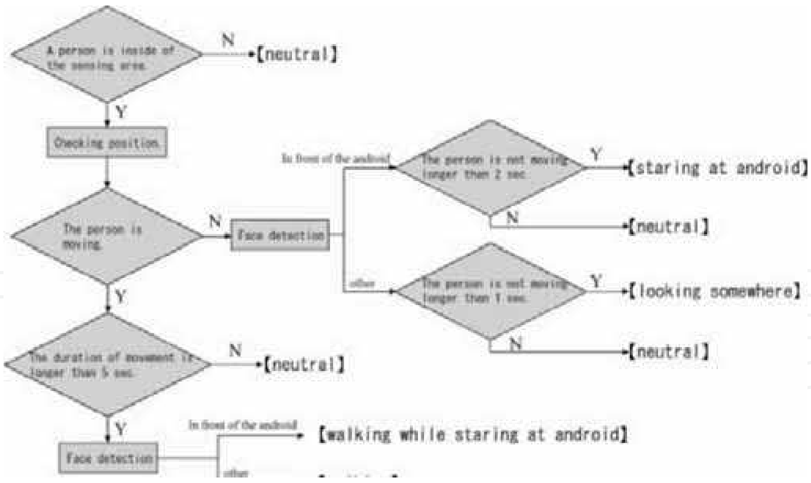


Fig. 11. Procedure for recognizing a person's behavior.

The scenario of the android's behavior in a quietly sitting state is constructed based on the mental state model and sub-scenarios we constructed. The process for recognizing the behavior of subject A in Table 1 is illustrated in Figure 11. For example, first, when a person is in the floor sensor region, his/her position is measured by the floor sensor.

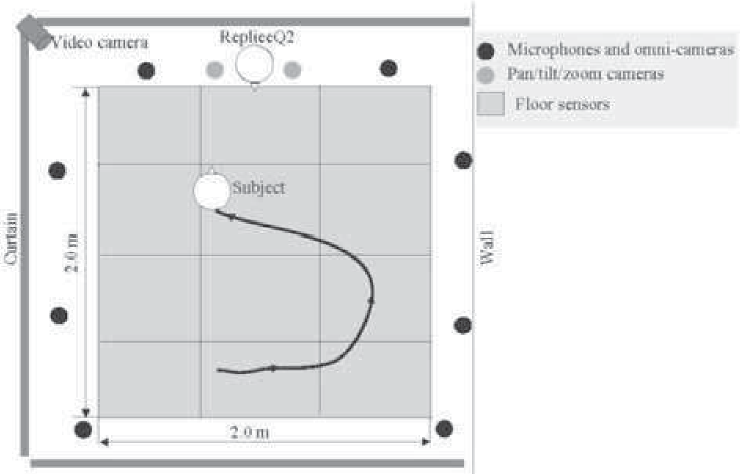


Fig. 12. Arrangement of the sensor network.

Next, whether he/she is moving or not is determined from time-series position data. When he/she is not moving, his/her face direction is measured by pan-tilt cameras. When his/her face is pointed toward the android, whether he/she is looking at the android for more than 2 sec. from the time-series data for the face direction. When he/she is looking at the android for more than 2 sec., he/she is assumed to be staring at the android. The mental state of the android is decided based on the behavior and position of a person. As mentioned in 2.2, a recognition of the situation (i.e., in this case a recognition of the person's behavior) is executed in each module. When the behavior of a person cannot be recognized, the mental state is assumed to be neutral (Figure 11). When a mental state is caused, a sub-scenario is executed accordingly. After each sub-scenario is executed, the mental state is neutral in some steps without shift to other state in order to avoid same motion is executed in sub-scenario.

Motion modules are manually made with motion-making software. Thus, the idling motion scenario is made by describing these procedures in XML.

3.7 The evaluation experiment

3.7.1 Purpose of examination

An experiment for evaluating the effectiveness of the idling scenario based on the sensor networks was conducted. The subjects evaluated the subjective human-like nature when took action around the android in an idling state. We also placed the android in idling motion without sensor information to compare. In this idling state, the android chose and repeated random motions among 15 kinds of motion modules, including "do nothing particular", "look around" and "sleepy motion". In this experiment, each subject evaluated the human-like nature of the android in both conditions. We predicted that the android that chose reactions based on the data of sensor networks would have a more human-like nature even in the situation where there was communication between the human and the android.

3.7.2 Experimental procedure

The experiment was conducted in the environment illustrated in Figure 12.. Each sensor was arranged depicted in this figure. Each subject was in the two conditions below. **Random condition:** The subject acted for 3 minutes near the android with randomly selected motions.

Sensor condition: The subject acted for 3 minutes near the android with motions selected by scenario.

The subjects were instructed before the experiment started: they were to act freely around the android and evaluate how human-like the android was. The subjects could also touch and talk to it when they acted. After participating the experiments in these two conditions, the subjects evaluate human-like nature of the android in each condition. The evaluation was on a scale from -3 to 3. Free-answer questionnaire about the impression of the android was also conducted. The order of the two conditions was changed for each subject. This was to counterbalance the order of the conditions. There was a one minute interval between the first condition and the second condition. The subjects were 16 male and female university students.

3.7.3 Result of experiment

Figure 13. shows the experimental scene. Each subject only walked around the android, and did not touch or talk to the android. The behavior exhibited by most of the subjects was to walk close to the android or to stay at a distance from it and they looked to be interested (Figure 9.). Some subjects fixed their eyes on the android almost without moving. In addition, some of them waved their hand or extended their hands toward the android's face. They appeared to be trying to confirm whether or not the android could recognize them.

Next, we verified a hypothesis that the android in the sensor condition is more humanlike than the android in the random condition. In this experiment, we compared the evaluation values of human-like nature by testing only the effect the order that the viewing condition had on the subjects.

In this experiment, the subjects were divided into two groups; the first group was first presented with the sensor condition (S-R group, 8 persons) and the second group was first

Fig. 13. Experimental scene.

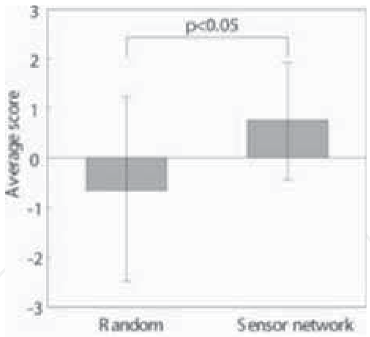


Fig. 14. Evaluation of human-like nature of the android(between-subject comparison).

presented with the random condition (R-S group, 8 persons). We compare the evaluation value of the S-R group and the R-S group. We compare the evaluation value of S-R group and R-S group. That is, we use only values of first presented condition. Figure 14. depicted the average of each condition. The error bar indicates standard deviation of each evaluation value. The independent t-test on the difference of average values shows significant difference (t = 1.781, degree of freedom 14, p = 0.0483). This result shows that the reactions based on sensor information gave the android a more human-like nature. In the free-answer questionnaire, the main reasons for the human-like nature of the android's motions were as given below.

- The motion of turning its gaze toward a person gave a human-like impression.
- When the subjects came close to the android, they felt that the android was nervous as it turned its gaze around. (This may indicate a quickening of the eye movement.)

When the subjects came close to the android, the frequency of blinking was augmented. This gave them a human-like impression.

However, some subjects mentioned these reasons as being motions which they deemed to be non human-like. This questionnaire indicates that the subjects detected changes in the emotional reaction. In addition, they used the word "nervous" and they attributed this change in emotional reaction to the change of mental state. This also indicates that the change of emotional reaction gave a human-like nature to the android.

3.8 Discussion

Only the floor sensor and pan-tilt cameras were used in this experiment, because the android's scenarios of the android were constructed based on the observation of quietly sitting real humans. As such, this experiment is not an evaluation of the entire system. However, we could verify that a quite different system from human perception can realize human-like behavior in the android. The gesture recognition with omni-cameras and the voice recognition with microphone described in section 2.3 shown made sure to function, but the impression evaluations with all of these functions are issues in the future. We built a mental state model based on the observation of two subjects in this experiment but this is only one example of a mental state model. Sensor networks allow us to realize human-like reaction in the android, but more studies are required to see: which relationships are between elements (e.g., sensor-mental state, mental state-behavior) affect the human-like nature. This mental state model should be more objective because it is constructed based on observations of the experimenter. As some methods for construct human emotion models have been proposed (e.g., (Bianchi-Berthouze & Lisetti, (2002)), it is possible to construct a more objective mental state model using these methods. In this experiment, different motion modules were used in the sensor condition and the random condition. Therefore, the differences in the evaluation of the impression for each condition were possibly not produced by reactions based on sensor information, they were possibly differences in the evaluation of the modules themselves. In other words, the motion modules in the random condition possibly obtained a low score. But we believe that the humanlike nature of these motion modules themselves were not low because they were constructed by combining motions for other studies (Shimada et al., 2006)(Noma et al., 2006). However, to clarify this point, it is necessary to conduct experiments to compare the sensor and random conditions in the future.

In the scenarios of the idling state of this experiment, human positions are detected by the floor sensor. The floor sensor is quite different from human modality and obtain the positions more precisely. The obtained data from the floor sensor were translated into the 12 level resolution and we believe that the motions of "see" of the android made the subjects believe it knew the human positions. However, suitable resolution and manner of information translation have not been revealed, studies of this issue are still necessary. In this experiment, we built the motion generation model based on the mental state aiming to make subjects attribute the motion of the android to a mental state, including the intention. Several studies aim to attribute motions of non humanoid robot to mental states to realize natural communication with human (Terada et al., 2007)(Kobayashi & Yamada, 2007). In these studies, it is possible that robots are anthropomorphized by the motions attributed to mental state. In contrast, the android would be anthropomorphized and attributed to the mental state, before observing the motions of the android because the android has a very humanlike appearance. It is possible that the appearance may promote natural communication. It is necessary to perform more study comparing the android to other robots to clarify such a supposition.

4. Conclusion

This paper proposed a system integrated an android and sensor networks in order to achieve a natural communication with human keeping human-like nature of the android. An android control system and computer networks for sensor processing of a large number of various sensors are integrated into a system. In addition, a software system developed, which achieves to obtain of sensor information and description of android's behavior in XML. Existing study already indicated that human-like appearance and motion of android causes human-like impression. In addition, reaction according to situation is required in order to cause more human-like presence in more complicated interaction. This study achieved to construct the world's first humanoid robot system which can achieve a human-like nature integrating the appearance, the movement and the perception. Configuring idling behavior of the android as an example issue indicated the developed system is effective to achieve a subjective android's human-like nature.

In order to achieve a subjective android's human-like nature in communication, it is not required to implement abilities equivalent to human including abilities to exercise and abilities to percept. The result of experiment in this paper suggests that reactions for human-like presence are efficient to achieve the subjective android's human-like nature in communication. Studies on minimum required abilities to exercise and percept for achieving human-like nature, i.e. a boundary condition of natural interaction should conduce a principle of natural communication. The system we built in this paper allow us to tackle issues focused on appearance, motion and perception simultaneously, such studies have not been started yet.

5. Acknowledgement

We developed the android ReplieeQ2 in collaboration with Kokoro Company, Ltd..

6. References

Kanda, T.; Ishiguro, H.; Ono, T.; Imai, M.; Maeda, T.; & Nakatsu, R. (2004). Development of robovie as a platform for everyday-robot research, *Electronics and Communications in Japan,* 87(4).

Kajita, S (2002). Research of biped robot and humanoid robotics project (HRP) in Japan, *The Fourth International Conference onMachine Automation (ICMA '02),* pp. 1-8.

Daibo, I. (1993). Psychology of intimacy and function of communication. *Technical Report of IEICE (The Institute of Electronics, Information and Communication Engineers),* HC-93-52:33-40, (In Japanese with English abstract).

Duffy, B. R. (2003). Anthropomorphism and the social robot. *Robotics and Autonomous Systems,* 42(3-4),177-190.

Ishiguro, H. (2005). Android science - toward a new cross - interdisciplinary framework.*Proceedings of the International Symposium of Robotics Research.*

Shimada, M.; Minato, T.; Itakura, S. & Ishiguro, H. (2006). Evaluation of android usingunconscious recognition, *Proceedings of the IEEE-RAS International Conference onHumanoid Robots,* pp. 157-162.

Noma, M.; Saiwaki, N.; Itakura, S. & Ishiguro, H. (2006). Composition and evaluation of the humanlike motions of an android, *Proceedings of the IEEE-RAS International Conference on Humanoid Robots,* pp. 163-168.

Okuno, H. G. ; Nakadai, K. ; Lourens, T. & Kitano, H. (2001). Human-robot interaction through real-time auditory and visual multiple-talker tracking, *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS-2001),* pp. 1402-1409.

Morishita, H.; Watanabe, K.; Kuroiwa, T.; Mori, T. & Sato, T. (2003). Development of robotic kitchen counter: A kitchen counter equipped with sensors and actuator for action-adapted and peronally-fit assistance, *Proceedings of the 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems,* pp. 1839-1844.

Mori, T.; Noguchi, H. & Sato, T. (2005). Daily life experience reservoir and epitomization with sensing room, *Proceedings of Workshop on Network Robot Systems: Toward Intelligent Robotic Systems Integrated with Environments.*

Ishiguro, H. (1997). Distributed vision system: A perceptual information infrastructure for robot navigation, *International Joint Conference on Artificial Intelligence (IJCAI-97),* pp. 36-41.

Ishiguro, H. (2007). Scientific issues concerning androids, *International Journal of Robotics Research,* 26(1),105-117.

Murakita, T.; Ikeda, T. & Ishiguro, H. (2004). Human tracking using floor sensors based on the markov chain monte carlo method, *Proceedings of the 7th International Conference on Pattern Recognition,* pages 917-920.

Nishimura, T.; SoGo, T.; Ogi, S.; Oka, R. & Ishiguro, H. (2001). Recognition of human motion behaviors using view-based aspect model based on motion change, *Transactions of the IEICE,* J84-DII(10):2212-2223.

Ishiguro, H. & Nishimura, T. (2001). Vambam: View and motion based aspect models for distributed omnidirectional vision systems, *Proceedings of the International Joint Conference on Artificial Intelligence,* pages 1375- 1380.

Adachi, Y.; Imai, A.; Ozaki, M. & Ishii, N. (2001). Study on extraction of flesh-colored region and detection of face direction, *Int.J. Knowledge-Base Intelligent Eng. Syst.,* 5-2.

Miwa, H.; Itoh, K.; Takanobu, H. & Takanishi, A. (2003). Construction of mental model of humanoid robot for natural communication with human, *Proceedings of the 18th Meeting of Special Interest Group on AI Challenges,* pages 39- 44.

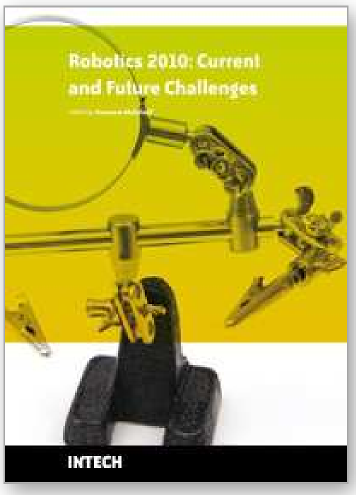
Breazeal, C. (2003). Emotion and sociable humanoid robots, *International Journal of Human-Computer Studies,* 59(1-2):119-155.

Plutchik, R. (1962). *The emotion: facts, theories and new model.* Random House, New York.

Bianchi-Berthouze, N. & Lisetti, C. L. (2002). Modeling multimodal expression of user's affective subjective experience, *User Modeling and User-Adapted Interaction,* 12:49-84.

Terada, K.; Shamoto, T.;Mei,H. & Ito, A. (2007). Reactive Movements of Non-humanoid Robots Cause Intention Attribution in Humans. *Proc. of IEEE/RSJ International Conference on Intelligent Robots and Systems 2007 (IROS 2007).*

Kobayashi, K. & Yamada, S. (2007) Motion Overlap for a Mobile Robot to Express its Mind, *Journal of Advanced Computational Intelligence and Intelligent Informatics,* Vol.11, No.8, pp. 964-971.



Robotics 2010 Current and Future Challenges

Edited by Houssem Abdellatif

ISBN 978-953-7619-78-7
Hard cover, 494 pages
Publisher InTech

Published in print edition February, 2010

Without a doubt, robotics has made an incredible progress over the last decades. The vision of developing, designing and creating technical systems that help humans to achieve hard and complex tasks, has intelligently led to an incredible variety of solutions. There are barely technical fields that could exhibit more interdisciplinary interconnections like robotics. This fact is generated by highly complex challenges imposed by robotic systems, especially the requirement on intelligent and autonomous operation. This book tries to give an insight into the evolutionary process that takes place in robotics. It provides articles covering a wide range of this exciting area. The progress of technical challenges and concepts may illuminate the relationship between developments that seem to be completely different at first sight. The robotics remains an exciting scientific and engineering field. The community looks optimistically ahead and also looks forward for the future challenges and new development.

Takenobu Chikaraishi, Takashi Minato and Hiroshi Ishiguro (2010). Development of an Android System Integrated with Sensor Networks, Robotics 2010 Current and Future Challenges, Houssem Abdellatif (Ed.), ISBN: 978-953-7619-78-7, InTech, Available from: http://www.intechopen.com/books/robotics-2010-current-and-future-challenges/development-of-an-android-system-integrated-with-sensor-networks

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821